Understanding the Dark Universe

Everything you always wanted to know about the cosmological constant problem (but were afraid to ask)

Tout ce que vous avez toujours voulu savoir sur le problème de la constante cosmologique sans jamais avoir osé le demander

Jérôme Martin

Institut d'astrophysique de Paris, UMR7095-CNRS, université Pierre-et-Marie-Curie, 98 bis, boulevard Arago, 75014 Paris, France

A B S T R A C T

This article aims at discussing the cosmological constant problem at a pedagogical but fully technical level. We review how the vacuum energy can be regularized in flat and curved space–time and how it can be understood in terms of Feynman bubble diagrams. In particular, we show that the properly renormalized value of the zero-point energy density today (for a free theory) is in fact far from being 122 orders of magnitude larger than the critical energy density, as often quoted in the literature. We mainly consider the case of scalar fields but also treat the cases of fermions and gauge bosons which allows us to discuss the question of vacuum energy in super-symmetry. Then, we discuss how the cosmological constant can be measured in cosmology and constrained with experiments such as measurements of planet orbits in our solar system or atomic spectra. We also review why the Lamb shift and the Casimir effect seem to indicate that the quantum zero-point fluctuations are not an artifact of the quantum field theory formalism. We investigate how experiments on the universality of free fall can constrain the gravitational properties of vacuum energy and we discuss the status of the weak equivalence principle in quantum mechanics, in particular the Colella, Overhauser and Werner experiment and the quantum Galileo experiment performed with a Salecker–Wigner–Peres clock. Finally, we briefly conclude with a discussion on the solutions to the cosmological constant problem that have been proposed so far.

© 2012 Published by Elsevier Masson SAS on behalf of Académie des sciences.

R É S U M É

Cet article traite du problème de la constante cosmologique d'une manière pédagogique aussi bien que technique. Nous passons en revue comment l'énergie du vide peut être régularisée en espace temps plat et courbe, et comment son calcul peut être compris en termes de diagrammes de Feynman en bulles. En particulier, nous montrons que la valeur correctement renormalisée de l'énergie de point zéro maintenant (pour une théorie libre) en en fait bien au dessous des 122 ordres de grandeur de l'énergie critique de l'univers. Nous nous concentrons sur le cas des champs scalaires bien que nous considérions le cas des fermions et des bosons de jauge également afin de traiter l'énergie du vide en supersymétrie. Nous discutons aussi comment la constante cosmologique peut être mesurée cosmologiquement et contrainte par la mesure de l'orbite des planètes du système solaire ou encore des spectres atomiques. Nous passons aussi en revue pourquoi la mesure du Lamb shift et de l'effet Casimir semblent indiquer que les fluctuations quantiques de

E-mail address: jmartin@iap.fr.

l'énergie du vide ne sont pas un artefact du formalisme de la théorie des champs. Nous montrons ensuite comment les expériences sur l'universalité de la chute libre peuvent aider à contraindre les propriétés gravitationnelles de l'énergie du vide et nous discutons le statut du principe d'équivalence faible en mécanique quantique, en particulier l'expérience de Colella, Overhauser et Werner ainsi que l'expérience de Galilée quantique faite avec une horloge de Salecker–Wigner–Peres. Enfin, nous concluons brièvement avec une discussion sur les différentes solutions au problème de la constante cosmologique qui ont été proposées jusqu'à présent.

© 2012 Published by Elsevier Masson SAS on behalf of Académie des sciences.

1. Introduction

This review article aims at discussing the cosmological constant problem [1–15]. This question is central in modern physics because its resolution would certainly mean a very important step forward in our attempts to understand physics beyond the current standard model.

The history of the cosmological constant problem is a long and rich one [5]. Its premises were in fact already present immediately after the birth of quantum field theory. Indeed, from the Heisenberg uncertainty principle, we know that the ground state energy of the quantum mechanical oscillator cannot be zero because the potential and the kinetic energies cannot vanish at the same time. Since a (free) quantum field can be viewed as an infinite collection of harmonic oscillators, it immediately comes that its ground state energy must be infinite. Of course, this is not yet the cosmological constant problem because gravity does not enter the stage. But, clearly, potentially, a very severe problem is already present.

The infinity mentioned above is the first infinity that one encounters in quantum field theory [16–22]. However, the presence of this infinity did not prevent the founding fathers of quantum field theory to develop the theory since, in absence of gravity (a working assumption of quantum field theory justified by the weakness of gravity), only differences in energies are observable. Therefore, in this context, this infinity can be merely ignored and the rest of the formalism can be worked out without any problem. As is well known, other types of infinities appear and, this time, there is no way to get around them. Treating and taming them is the goal of renormalization. The very impressive agreement between high precision measurements in accelerators and the predictions of quantum field theory in presence of radiative corrections is the proof that renormalization is able to correctly regulate these infinities.

Therefore, we see that these two types of infinities are treated very differently. The problem of the zero-point energy is just avoided while the problem of the radiative corrections is directly and explicitly addressed. Clearly, one cannot help thinking that the first problem is in fact swept under the carpet. And, indeed, as soon as gravity is turned on, it badly strikes back.

At this point, it is worth noticing the following. One should not get the impression that the zero-point energy cannot be renormalized. As a matter of fact, as will be discussed in this review, it can be made perfectly finite. However, this finite, renormalized, value of the zero-point energy seems to be too large to be compatible with the observations. Again, this is very different from the usual case of quantum field theory where the finite part extracted from a divergent expression always leads to a good agreement with the experiments. At first sight, the cosmological constant problem is therefore neither the presence of a new infinity nor our inability to regularize it but rather the apparent failure of the renormalization scheme to produce, at the quantitative level, a finite vacuum energy compatible with the observational data.

In fact, a failure of renormalization is not the only logical possibility. The vacuum fluctuations could also be a mathematical artifact of the quantum field theory framework and/or their gravitational properties could be abnormal. However, when one tries to test (theoretically or experimentally) these ideas, it seems that one never encounters a problem of the kind mentioned above. This fact must also be considered as a part of the cosmological constant problem which, therefore, appears to have many different ramifications. We see that this problem is in fact a very deep problem which lies at the cross roads between different branches of physics (gravitational physics, quantum field theory, cosmology etc., . . .). In brief, it has to do with the gravitational properties of the quantum vacuum.

The importance of this question has recently been reinforced by the discovery that the expansion of our Universe is accelerated [23,24]. According to standard lore, this could be the first observational evidence that the quantum vacuum is able to curve space–time. On the other hand, the effect of a cosmological constant is visible only on large scales and, therefore, it seems to be problematic to check this result elsewhere than in cosmology. But, if true, the cosmological constant problem represents a unique situation where some aspects of quantum gravity are at play and where, at the same time, corresponding observational signatures are not hopelessly beyond our technical capabilities. This makes it a valuable opportunity to go beyond our current understanding of theoretical physics. The only other situation where the three above mentioned aspects (quantum mechanics, gravity and the possibility to detect sizable observational effects) are mixed is the theory of cosmological inflation [25–34] (for a review, see Refs. [35–37]) where the Cosmic Microwave Background (CMB) anisotropies, first detected by the COBE (COsmic Background Explorer) satellite [38], are supposed to originate from the quantum fluctuations of the inflation and gravitational fields (for the observational status of inflation, see Ref. [39]).

In this review article, we go through all the issues mentioned before. Our prejudice is to present the question at the technical level, in full details, in order for the article to be reasonably self-contained. The price to pay for this approach

is that, sometimes, we are led to consider problems that are well treated in standard textbooks and/or are not, strictly speaking, directly concerned with the cosmological constant problem. On the other hand, this is also an opportunity to review many different subjects belonging to different types of physics.

The outline of the article is the following one. In Section 2, we discuss how the cosmological constant Λ is introduced in the Einstein field equations and we explain why the vacuum energy also participates in the value of Λ. In Section 3, we treat the classical cosmological constant problem, i.e. we show that, in the presence of gravity, it is no longer true that only differences of energy are observable. We illustrate this discussion by means of the electro-weak phase transition. In Section 4, we demonstrate that, even if the classical ground state is tuned to zero, the quantum vacuum fluctuations still give a contribution to the cosmological constant. We discuss how the corresponding quantum vacuum energy can be calculated and argue that the method often used in the literature and which consists in introducing a cut-off, is not appropriate. We also investigate how the quantum vacuum energy can be expressed in terms of Feynman diagrams (bubble diagrams) and investigate the case where interactions are present. In Section 5, we focus on the so-called bubble diagrams. We discuss them in the context of quantum field theory but also in ordinary quantum mechanics. We then show that they have very different properties than the other loop diagrams. In Section 6, we use the Gaussian effective potential approach to calculate the vacuum energy in a situation where non-perturbative effects are present. Since all the calculations are carried out in the case of a scalar field, we treat in Section 7 the case of other types of fields (spinor and vector fields). This also gives us the opportunity to discuss the idea of super-symmetry. In Section 8, we present the cosmological constant problem in a more rigorous way. In particular, we show that a flat space–time calculation is a good approximation although a fully consistent approach should be formulated in curved space–time. One concludes the first part of this review article with Section 9 where we estimate the vacuum energy and find a value very far from the often quoted "122 orders of magnitude". In Section 10, we explain, from a theoretical point of view, how the value of the cosmological constant was recently measured in cosmology. In particular, we discuss the hypothesis that are implicitly assumed in order to obtain this result. In Section 11, we discuss whether the cosmological constant can be measured elsewhere than in cosmology and argue that no other experimental context can compete with cosmology. However, we also show that other experiments can put constraints on the vacuum energy that are interesting from the point of view of the cosmological constant problem. In Section 12, we review the experiments (Lamb shift, Casimir effect) that seem to prove the existence of the vacuum fluctuations. In Section 13, we investigate whether there are observations that can probe their gravitational properties. In Section 14, we discuss whether the weak equivalence principle still holds in quantum mechanics since this has obvious implications for the question of the vacuum weight. Finally, in Section 15, we present our conclusions and very briefly review the solutions to the cosmological constant problem that have been proposed so far. Let us mention that this article is part of a more general review on the dark energy question, see Refs. [40–43].

Before concluding this introduction, we would like to make the following remark. This article aims at discussing what the cosmological constant problem is. This article does not aim at discussing what solutions to this problem have been proposed even if, as already mentioned, in Section 15, we say a few words on this topic (and choose, in a totally arbitrary way which only reflects the author prejudices, to focus more on some of them). The justification for the above mentioned choice is that the question of the gravitational properties of vacuum zero-point fluctuations is already a highly non-trivial one as the variety of the subjects listed above demonstrate. Therefore, it seems to us better to deeply understand what is at play before trying to propose a way out. Of course, it is because this problem is very difficult and has many different ramifications in many different branches of physics that it is so interesting.

2. The cosmological constant

In this first section, we introduce the cosmological constant as a free parameter in the classical action of the gravitational field. This action, together with the action describing matter, can be expressed as

$$S = \frac{1}{2\kappa} \int d^4x \sqrt{-g}(R - 2\Lambda_{\rm B}) + S_{\rm matter}[g_{\mu\nu}, \Psi] \tag{1}$$

where $\kappa \equiv 8\pi G/c^4 \equiv 8\pi/m_{\rm Pl}^2 \equiv 1/M_{\rm Pl}^2$, $m_{\rm Pl}$ and $M_{\rm Pl}$ being the Planck mass and the reduced Planck mass, respectively. The first term is the standard Einstein–Hilbert action and represents the gravitational part of the total action given by Eq. (1). The cosmological constant appears in the second term of the above expression. As announced, at this level, it is merely a new parameter of the total action and for this reason we write it as $\Lambda_{\rm B}$ for "bare cosmological constant". It has the dimension of the inverse of a square length. It is compatible with general covariance and is of course compatible with a conserved energy momentum tensor since $\nabla^\mu g_{\mu\nu} = 0$. Therefore, this term appears to be totally natural from the relativistic point of view and there is a priori no reason to discard it. Since, according to the standard lore of field theory, everything which is not forbidden should be considered, the cosmological constant should clearly be included in our description of the gravitational field. Finally, the third term in the above equation denotes the matter action where Ψ represents a generic matter field that we do not need to specify at this stage. Variation of the total action with respect to the metric tensor leads to the Einstein equations of motion which read

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda_{\rm B}g_{\mu\nu} = \kappa T_{\mu\nu} \tag{2}$$

where the stress–energy tensor is defined by

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta S_{\text{matter}}}{\delta g^{\mu\nu}} \tag{3}$$

As already mentioned, Λ_{B} appears to be just a parameter of the model. At this level, the only thing that can be done is to try to constrain it using various observations. But, clearly, in the present framework, there is no way to calculate its value from more fundamental considerations.

However, as originally shown by Sakharov [44], the nature of the discussion is crucially changed when one takes into account quantum field theory. The point is that the stress–energy tensor of a field placed in the vacuum state must be given by

$$\langle 0|T_{\mu\nu}|0\rangle = -\rho_{\text{vac}} g_{\mu\nu} \tag{4}$$

where ρ_{vac} is the constant energy density of the vacuum. This equation is valid for all the fields present in the Universe, as will be shown below. There are several ways to prove the above result. Firstly, one can use the fact that, in flat space–time (in Minkowski space–time), the only invariant tensor is $\eta_{\mu\nu}$. Since the vacuum state must be the same for all observers, one necessarily has $\langle T_{\mu\nu}\rangle \propto \eta_{\mu\nu}$. In curved space–time, this means that

$$\langle T_{\mu\nu}\rangle = -\rho_{\text{vac}}(t, \mathbf{x}) g_{\mu\nu} \tag{5}$$

and from the fact that the stress–energy tensor must be conserved, we reach the conclusion that ρ_{vac} must be a constant. Therefore, one obtains the formula shown above.

Another way to obtain the same result is to consider a specific example, for instance a scalar field Φ. The corresponding action reads

$$S_\Phi = -\int \mathrm{d}^4 x \sqrt{-g} \left[\frac{1}{2} g^{\mu\nu} \partial_\mu \Phi \partial_\nu \Phi + V(\Phi) \right] \tag{6}$$

where $V(\Phi)$ is the potential. Using the definition (3), the corresponding stress–energy tensor can be written explicitly as

$$T_{\mu\nu} = \partial_\mu \Phi \partial_\nu \Phi - g_{\mu\nu} \left[\frac{1}{2} g^{\alpha\beta} \partial_\alpha \Phi \partial_\beta \Phi + V(\Phi) \right] \tag{7}$$

From this expression, one sees that the scalar field is in fact a perfect fluid. Now, the vacuum state is the minimum energy state. Clearly, in order to minimize the energy we have to consider a situation where the kinetic energy vanishes and where the field sits at the minimum of its potential. In this case, the stress–energy tensor reduces to

$$\langle T_{\mu\nu}\rangle = -V(\Phi_{\text{min}}) g_{\mu\nu} \tag{8}$$

which has exactly the expected form with $\rho_{\text{vac}} = V(\Phi_{\text{min}})$.

In fact what was shown before is that there are at least two sources for the vacuum energy. There is a "classical" contribution given by Eq. (8) which originates from the value of the potential at its minimum. The corresponding vacuum energy will be calculated in Section 3. There is also a "quantum-mechanical" source given by Eq. (5) which originates from the zero point fluctuations of the ground state. This problem will be treated in Section 4.

Having established the form of the stress–energy–momentum tensor in the vacuum, one can now proceed with the Sakharov argument. The next step consists in assuming that the equivalence principle applies to the zero-point fluctuations (here, and from now on, we mean to the "classical" and "quantum-mechanical" contributions as discussed in the last paragraph), that is to say that the zero-point fluctuations gravitate. Clearly, this is not a trivial step but, after all, the vacuum fluctuations are just a specific type of energy and, in general relativity, all forms of energy gravitate. Therefore, a consistent way of writing the Einstein equations when quantum field theory is taken into account seems to be

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda_{\text{B}} g_{\mu\nu} = \kappa T_{\mu\nu}^{\text{matter}} + \kappa \langle T_{\mu\nu}\rangle \tag{9}$$

where on the right-hand side the first contribution comes from ordinary matter while the second one represents the contribution originating from the vacuum (of course, it is to be understood that we sum up the contributions coming form all the fields present in the Universe). Using the form of $\langle T_{\mu\nu}\rangle$ established before, one arrives at

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda_{\text{eff}} g_{\mu\nu} = \kappa T_{\mu\nu}^{\text{matter}} \tag{10}$$

where

$$\Lambda_{\text{eff}} = \Lambda_{\text{B}} + \kappa \rho_{\text{vac}} \tag{11}$$

Therefore, we conclude that the effective cosmological constant is the sum of the bare cosmological constant and of a contribution originating from the vacuum fluctuations. The effective cosmological constant Λ_{eff} is the quantity that one can observe and constrain when tests of the Einstein equations are carried out.

As we now discuss, the problem is that $\kappa \rho_{\text{vac}}$ is made of several terms which are all huge in comparison with the observed value of Λ_{eff}.

3. The classical cosmological constant problem

The cosmological constant problem is in fact a multi facets question as the quantity Λ_{eff} receives contributions from different origins. As was demonstrated in the last section on the example of a scalar field, a classical contribution to Λ_{eff} comes from the value of the potential at its minimum. If this is not zero, then it affects the value of the effective cosmological constant. This point is discussed in this section and is exemplified with the electroweak transition. Another contribution comes from the quantum zero point fluctuations but this will be discussed in the next sections.

3.1. Phase transition

It is possible to study the dynamics of a phase transition with the help of the following simple model. Let us consider a scalar field Φ in interaction with another scalar field Ψ such that

$$V(\Phi, \Psi) = V(\Phi) + \frac{\bar{g}}{2}\Phi^2\Psi^2 \tag{12}$$

where $\bar{g}$ is a dimensionless coupling constant and where the self-interacting potential is given by

$$V(\Phi) = V_0 + \frac{\lambda}{4}(\Phi^2 - v^2)^2 \tag{13}$$

λ being a coupling constant describing the self-interaction and v the value of the scalar field at the minimum of its potential in absence of an interaction with Ψ. The quantity V_0 denotes the classical off-set. If the field Ψ is in thermal equilibrium, then one is entitled to replace Ψ^2 with $\langle\Psi^2\rangle_T$ where the average is taken in a thermal state with temperature T. Since $\langle\Psi^2\rangle_T \propto T^2$ [45], the effective potential becomes

$$V_{\text{eff}}(\Phi) = V_0 + \frac{\lambda}{4}(\Phi^2 - v^2)^2 + \frac{\bar{g}}{2}T^2\Phi^2 \tag{14}$$

where we have slightly redefined the coupling constant $\bar{g}$ in order to take into account the proportionality constant between the thermal average and the temperature. This potential can also be expressed as

$$V_{\text{eff}}(\Phi) = V_0 + \frac{\lambda v^4}{4} + \frac{\lambda v^2}{2}\left(\frac{T^2}{T_{\text{cri}}^2} - 1\right)\Phi^2 + \frac{\lambda}{4}\Phi^4 \tag{15}$$

where we have defined $T_{\text{cri}} = v\sqrt{\lambda/\bar{g}}$. One sees on the above equation that the interaction with the field Ψ gives a temperature dependence to the effective mass of the Φ field, namely

$$m_{\text{eff}}^2(T) \equiv \lambda v^2\left(\frac{T^2}{T_{\text{cri}}^2} - 1\right) \tag{16}$$

As a consequence, when $T > T_{\text{cri}}$, that is to say before the transition, the square of the effective mass is positive while, after the transition, when $T < T_{\text{cri}}$, it becomes negative. In the first situation, the minimum is located at $\Phi = 0$ and the corresponding value of the vacuum energy is $V_0 + \lambda v^4/4$, see Fig. 1. In the second situation, the minimum is located at $\Phi = v$ and the off-set is simply given by V_0, see Fig. 1. We now clearly see the problem. If we require the vacuum energy to vanish before the transition, we must choose $V_0 = -\lambda v^4/4$. But, then, after the transition, the vacuum energy is no longer zero and is given by the negative value $\rho_{\text{vac}} = -\lambda v^4/4$. On the other hand, one can choose ρ_{vac} to vanish after the transition. In this case, one must choose $V_0 = 0$. But this means that it was not zero before the transition, $\rho_{\text{vac}} = \lambda v^4/4$. This last option can maybe viewed as the preferred solution since we do not have direct observational constraints on the matter content of the Universe prior to Big Bang Nucleosynthesis (BBN). Anyhow, we see that the choice of a parameter in the potential allows us to change the vacuum energy at the classical level, hence the name "classical cosmological constant problem". But, clearly, the off-set cannot be zero before and after the transition. Of course, in order to really establish that there is a problem, the question remains to estimate this vacuum energy for realistic phase transition and to compare it with the observational constraints. In the next subsection, we turn to the first issue and consider the case of the electroweak phase transition.
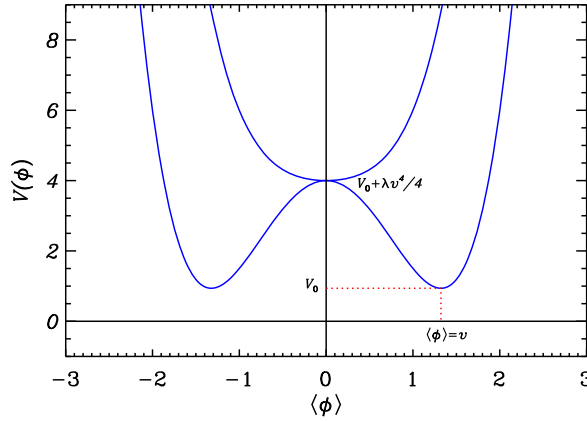
Fig. 1. The effective potential given by Eq. (14). Before the transition, for $T > T_{\mathrm{cri}}$, the minimum of the potential is located at the origin and the vacuum energy is given by $V_0 + \lambda v^4/4$. After the transition, for $T < T_{\mathrm{cri}}$, the minimum is located at $\Phi = v$ and the corresponding vacuum energy has changed and now equals V_0. It is clear that V_0 can always be chosen such that the vacuum energy vanishes either before or after the transition. It is equally clear that one cannot choose the parameters of the potential such that ρ_{vac} is zero before and after the phase transition.

3.2. The electroweak phase transition

In this section, we calculate the vacuum energy "induced" by the electroweak phase transition. In order to achieve this goal, we first must recall some basics facts about the standard model of particle physics, see Refs. [16–22].

The Higgs field is a doublet of complex scalar fields. It is charged under the group $U(1)_Y \times SU(2)_L$ characterized by two coupling constant g' and g. The Higgs field Lagrangian reads

$$\mathcal{L}_{\mathrm{Higgs}} = -(D_\mu \Sigma)^\dagger D^\mu \Sigma - V(\Sigma, \Sigma^\dagger) \tag{17}$$

where the covariant derivative can be expressed as

$$D_\mu \Sigma = \partial_\mu \Sigma + i \frac{g'}{2} Y_H B_\mu \Sigma + i g T_a W_\mu^a \Sigma \tag{18}$$

where a runs from one to three. The quantities B_μ and W_μ^a are the gauge bosons, Y_H is the Higgs weak hyper-charge and T_a are the generators of the $SU(2)$ group (see below for more details about these quantities). The potential is chosen to be

$$V(\Sigma, \Sigma^\dagger) = \frac{m^2}{2} \Sigma^\dagger \Sigma + \frac{\lambda}{4} (\Sigma^\dagger \Sigma)^2 \tag{19}$$

and clearly resembles Eqs. (14) and (15) for $T < T_{\mathrm{cri}}$. Indeed, in this expression, the quantity m^2 is negative (as explained before, m^2 is in fact the square of the effective mass). This also means that the minimum of the potential is not located at $\langle \Sigma \rangle = 0$ but at

$$\langle \Sigma \rangle = v = \sqrt{-\frac{m^2}{\lambda}} \tag{20}$$

obtained from the condition $\partial V / \partial (\Sigma^\dagger \Sigma) = 0$. The quantity v is in fact the new vacuum expectation value of the Higgs after spontaneous symmetry breaking. In particle physics, one studies the theory after the transition, when the Higgs field has stabilized at its new minimum. But, in reality, as explained before, we must consider that the transition is a dynamical process. Before the transition we had $m^2 > 0$ and the minimum was located at $\langle \Sigma \rangle = 0$ and after the transition, m^2 has become negative and, consequently, the new vacuum is given by v. Returning to the minimum after the transition, it is easy to show that the value of the potential at $\langle \Sigma \rangle = v$ is given by

$$V(\langle \Sigma \rangle = v) = -\frac{m^4}{4\lambda} \tag{21}$$

and is negative. Clearly, this is because we have chosen $V_0 = 0$ in Eq. (19). As already explained in Section 3.1, this means that the vacuum energy vanishes before the transition and becomes negative after, see also Fig. 1. This means that we have considered the situation corresponding to the left panel in Fig. 2. Of course, this is arbitrary and we could also have considered the situation corresponding to the right panel in Fig. 2. Again, as explained in the previous subsection, in this case, the vacuum energy would vanish after the transition but not before.

The numerical value of the vacuum energy given by Eq. (21) is fixed by the electroweak physics in particular by the parameters m and λ. Therefore, in order to compute ρ_{vac} explicitly, we must explain how m and λ are evaluated. For this
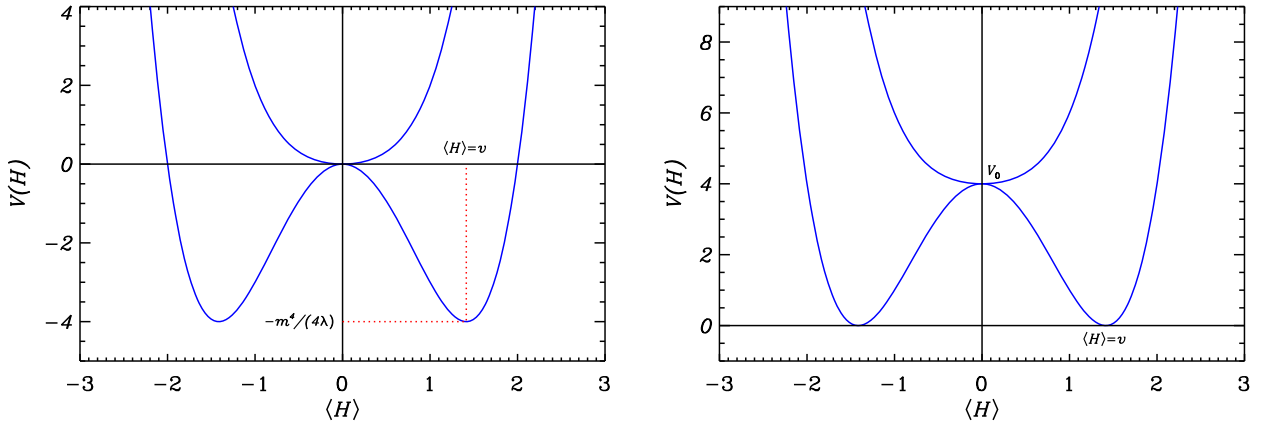
Fig. 2. Effective potential of the Higgs boson before and after the electroweak phase transition. The left panel corresponds to a situation where the vacuum energy vanishes at high temperature. As a consequence ρ_{vac} is negative at temperature smaller than the critical temperature. This is the situation treated in the text where the quantity $-m^4/(4\lambda)$ is explicitly calculated. On the right panel, the off-set parameter V_0 is chosen such that the vacuum energy is zero after the transition. As a consequence, it does not vanish at high temperatures.

purpose, we now briefly recall the main features of the standard model of particle physics. Let us come back to the Higgs Lagrangian given at the beginning of this section, see Eq. (17). In the expression of $\mathcal{L}_{\text{Higgs}}$, the quantities T_a are the generator of SU(2)$_L$ and are given by $T_a = \sigma_a/2$ where σ_a are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{22}$$

Using the gauge invariance, one can always rewrite the complex doublet as

$$\Sigma = \left(v + \frac{H}{\sqrt{2}} \right) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{23}$$

where H is a real scalar field, the "Higgs boson". Therefore, the covariant derivative (18) can be expressed as

$$D_\mu \Sigma = \begin{pmatrix} i\frac{g}{2}(v + \frac{H}{\sqrt{2}})(W_\mu^1 - iW_\mu^2) \\ \frac{1}{\sqrt{2}}\partial_\mu H + i(v + \frac{H}{\sqrt{2}})(\frac{g'}{2}Y_H B_\mu - \frac{g}{2}W_\mu^3) \end{pmatrix} \tag{24}$$

and the Higgs Lagrangian (17) becomes

$$\mathcal{L}_{\text{Higgs}} = -\frac{1}{2}\partial_\mu H \partial^\mu H - V\left(\Sigma, \Sigma^\dagger\right) - \left(v^2 + \sqrt{2}vH + \frac{1}{2}H^2 \right)$$
$$\times \left(\frac{g^2}{4}W_\mu^1 W^{1\mu} + \frac{g^2}{4}W_\mu^2 W^{2\mu} + \frac{g^2}{4}W_\mu^3 W^{3\mu} + \frac{g'^2}{4}Y_H^2 B_\mu B^\mu - \frac{gg'}{2}Y_H B_\mu W^{3\mu} \right) \tag{25}$$

where the scalar potential can be written as

$$V\left(\Sigma, \Sigma^\dagger\right) = -\frac{\lambda v^4}{4} + \frac{1}{2}\lambda v^2 H^2 + \frac{\lambda}{2}\frac{v}{\sqrt{2}}H^3 + \frac{\lambda}{16}H^4 \tag{26}$$

We now introduce the physical gauge bosons. As is well known, they are defined by

$$B_\mu = -\sin\theta Z_\mu + \cos\theta A_\mu, \qquad W_\mu^3 = \cos\theta Z_\mu + \sin\theta A_\mu, \qquad W_\mu^2 = \frac{W_\mu^+ - W_\mu^-}{i\sqrt{2}}, \qquad W_\mu^1 = \frac{W_\mu^+ + W_\mu^-}{\sqrt{2}} \tag{27}$$

where θ is the Weinberg angle, A_μ represents the photon and Z_μ, $W_\mu^\pm$ the massive gauge bosons. Expressing the Higgs Lagrangian (25) in terms of these new fields, one obtains the following expression

$$\mathcal{L}_{\text{Higgs}} = -\frac{1}{2}\partial_\mu H \partial^\mu H - V\left(\Sigma, \Sigma^\dagger\right) - \left(v^2 + \sqrt{2}vH + \frac{1}{2}H^2 \right)\left[\frac{g^2}{4}W_\mu^+ W^{-\mu} + \frac{g^2}{4}W_\mu^- W^{+\mu} \right.$$
$$+ Z_\mu Z^\mu \left(\frac{g^2}{4}\cos^2\theta + \frac{g'^2}{4}Y_H^2 \sin^2\theta + \frac{gg'}{2}Y_H \sin\theta \cos\theta \right)$$
$$+ A_\mu A^\mu \left(\frac{g^2}{4}\sin^2\theta + \frac{g'^2}{4}Y_H^2 \cos^2\theta - \frac{gg'}{2}Y_H \sin\theta \cos\theta \right)$$

$$+ Z_\mu A^\mu \left(\frac{g^2}{2} \cos\theta \sin\theta - \frac{g'^2}{2} Y_H^2 \cos\theta \sin\theta - \frac{gg'}{2} Y_H \cos^2\theta + \frac{gg'}{2} Y_H \sin^2\theta \right) \Bigg] \tag{28}$$

Of course $\mathcal{L}_{\text{Higgs}}$ is not the only term in the standard model Lagrangian. There are also the terms describing the kinetic terms of the gauge fields

$$\mathcal{L}_{\text{K-Gauge}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{a\mu\nu} W^{a\mu\nu}$$
$$= -\frac{1}{4} A_{\mu\nu} A^{\mu\nu} - \frac{1}{4} Z_{\mu\nu} Z^{\mu\nu} - \frac{1}{4} W_{+\mu\nu} W^{+\mu\nu} - \frac{1}{4} W_{-\mu\nu} W^{-\mu\nu} + \cdots \tag{29}$$

where $B_{\mu\nu}$ is the B field strength, namely $B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu$ and where a similar definition applies to the other gauge fields.

Finally, one must also specify how leptons are described. In the standard model of particle physics, a fermion is represented by a four-component Dirac spinor Ψ, see also Section 7.1 where the zero-point fluctuations of spinors are calculated. Let us now recall what chirality is. For this purpose, we introduce the matrix γ^5 such that $(\gamma^5)^2 = \mathbb{I}$ and $(\gamma^5)^\dagger = \gamma^5$. This allows us to define two chiral projectors by

$$P_R \equiv \frac{\mathbb{I}_4 + \gamma_5}{2}, \qquad P_L = \frac{\mathbb{I}_4 - \gamma_5}{2} \tag{30}$$

where $\mathbb{I}_4$ denotes the identity four by four matrix. We now define the Dirac matrices, see Section 7.1 for a more general approach. In the chiral representation, they are given by

$$\gamma^0 = \begin{pmatrix} 0 & -\mathbb{I}_2 \\ -\mathbb{I}_2 & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$$
$$\gamma^5 = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix} \tag{31}$$

and the two projectors can be written as

$$P_R = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & 0 \end{pmatrix}, \qquad P_L = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I}_2 \end{pmatrix} \tag{32}$$

Given a Dirac spinor Ψ (we denote the four-component spinors by a capital Greek letter and the two-component spinors by an ordinary Greek letter)

$$\Psi = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} \tag{33}$$

one can thus define the two states of chirality (right- and left-handed) by means of the following two expressions:

$$\Psi_R \equiv P_R \Psi = \begin{pmatrix} \psi_R \\ 0 \end{pmatrix}, \qquad \Psi_L \equiv P_L \Psi \begin{pmatrix} 0 \\ \psi_L \end{pmatrix} \tag{34}$$

The quantities $\psi_{R,L}$ are two-component spinors known as Weyl spinors. They play a very important role in super-symmetry, see Section 7.4. Notice that we also have $\overline{\Psi} \equiv \Psi^\dagger \gamma^0 = -(\psi_L^\dagger, \psi_R^\dagger)$. This implies that the (mass-less) Dirac Lagrangian can be expressed as

$$\mathcal{L}_{\text{Dirac}} = -i\overline{\Psi} \gamma^\mu \partial_\mu \Psi = -i\psi_L^\dagger \overline{\sigma}^\mu \partial_\mu \psi_L - i\psi_R^\dagger \sigma^\mu \partial_\mu \psi_R \tag{35}$$

where we have defined the covariant Pauli matrices $\sigma^\mu = (\mathbb{I}_2, \sigma^i)$ and $\overline{\sigma}^\mu = (\mathbb{I}_2, -\sigma^i)$.

After these brief reminders, we now consider the first family of the standard model leptons (the other two families can be treated in the same fashion and the quarks must be considered separately), i.e. the electron and the neutrino. Following the usual convention, we denote by e_R the right-handed electron Weyl spinor and by e_L the left-handed one. With regards to the neutrino, only the left-handed particle is present in the standard model and will be denoted as ν_L. Then, we have the following properties. The particles e_L and ν_L are charged under $U(1)_Y$ but also under $SU(2)_L$, i.e. they are considered as a complex doublet of two-component spinors

$$L_e = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{36}$$

On the other hand, e_R is charged under $U(1)_Y$ but is postulated to be a $SU(2)_L$ singlet. Therefore, concretely, the Lagrangian can be written as

$$\mathcal{L}_{\text{Leptons}} = -iL_e^\dagger \bar{\sigma}^\mu \left(\partial_\mu + i\frac{g'}{2} Y_L B_\mu + igT_a W_\mu^a \right) L_e$$
$$- ie_R^\dagger \bar{\sigma}^\mu \left(\partial_\mu + i\frac{g'}{2} Y_R B_\mu \right) e_R \tag{37}$$

where $L_e^\dagger$ means $(\nu_L^\dagger, e_L^\dagger)$. Notice the factor $1/2$ in front of the gauge boson B_μ which is the usual convention [this factor was already present in Eq. (18)]. The quantities Y_L and Y_R are the leptonic weak hyper-charges. Let us focus on the first term in the above expression. The term between the parenthesis is a two by two matrix that can be written in terms of the physical gauge bosons, using the definitions (27). Similarly, the second term can be expressed in terms of the fields A_μ and Z_μ. Then, the Lagrangian for the leptons can be rewritten as

$$\mathcal{L}_{\text{Leptons}} = -i\nu_L^\dagger \bar{\sigma}_\mu \left[\partial_\mu - i\left(\frac{g'}{2} Y_L \sin\theta - \frac{g}{2}\cos\theta \right) Z_\mu + i\left(\frac{g'}{2} Y_L \cos\theta + \frac{g}{2}\sin\theta \right) A_\mu \right] \nu_L$$
$$- ie_L^\dagger \bar{\sigma}_\mu \left[\partial_\mu - i\left(\frac{g'}{2} Y_L \sin\theta + \frac{g}{2}\cos\theta \right) Z_\mu + i\left(\frac{g'}{2} Y_L \cos\theta - \frac{g}{2}\sin\theta \right) A_\mu \right] e_L$$
$$+ \frac{g}{\sqrt{2}} \nu_L^\dagger \bar{\sigma}^\mu W_\mu^- e_L + \frac{g}{\sqrt{2}} e_L^\dagger \bar{\sigma}^\mu W_\mu^+ \nu_L - ie_R^\dagger \bar{\sigma}^\mu \left(\partial_\mu - i\frac{g'}{2} Y_R \sin\theta Z_\mu + i\frac{g'}{2} Y_R \cos\theta A_\mu \right) e_R \tag{38}$$

Therefore, the total Lagrangian is now given by $\mathcal{L} = \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{K-Gauge}} + \mathcal{L}_{\text{Leptons}}$. This Lagrangian depends on various free parameters that we now discuss and choose. Firstly, the neutrino is not electromagnetically charged and from Eq. (38) one has that

$$g' Y_L \cos\theta + g\sin\theta = 0 \tag{39}$$

Secondly, the electron carries the charge e. This means that one must recover the following Lagrangian,

$$\mathcal{L}_{\text{EM}} = -i\overline{\Psi}\gamma^\mu (\partial_\mu + iQ A_\mu)\Psi \tag{40}$$

where Q is the electromagnetic charge, $Q = e$ in our case. Working out the interaction part of the above Lagrangian, one obtains

$$\mathcal{L}_{\text{EM}}^{\text{int}} = -Q\, e_L^\dagger \bar{\sigma}_\mu A_\mu e_L - Q\, e_R^\dagger \sigma_\mu A_\mu e_R, \tag{41}$$

and, looking at Eq. (38), one deduces that

$$\frac{g'}{2} Y_L \cos\theta - \frac{g}{2}\sin\theta = \frac{g'}{2} Y_R \cos\theta = e \tag{42}$$

As a consequence, comparing Eqs. (39) and (42), one can establish that $2Y_L = Y_R$. It is conventional to choose $Y_L = -1$ and therefore $Y_R = -2$. Then, Eq. (39) leads to $g' = g\tan\theta$, from which one can write $\cos\theta = g/\sqrt{g^2 + g'^2}$ and $\sin\theta = g'/\sqrt{g^2 + g'^2}$. Finally, Eq. (42) implies that $g\sin\theta = e$. Everything can be summarized by the Gell-Mann formula

$$Q = e\left(I_3 + \frac{Y}{2} \right) \tag{43}$$

if the leptonic isospin doublet is assigned a "weak isospin" $I = 1/2$ such that the neutrino has a third component $I_3 = 1/2$ while the electron is such that $I_3 = -1/2$. Of course the right-handed particles are such that $I = 0$. Returning to Eq. (28) and using the above equations, we see that the photons remains mass-less if $Y_H = 1$. This also allows us to determine the mass of the gauge bosons and one arrives at

$$m_{W^\pm}^2 = \frac{g^2 v^2}{2} \tag{44}$$
$$m_Z^2 = 2v^2 \left(\frac{g}{2}\cos\theta + \frac{g'}{2}\sin\theta \right)^2 = \frac{m_{W^\pm}^2}{\cos^2\theta} \tag{45}$$

One also easily proves that the coefficient of the cross-term $Z_\mu A^\mu$ vanishes. Finally, looking at Eq. (26), one establishes that the Higgs mass is given by

$$m_H^2 = \lambda v^2 \tag{46}$$

and, from the expression of the potential (26), one finds the expression of its value at the minimum

$$V(\langle \Sigma \rangle = v) = -\lambda \frac{v^4}{4} = -\frac{m_H^4}{4\lambda} \tag{47}$$

in agreement with what we had before. Using the expression of the Higgs mass, one can re-write this expression as

$$\rho_{\text{vac}} = V\left(\langle \Sigma \rangle = v\right) = -\frac{1}{4}m_{\text{H}}^2 v^2 \tag{48}$$

We see that in order to calculate the vacuum energy, we need to know v and the Higgs mass. Let us first discuss how the vacuum expectation value of the Higgs can be obtained. For this purpose, we focus on the term in the electroweak Lagrangian describing the interaction between leptons and charged gauge bosons (the so-called charged currents). From the above considerations, see Eq. (38), they are given by

$$\mathcal{L}_{\text{intL-W}} = \frac{g}{\sqrt{2}} v_{\text{L}}^\dagger \bar{\sigma}^\mu W_\mu^- e_{\text{L}} + \frac{g}{\sqrt{2}} e_{\text{L}}^\dagger \bar{\sigma}^\mu W_\mu^+ v_{\text{L}} \tag{49}$$

It is easy to verify by an explicit calculation that it can be re-expressed in terms of Dirac spinors[1] as

$$\begin{aligned}
\mathcal{L}_{\text{intL-W}} = {} & \frac{g}{\sqrt{2}} \overline{\Psi}_\nu \gamma^\mu \frac{1-\gamma_5}{2} W_\mu^- \Psi_e \\
& + \frac{g}{\sqrt{2}} \overline{\Psi}_e \gamma^\mu \frac{1-\gamma_5}{2} W_\mu^+ \Psi_\nu
\end{aligned} \tag{50}$$

which is similar to the Fermi theory for the weak interaction described by the following Lagrangian

$$\mathcal{L}_{\text{Fermi}} = \frac{G_{\text{F}}}{\sqrt{2}} J_\lambda J^{\lambda\dagger} \tag{51}$$

with $J_\lambda = J_\lambda^{(e)} + J_\lambda^{(\mu)}$ (i.e. an electronic and a muonic part) where the current is given by $J_\lambda^{(e)} = \overline{\Psi}_e \gamma_\lambda (1 - \gamma_5) \Psi_{\nu_e}$ and a similar definition for $J_\lambda^{(\nu)}$. If, for instance we consider the decay of muon into an electron, a muonic neutrino and an anti-electronic neutrino, described by the graph

(52)

the two theories will lead to the same leading contribution at small energies if the following identification is made

$$\frac{g^2}{8m_W^2} = \frac{G_{\text{F}}}{\sqrt{2}} \tag{53}$$

Indeed, the appearance of the gauge boson mass comes from the propagators of W_μ. At small momenta compared to the mass, only m_W remains and the above graph is in fact equivalent to the graph corresponding to the Fermi theory. Then, using the expression of the W-bosons mass established before, $m_W^2 = g^2 v^2/2$, one arrives at

$$v^2 = \frac{\sqrt{2}}{4G_{\text{F}}^2} \tag{54}$$

Notice that there is a factor $\sqrt{2}$ difference with respect to the textbook [20] because we have defined the Higgs as $v + H/\sqrt{2}$ instead of $(v + H)/\sqrt{2}$. The Fermi constant is given by $G_{\text{F}} \simeq 1.16 \times 10^{-5}$ (GeV)$^{-2}$. Therefore,

$$v \simeq 175 \text{ GeV} \tag{55}$$

or $\sqrt{2}v \simeq 246$ GeV. We see that we have been able to calculate the Higgs vacuum expectation value v. As a consequence, Eq. (48) reads

$$\rho_{\text{vac}} = -\frac{\sqrt{2}}{16} \frac{m_{\text{H}}^2}{G_{\text{F}}^2} \tag{56}$$

The mass of the Higgs boson is not known although, at the time of writing, there are reasons to believe that $m_{\text{H}} < 129$ GeV at 95% CL [46] (there are even reasons to believe that $m_{\text{H}} \simeq 125$ GeV but this is not yet established at a sufficient statistical level at the time of writing). Therefore, we find that

---

[1] Let us recall that they are defined by the following expression

$$\Psi_e = \begin{pmatrix} e_{\text{L}} \\ e_{\text{R}} \end{pmatrix}$$

and a similar expression for Ψ_ν.

$$\rho_{\text{vac}}^{\text{EW}} \simeq -1.2 \times 10^8 \text{ GeV}^4 \simeq -10^{55} \rho_{\text{cri}} \tag{57}$$

where $\rho_{\text{cri}} \simeq 10^{-47}$ GeV^4 is the critical energy density today. As will be discussed in the following, this result is in fact in contradiction with various observations.

To conclude this section, let us recall that one can always adjust the vacuum energy today to zero by tuning the parameter V_0. This is clearly not a very satisfactory method and this implies that the vacuum energy density was huge prior to the electroweak phase transition. It is also worth noticing that this problem is not specific to the electroweak transition. Exactly the same discussion could be presented for the Quantum Chromo Dynamics (QCD) transition [3,8] (for instance) and would lead to the same problem since

$$\rho_{\text{vac}}^{\text{QCD}} \simeq 10^{-2} \text{ GeV}^4 \simeq 10^{45} \rho_{\text{cri}} \tag{58}$$

Moreover, as we are now going to discuss in detail, even if we "solve" the classical cosmological constant by tuning the vacuum energy to zero, see Eq. (8), the problem reappears at the quantum level. Indeed, we have seen that the zero-point quantum fluctuations also give a contribution to the cosmological constant, see Eq. (5). The goal of the next section is to estimate the corresponding energy density.

4. The quantum-mechanical cosmological constant problem

We have seen that, even if we solve the classical cosmological constant problem and find a convincing reason to put the minimum of the potential to zero, ρ_{vac} is a quantity which also receives contributions from the zero-point fluctuations of all the quantum fields present in the Universe. We now discuss this "quantum-mechanical" cosmological constant problem. In principle, the corresponding vacuum energy density can be estimated from first principles. The goal of this section (and of Section 6) is to carry out this task. We will first discuss how the Lorentz invariance affects the calculation of the vacuum energy, see the next Section 4.1. Then, in Section 4.2, we will show how this question can be formulated in terms of Feynman diagrams. In Section 4.3, we will also evaluate the vacuum energy not only for a free theory as usually done but also in the more realistic case where interactions are present. Finally, in order to understand better the origin of the vacuum energy density, in Sections 5.1 and 5.2, we will address the same question but in the context of ordinary quantum mechanics. This will allow us in Sections 5.3 and 5.4 to discuss in more detail the properties of the so-called bubble diagrams from which the cosmological constant originates. As mentioned above, all the fields present in the Universe participate in the vacuum energy. Therefore, we need to do the calculation for scalar, fermion and vector fields. Here, in a first step, we carry out the calculation for a real scalar field. The other types will be treated later on in Section 7. For simplicity, we will also consider that the metric is flat and that the fields live in Minkowski space–time. A sloppy justification is that the main contribution to the vacuum energy density comes from modes with very high momenta, corresponding to scales at which the curvature of space–time is negligible. In fact, we will come back to this very important question in more detail in what follows and will try to discuss this point in a more rigorous way, see Section 8.

4.1. The zero-point energy density

Let us consider a simple real free scalar field with the potential $V(\Phi) = m^2 \Phi^2 / 2$ where m is the mass of the scalar particle. In flat space–time, the equation of motion is nothing but the Klein–Gordon equation, namely

$$-\ddot{\Phi} + \delta^{ij} \partial_i \partial_j \Phi - m^2 \Phi = 0 \tag{59}$$

Since the scalar field is free, the equation of motion is linear and, as a consequence, one can Fourier expand $\Phi(t, \boldsymbol{x})$ as

$$\Phi(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 \boldsymbol{k}}{\sqrt{2\omega(k)}} \left(c_{\boldsymbol{k}} \mathrm{e}^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + c_{\boldsymbol{k}}^{\dagger} \mathrm{e}^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}} \right) \tag{60}$$

with

$$\omega(k) \equiv \sqrt{k^2 + m^2} \tag{61}$$

The four-dimensional momentum has been written as $k^\mu = (k^0, \boldsymbol{k})$ and the integration is performed over its spatial part only. We have used the notation $|\boldsymbol{k}| = k$. Its time component is given by $k^0 = \omega$. As usual, the field has been quantized by considering $c_{\boldsymbol{k}}$ and $c_{\boldsymbol{k}}^{\dagger}$ (the so-called annihilation and creation operators) as quantum operators satisfying the following commutation relation:

$$\left[c_{\boldsymbol{k}}, c_{\boldsymbol{k'}}^{\dagger} \right] = \delta^{(3)} \left(\boldsymbol{k} - \boldsymbol{k'} \right) \tag{62}$$

This equation is in fact equivalent to the standard commutation rule between the field operator and its conjugate momentum.

We are now in a position where one can compute $\langle 0|T_{\mu\nu}|0\rangle$. From Eq. (7), we see that we must first evaluate the mean values of various quantities depending on the field operator and its derivatives. Straightforward calculations lead to

$$\langle 0|\dot{\Phi}^2|0\rangle = \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}^3\boldsymbol{k}}{2\omega(k)} \, \omega^2(k) \tag{63}$$

$$\langle 0|\delta^{ij}\partial_i\Phi\partial_j\Phi|0\rangle = \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}^3\boldsymbol{k}}{2\omega(k)} \, \boldsymbol{k}^2 \tag{64}$$

$$\langle 0|\Phi^2|0\rangle = \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}^3\boldsymbol{k}}{2\omega(k)} \tag{65}$$

These equations can then be used to determine the energy density and pressure of the vacuum. From the above expression (7) of the stress–energy tensor, one has

$$T_{00} = \mathcal{H} = \frac{1}{2}\dot{\Phi}^2 + \frac{1}{2}\delta^{ij}\partial_i\Phi\partial_j\Phi + \frac{1}{2}m^2\Phi^2 \tag{66}$$

where $\mathcal{H}$ denotes the Hamiltonian density. This leads to the following expression for the energy density $\rho = u^\mu u^\nu T_{\mu\nu}$ measured by a fundamental observer characterized by its four-velocity vector $u^\mu = (1, \mathbf{0})$

$$\langle\rho\rangle = \langle 0|u^0 u^0 T_{00}|0\rangle = \frac{1}{(2\pi)^3}\frac{1}{2}\int \mathrm{d}^3\boldsymbol{k}\,\omega(k) \tag{67}$$

In the same manner, one can easily estimate the pressure. One obtains

$$\langle p\rangle = \left\langle 0\left|\frac{1}{3}\perp^{\mu\nu}T_{\mu\nu}\right|0\right\rangle$$
$$= \left\langle 0\left|\frac{1}{2}\dot{\Phi}^2 - \frac{1}{6}\delta^{ij}\partial_i\Phi\partial_j\Phi - \frac{1}{2}m^2\Phi^2\right|0\right\rangle \tag{68}$$

where $\perp^{\mu\nu} \equiv g^{\mu\nu} + u^\mu u^\nu$ is a projector. Upon using the expressions (63), (64) and (65), one arrives at the following expression:

$$\langle p\rangle = \frac{1}{(2\pi)^3}\frac{1}{6}\int \mathrm{d}^3\boldsymbol{k}\frac{k^2}{\omega(k)} \tag{69}$$

Notice that we also have

$$\langle 0|T_{0i}|0\rangle = -\frac{1}{(2\pi)^3}\frac{1}{2}\int \mathrm{d}^3\boldsymbol{k}k_i = 0 \tag{70}$$

and, hence, we do not need to consider to off-diagonal elements. This completes our calculation of the vacuum stress–energy tensor.

From the above considerations, using Eq. (11), we deduce that

$$\Lambda_{\mathrm{eff}} = \Lambda_{\mathrm{B}} + \frac{\kappa}{(2\pi)^3}\int \mathrm{d}^3\boldsymbol{k}\frac{1}{2}\omega(k) \tag{71}$$

The obvious problem with the previous calculation is that the integrals expressing the energy density $\langle\rho\rangle$ and the pressure $\langle p\rangle$ blow up in the ultra-violet regime (and, therefore, strictly speaking, the effective cosmological constant is in fact infinite). This is of course a well-known problem in quantum field theory. Usually, this type of divergences is ignored (or sometimes removed by mean of the normal ordering product) because, in absence of gravity, only differences of energy can be detected [16,17]. However, in the present context, this is clearly no longer the case since the absolute value of the vacuum energy weighs and, therefore, can be measured. This is why regularizing these infinities becomes a crucial issue that we now discuss in some details.

The common method found in the literature is to introduce a cut-off at $k = M$, where the physical interpretation of M is that this is the scale at which the effective theory used before breaks down [2,4,5,7,9]. The quantity M is not known: it could be the Planck scale, the string scale or even the super-symmetric breaking scale. As we will see later on, this is in fact not important because the cosmological constant problem is present regardless of the precise value of the cut-off. In this approach, the energy density becomes

$$\langle\rho\rangle = \frac{1}{4\pi^2}\int\limits_0^M \mathrm{d}k\,k^2\sqrt{k^2 + m^2} \tag{72}$$

$$= \frac{M^4}{16\pi^2}\left[\sqrt{1 + \frac{m^2}{M^2}}\left(1 + \frac{1}{2}\frac{m^2}{M^2}\right) - \frac{1}{2}\frac{m^4}{M^4}\ln\left(\frac{M}{m} + \frac{M}{m}\sqrt{1 + \frac{m^2}{M^2}}\right)\right] \tag{73}$$

$$= \frac{M^4}{16\pi^2}\left(1 + \frac{m^2}{M^2} + \cdots\right) \tag{74}$$

where, in the last expression, we have expanded the exact expression in terms of the small parameter m/M. We see that the divergence is quartic in the cut-off scale. Let us now perform the same calculation for the pressure. Using Eq. (69), one obtains

$$\langle p \rangle = \frac{1}{3} \frac{1}{4\pi^2} \int_0^M dk \frac{k^4}{\sqrt{k^2 + m^2}} \tag{75}$$

$$= \frac{1}{3} \frac{M^4}{16\pi^2} \left[\sqrt{1 + \frac{m^2}{M^2}} \left(1 - \frac{3}{2} \frac{m^2}{M^2} \right) + \frac{3}{2} \frac{m^4}{M^4} \ln\left(\frac{M}{m} + \frac{M}{m} \sqrt{1 + \frac{m^2}{M^2}} \right) \right] \tag{76}$$

$$= \frac{1}{3} \frac{M^4}{16\pi^2} \left(1 - \frac{m^2}{M^2} + \cdots \right) \tag{77}$$

It is clear from the previous expressions that $\langle p \rangle / \langle \rho \rangle \neq -1$ which indicates that the stress–energy tensor is not of the form $\propto -\rho g_{\mu\nu}$. In the limit $m \to 0$, as can be easily shown from Eqs. (74) and (77), the equation of state is in fact $\langle p \rangle / \langle \rho \rangle = 1/3$. This would mean that the zero point fluctuations do not behave like a cosmological constant but rather like radiation. On the other hand, this result seems to be strange since we know from fundamental considerations that the stress–energy tensor of a system in its ground state must be proportional to the metric tensor. Clearly, something has gone wrong is our calculation.

Interestingly enough, let us also note in passing that if one only focuses on the logarithmic terms, namely

$$\langle \rho \rangle^{\log} = -\frac{m^4}{32\pi^2} \ln\left(\frac{M}{m} + \frac{M}{m} \sqrt{1 + \frac{m^2}{M^2}} \right) \tag{78}$$

$$\langle p \rangle^{\log} = \frac{m^4}{32\pi^2} \ln\left(\frac{M}{m} + \frac{M}{m} \sqrt{1 + \frac{m^2}{M^2}} \right) \tag{79}$$

then the equation of state of the vacuum is correctly reproduced (even if the energy density can be negative). In fact, as we explain below, this is not a coincidence [47].

It is easy to check that the previous considerations are not an artifact of the fact that we have introduced a sharp cut-off. Indeed, if rather than a sharp cut-off in the upper limit of the integrals, one introduces a smooth exponential cut-off, one obtains (for simplicity, one only considers the case $m = 0$),

$$\langle \rho \rangle = \frac{1}{4\pi^2} \int_0^\infty dk \, k^3 e^{-\alpha k} = \frac{1}{4\pi^2} \frac{\Gamma(4)}{\alpha^4} \tag{80}$$

$$= \frac{1}{4\pi^2} 6 M^4 \tag{81}$$

where we have written $\alpha = 1/M$ and where $\Gamma(z)$ is the Euler's integral of the second kind [48,49]. One recovers our familiar quartic divergence. Doing the same manipulation for the pressure, one obtains again that $\langle p \rangle / \langle \rho \rangle = 1/3$.

In fact, what has gone wrong is that we have used schemes of regularization that do not respect the symmetry of our underlying theory [47,50,51]. Indeed, the Lorentz invariance is broken when we impose a cut-off on the spatial momentum only. As exemplified by Eq. (81), it is easy to show that this is a feature of any regularization scheme breaking Lorentz invariance and not a specific property of the particular cut-off scheme used above (and usually used in the literature). In fact, this property is well known in the context of gauge theories. If one regulates divergent graphs with a method that is not gauge invariant, then one obtains incorrect results [52]. Here, the situation is very similar.

Therefore, in order to meaningfully evaluate the zero-point energy density, one must use a regularization scheme that does not break Lorentz invariance [47,50,51]. An obvious choice is dimensional regularization and we now explore this route. In order to use this method, we first reformulate our problem in a d-dimensional Minkowski space–time. Obviously, the Klein–Gordon equation (59) remains unchanged except that the indices of the Krönecker symbol now run from 0 to $d - 1$. The Fourier expansion of the field becomes [compare with Eq. (60)]

$$\Phi(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{(d-1)/2}} \int \frac{d^{d-1}\boldsymbol{k}}{\sqrt{2\omega(k)}} \left(c_{\boldsymbol{k}} e^{-i\omega t + i\boldsymbol{k} \cdot \boldsymbol{x}} + c_{\boldsymbol{k}}^\dagger e^{i\omega t - i\boldsymbol{k} \cdot \boldsymbol{x}} \right) \tag{82}$$

and, as a consequence, the energy density can be expressed as

$$\langle \rho \rangle = \frac{\mu^{4-d}}{(2\pi)^{(d-1)}} \frac{1}{2} \int d^{d-1}\boldsymbol{k} \, \omega(k) \tag{83}$$

$$= \frac{\mu^{4-d}}{(2\pi)^{(d-1)}} \frac{1}{2} \int\limits_0^\infty dk \, k^{d-2} \, d^{d-2}\Omega \, \omega(k) \tag{84}$$

where we have introduced a scale μ in order for the equation to be dimensionally correct. This equation should be compared with Eq. (67). Then, using the fact that the angular integrals can be easily performed, namely $\int d^{d-2}\Omega = 2\pi^{(d-1)/2}/\Gamma(d/2 - 1/2)$, where, again, $\Gamma(z)$ is the Euler's integral of the second kind [48,49], one obtains

$$\langle\rho\rangle = \frac{\mu^4}{2(4\pi)^{(d-1)/2}} \frac{\Gamma(-d/2)}{\Gamma(-1/2)} \left(\frac{m}{\mu}\right)^d \tag{85}$$

Performing the same calculation for the pressure, one arrives at a similar expression

$$\langle p\rangle = \frac{\mu^{4-d}}{(2\pi)^{(d-1)}} \frac{1}{2(d-1)} \int d^{d-1}\boldsymbol{k} \, \frac{k^2}{\omega(k)} \tag{86}$$

$$= \frac{\mu^4}{4(4\pi)^{(d-1)/2}} \frac{\Gamma(-d/2)}{\Gamma(1/2)} \left(\frac{m}{\mu}\right)^d \tag{87}$$

But $\Gamma(-1/2) = -2\Gamma(1/2)$ and, as a consequence,

$$\langle p\rangle = -\langle\rho\rangle \tag{88}$$

This time, as expected, we have obtained the equation of state of the vacuum. The fact that the regularization method used satisfies the Lorentz symmetry has led us to the correct result.

The consistency of the above result can be checked in a different manner [47]. Using Eqs. (83) and (86) for the energy density and the pressure, one sees that they satisfy the following differential equation:

$$\langle\rho\rangle - 2m^2 \frac{d\langle\rho\rangle}{dm^2} = (d-1)\langle p\rangle \tag{89}$$

Using the fact that $\langle p\rangle = -\langle\rho\rangle$ one obtains

$$m^2 \frac{d\langle\rho\rangle}{dm^2} = \frac{d}{2}\langle\rho\rangle \tag{90}$$

which can be integrated to give $\langle\rho\rangle \propto m^d$ in full agreement with Eq. (85). Therefore, we conclude that the above considerations are all consistent.

Let us now analyze in more details the structure of the divergences. For this purpose one writes $d = 4 - \epsilon$, where ϵ is supposed to be a small parameter, and we expand $\langle\rho\rangle$ in terms of ϵ. The Euler function is such that

$$\Gamma\left(-2 + \frac{\epsilon}{2}\right) = \frac{1}{-2 + \epsilon/2} \frac{1}{-1 + \epsilon/2} \frac{1}{\epsilon/2} \Gamma\left(1 + \frac{\epsilon}{2}\right) \tag{91}$$

with $\Gamma(1 + \epsilon/2) \simeq \Gamma(1) + \epsilon\Gamma'(1)/2$, $\Gamma'(1) = \Psi(1)\Gamma(1)$ where Ψ is the Polygamma function [48,49] and $\Psi(1) = -\gamma \simeq 0.57772$ the Euler–Mascheroni constant [48,49]. If, in addition, we use the two following expansions:

$$(4\pi)^{-3/2 + \epsilon/2} \simeq \frac{1}{(4\pi)^{3/2}} \left[1 + \frac{\epsilon}{2}\ln(4\pi)\right] \tag{92}$$

$$\left(\frac{m}{\mu}\right)^{4-\epsilon} \simeq \left(\frac{m}{\mu}\right)^4 \left(1 - \epsilon\frac{m}{\mu}\right) \tag{93}$$

then one arrives at the following expression for the energy density of the vacuum:

$$\langle\rho\rangle \simeq -\frac{m^4}{64\pi^2} \left[\frac{2}{\epsilon} + \frac{3}{2} - \gamma - \ln\left(\frac{m^2}{4\pi\mu^2}\right)\right] + \cdots \tag{94}$$

Using a $\overline{\text{MS}}$ renormalization scheme convention [$\overline{\text{MS}}$ means "modified minimal subtraction" and simply consists in subtracting the pole $\sim 1/\epsilon$ together with the accompanying terms γ and $\ln(4\pi)$, see Chapter 11 of Ref. [16]; in fact, here, we slightly modify the convention such that the term $3/2$ is subtracted as well], the regularized, finite, energy density of the vacuum now reads

$$\langle\rho\rangle = \frac{m^4}{64\pi^2} \ln\left(\frac{m^2}{\mu^2}\right) \tag{95}$$

This result is very different from the result obtained by imposing a sharp cut-off. In fact, one recovers the structures of the logarithm terms found in Eqs. (78) and (79) since it is well known that the dimensional regularization scheme removes the

power law terms. As discussed around Eqs. (78) and (79), this is the reason why those terms lead to the correct vacuum equation of state. We see that the result is not proportional to the cut-off scale to the power four, as usually claimed on dimensional grounds, but to the mass of the particle to the power four, which is also dimensionally correct. Physically, for instance, this means that the photon does not contribute to the vacuum energy density contrary to what a sharp cut-off calculation would predict. Therefore, the order of magnitude of the final result can a priori be very different. We also notice that the sign of $\langle \rho \rangle$ can change according to $m > \mu$ or $m < \mu$. Of course, it remains to be seen whether this could solve the cosmological constant problem and we will argue below that, unfortunately, the answer to this question is no. It should also be kept in mind that the above calculation is valid for free fields. It is clear that a more realistic version should take into account the interactions and we will also consider this issue in the following. However, before addressing these various points and in order to gain further physical insights, we would like to investigate again the above issues, this time from the Feynman diagrams point of view. We turn to this question in the next subsection.

4.2. The vacuum and the Feynman diagrams

We have seen in the last subsection that the vacuum contribution to the cosmological constant is given in terms of divergent integrals. A priori, there is no reason to be surprised since divergent integrals are routinely found in quantum field theory. We have the method of renormalization at our disposal to get rid of these infinities and the so accurate predictions of particle physics are here to remind us that this is a very convincing way of taming these divergences. However, we will see that the divergent integrals that we encounter when we calculate the vacuum energy are in fact of a different nature than the divergent integrals met when we calculate, say, a cross section. In order to discuss this point in more detail and to understand better the structure of these divergences, we now briefly return to some textbook considerations [16–22].

In quantum field theory, the basic object is the generating functional defined by the following path integral (we still restrict our considerations to a simple scalar field):

$$Z[J] = \mathcal{N} \int \mathcal{D}\Phi \exp\left\{ i \int d^4x \left[\mathcal{L} + J(x)\Phi(x) \right] \right\} \tag{96}$$

where $\mathcal{N}$ is a number chosen such that $Z[J=0]=1$. The quantity $\mathcal{L}$ is the Lagrangian density while J represents a source coupled to the scalar field. From the generating functional, one can compute the various correlation functions

$$G^{(n)}(x_1 \cdots x_n) = \langle \Omega | T\left[\Phi(x_1) \cdots \Phi(x_n) \right] | \Omega \rangle \tag{97}$$

$$= \left(\frac{1}{i} \right)^n \frac{\delta^n Z[J]}{\delta J(x_1) \cdots \delta J(x_n)} \bigg|_{J=0} \tag{98}$$

where $|\Omega\rangle$ denotes the vacuum state of the theory. We distinguish it from $|0\rangle$ which denotes the vacuum state of the theory without interaction. In this last case, the generating functional can be computed exactly. It reads (notice that we have $Z_0[0]=1$ and that the "0" subscript indicates that we refer to the free theory)

$$Z_0[J] = \exp\left[-\frac{1}{2} \int d^4x\, d^4y\, J(x) D_F(x-y) J(y) \right] \tag{99}$$

where D_F is the Feynman propagator defined by the following expression:

$$D_F(x_1 - x_2) = \langle \Omega | T\left[\Phi(x_1)\Phi(x_2) \right] | \Omega \rangle \tag{100}$$

T denoting the time ordered product. As is well known, the propagator can be conveniently expressed in Fourier space as

$$D_F(x_1 - x_2) = \frac{i}{(2\pi)^4} \int \frac{d^4k}{k^2 + m^2} e^{ik_\mu \cdot (x_1^\mu - x_2^\mu)} \tag{101}$$

In presence of interactions, one can write $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\text{int}}$ where $\mathcal{L}_{\text{int}}$ represents the self-interacting part of the Lagrangian density. In this case the generating functional can be written as

$$Z[J] = \mathcal{N} \exp\left\{ i \int d^4x\, \mathcal{L}_{\text{int}}\left[\frac{1}{i} \frac{\delta}{\delta J(x)} \right] \right\} Z_0[J] \tag{102}$$

where, in order to have the correct normalization, the factor $\mathcal{N}$ must be given by

$$\mathcal{N}^{-1} \equiv \exp\left\{ i \int d^4x\, \mathcal{L}_{\text{int}}\left[\frac{1}{i} \frac{\delta}{\delta J(x)} \right] \right\} Z_0[J] \bigg|_{J=0} \tag{103}$$

As explained in the standard quantum field theory textbooks, the above expressions (102) and (103) are particularly well suited for a perturbative expansion. For instance, if we want to calculate the generating functional for the case of a quartic theory $\mathcal{L}_{\text{int}} = \lambda \Phi^4/4!$, where λ is the coupling constant, we are led to the following expansion in λ

$$Z[J] = \mathcal{N}\left[1 - i\frac{\lambda}{4!}\int d^4x \frac{1}{i^4}\frac{\delta^4 Z_0[J]}{\delta J(x)^4} + \cdots\right]$$ (104)

As usual, the various terms that appear in this expansion can be graphically represented by Feynman diagrams. The Feynman rules of the theory in real space are given by

$$D(x_1 - x_2) \equiv \bullet\!-\!\!-\!\!-\!\!\bullet$$ (105)

$$-i\lambda \int d^4x \equiv \times$$ (106)

$$i\int d^4x\, J(x) \equiv \circ\!-\!\!-\!\!-\!\!-$$ (107)

and a loop for a propagator evaluated at the same point in space–time. This completes our quick reminder of the basics of field theory that we need in the following.

Now, we compute the energy density of a free scalar field (the interacting case is treated in the next subsection) using the tools that we have just presented. In fact, it is interesting to calculate the trace of the stress–energy tensor. Upon using Eq. (7), one obtains

$$\langle T\rangle = \langle \eta^{\mu\nu} T_{\mu\nu}\rangle = -\langle\rho\rangle + (d-1)\langle p\rangle$$ (108)

Then, since the energy density and the pressure are given by Eqs. (83) and (86), the above relation can be re-expressed as

$$\langle T\rangle = -\mu^{4-d}\int \frac{d^{d-1}\mathbf{k}}{(2\pi)^{d-1}}\frac{m^2}{2\omega(k)}$$ (109)

But, from Eqs. (100) and (101), and now specifying to the case $d = 4$, one has

$$D_F(0) = \frac{i}{(2\pi)^4}\int \frac{d^4k}{k^2 + m^2} = \frac{i}{(2\pi)^4}\int \frac{dk^0\, d^3\mathbf{k}}{-(k^0)^2 + \omega^2}$$ (110)

where we have distinguished the integration over the time and space components of the momentum. Using the identity $\int dk^0/[-(k^0)^2 + \omega^2] = -\pi/(i\omega)$, see also Eq. (154), this leads to

$$D_F(0) = -\int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega}$$ (111)

or, in other words,

$$\langle T\rangle = m^2 D_F(0)$$ (112)

But, if we now use the fact that $\langle p\rangle = -\langle\rho\rangle$, Eq. (108) can also be written as $\langle T\rangle = -d\langle\rho\rangle$. As a consequence, Eq. (112) implies that

$$\langle\rho\rangle = -\frac{m^2}{4}D_F(0)$$ (113)

This equation can also be obtained by a direct calculation. Indeed, one has

$$D_F(0) = \frac{i\mu^{4-d}}{(2\pi)^d}\int \frac{dk^0\, d^{d-1}\mathbf{k}}{-(k^0)^2 + \omega^2}$$ (114)

$$= -\frac{d}{m^2}\frac{\mu^4}{2(4\pi)^{(d-1)/2}}\frac{\Gamma(-d/2)}{\Gamma(-1/2)}\left(\frac{m}{\mu}\right)^d$$ (115)

$$= -\frac{d}{m^2}\langle\rho\rangle$$ (116)

where, in the second step of the calculation, we have used dimensional regularization. Now, using Eq. (11) and the Feynman rule stipulating that a propagator evaluated at the same point is represented by a loop, one can re-write the cosmological constant as

$$\Lambda_{\text{eff}} = \Lambda_B - \kappa\frac{m^2}{4}\bigcirc$$ (117)

In other words, the cosmological constant is given by a "bubble diagram", i.e. a diagram that has no external leg.

We have just established that the vacuum energy density, or equivalently the cosmological constant, can be expressed in term of a bubble diagram. This type of diagram plays a special role in quantum field theory and we will discuss this point in more detail in the following next subsections. In order to develop our intuition, let us first calculate explicitly the generating function in the case of a self-interacting scalar field. The calculation is straightforward and upon using Eq. (99), this leads to

$$
-i\frac{\lambda}{4!}\int d^4x \frac{1}{i^4}\frac{\delta^4 Z_0[J]}{\delta J(x)^4} = -i\frac{\lambda}{4!}\left\{3\int d^4x\, D^2(0) - 6\int d^4x\, D(0)\left[\int d^4y\, D(x-y)J(y)\right]^2\right.
$$
$$
\left. + \int d^4x\left[\int d^4y\, D(x-y)J(y)\right]^4\right\}Z_0[J] + \cdots
\tag{118}
$$

$$
= \frac{1}{4!}\left(3\, \text{} + 6\, \text{} + \text{} + \cdots\right)Z_0[J]
\tag{119}
$$

One notices the appearance of a "bubble diagram" (the first one), similar (i.e. with no external leg) but not identical to the one encountered in the calculation of the cosmological constant. We also have a diagram with a loop (the second one) but with external legs and a diagram with no loop (the last one). All the diagram with loops (that is to say the two first ones) are divergent in quantum field theory. Nevertheless, the bubble diagram and the loop diagram play in fact a very different role as we are going to see. Let us now calculate the normalization $\mathcal{N}$, see Eq. (103). It uses the same expression as before but with $J = 0$ which means that all the diagram containing a source disappear. As a consequence, we have

$$
\mathcal{N}^{-1} = 1 + \frac{3}{4!}\, \text{}
\tag{120}
$$

We can now gather everything and compute the generating functional. Upon using the expression (104) and the results (119) and (120), one arrives at

$$
Z[J] = \left[1 + \frac{1}{4!}\left(6\, \text{} + \text{}\right) + \cdots\right]e^{\frac{1}{2}\, \text{}}
\tag{121}
$$

We see that the bubble diagram has canceled in the final expression. Since all the predictions (cross-sections, etc. ...) follow from the calculation of the generating functional, we see that the bubble diagrams never contribute, in absence of gravity, to any observables. On the contrary, when the gravitational field is turned on, the bubble diagrams become important and can affect the predictions of the theory. In summary, the cosmological constant problem can be viewed as the appearance of a new type of divergent graphs. As demonstrated with Eq. (95), these graphs can be regularized but, as will be discussed below, the regularized value seems to be incompatible with some astrophysical observations.

4.3. The vacuum in presence of interactions

So far we have calculated the vacuum energy density for a free field. It is natural to ask what happens when one considers interaction [50] since we know that a realistic model necessarily contains such terms. For this purpose, let us therefore consider again a simple self-interacting scalar field with $\mathcal{L}_{int} = \lambda\Phi^4/4!$. As a consequence, this adds the following contribution to the energy density of the vacuum:

$$
\Delta\rho = \frac{\lambda}{4!}\langle\Phi^4\rangle = \frac{3\lambda}{4!}\langle\Phi^2\rangle^2 = \frac{\lambda}{8}D_F^2(0)
\tag{122}
$$
$$
= \frac{i}{8\int d^4x}\, \text{}
\tag{123}
$$

where, on the first line in the above expression, we have used the fact that the vacuum wave-functional is a Gaussian. As expected, this new term is also given by a bubble diagram. However, one should not forget that, when interaction is present, one must also renormalize the mass. As is well known, at one loop, the renormalization of the mass comes from the tadpole diagram. More precisely, one has to calculate the correction to the propagator. It is given by the following expression:

$$
\langle\Omega|T\left[\Phi(x_1)\Phi(x_2)\right]|\Omega\rangle = \text{} + \frac{1}{2}\, \text{} + \cdots
\tag{124}
$$

or, in terms of explicit expressions (see Eq. (101))

$$\langle\Omega|T\big[\varPhi(x_1)\varPhi(x_2)\big]|\Omega\rangle = D_{\mathrm{F}}(x_1 - x_2) - \frac{i\lambda}{2}\int \mathrm{d}^4x\, D_{\mathrm{F}}(x_1 - x)D_{\mathrm{F}}(0)D_{\mathrm{F}}(x - x_2) + \cdots \tag{125}$$

$$= \frac{i}{(2\pi)^4}\int \frac{\mathrm{d}^4p}{p^2 + m^2}\mathrm{e}^{ip_\mu\cdot(x_1^\mu - x_2^\mu)} + \frac{i\lambda}{2}D_{\mathrm{F}}(0)\frac{1}{(2\pi)^4}\int \mathrm{d}^4p\,\frac{\mathrm{e}^{ip_\mu\cdot(x_1^\mu - x_2^\mu)}}{(p^2 + m^2)^2} + \cdots \tag{126}$$

$$= \frac{i}{(2\pi)^4}\int \frac{\mathrm{d}^4p}{p^2 + m^2}\mathrm{e}^{ip_\mu\cdot(x_1^\mu - x_2^\mu)}\left[1 + \frac{\lambda}{2}D_{\mathrm{F}}(0)\frac{1}{p^2 + m^2}\right] + \cdots \tag{127}$$

$$= \frac{i}{(2\pi)^4}\int \frac{\mathrm{d}^4p}{p^2 + m^2 - \lambda D_{\mathrm{F}}(0)/2}\mathrm{e}^{ip_\mu\cdot(x_1^\mu - x_2^\mu)} + \cdots \tag{128}$$

where the last equation is equivalent to the previous one at leading order in the coupling constant. We see that this implies the following redefinition (or renormalization) of the mass:

$$m_{\mathrm{ren}}^2 = m^2 - \frac{\lambda}{2}D_{\mathrm{F}}(0) \tag{129}$$

As a consequence, the vacuum energy which is now the sum of the free contribution (113) and of the contribution (122), which originates from the self-interaction of the field, can be re-expressed as

$$\begin{aligned}\langle\rho\rangle &= -\frac{m^2}{4}D_{\mathrm{F}}(0) + \frac{\lambda}{8}D_{\mathrm{F}}^2(0)\\ &= -\frac{m_{\mathrm{ren}}^2}{4}D_{\mathrm{F}}(0) - \frac{1}{4}\left[\frac{\lambda}{2}D_{\mathrm{F}}(0)\right]D_{\mathrm{F}}(0) + \frac{\lambda}{8}D_{\mathrm{F}}^2(0)\\ &= -\frac{m_{\mathrm{ren}}^2}{4}D_{\mathrm{F}}(0)\end{aligned} \tag{130}$$

where, in the second line, we have used Eq. (129). The extra contribution coming from the interacting term is exactly canceled by the renormalization of the mass [50,53]. In other words, one sees that the presence of the interaction has not modified the expression of the vacuum energy density (113), at least at one loop, provided the final result is expressed in terms of the renormalized mass rather than the bare one.

5. Bubble diagrams

The previous sections have shown the central role played by bubble diagrams in the vacuum energy problem. These diagrams do not affect the physical predictions when the gravitational field is turned off and this is why, in standard quantum field theory, there are usually ignored. In this section, we wish to discuss in more detail this type of diagram. For this purpose, we will analyze them in ordinary quantum mechanics where the technical aspects are less complicated and where the interpretation is easier [54–56].

5.1. Quantum mechanics as a field theory

Let us start with the simplest quantum-mechanical system, namely the one-dimensional harmonic oscillator. Let us denote by $x(t)$ the position of the particle evolving into a parabolic potential. Then, the corresponding Lagrangian is given by

$$L(\dot{x}, x) = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2 \tag{131}$$

where m is the mass of the particle and ω is the effective frequency. Then, let us define the "scalar field" $\varPhi(t)$ by

$$\varPhi(t) \equiv \sqrt{m}x(t) \tag{132}$$

which implies that $\varPhi$ has dimension $-1/2$. As a consequence, the Lagrangian now reads

$$L = \frac{\dot{\varPhi}^2}{2} - \frac{\omega^2}{2}\varPhi^2 \tag{133}$$

This Lagrangian is similar to that which one encounters in field theory for a free scalar field. This means that all the techniques of quantum field theory can in fact be used in ordinary quantum mechanics. Of course, this does not lead to genuine new results since both approaches are equivalent but this can shed a new light of some physical results. This is the strategy that we use here applied to the vacuum energy problem. We now quickly remind how the formalism of quantum field theory can be implemented in quantum mechanics. We will also compare this approach to the standard approach of quantum mechanics. The first step is to define the conjugate momentum. It is given by

$$\Pi(t) = \frac{\partial L}{\partial \dot{\Phi}} = \dot{\Phi}(t) \tag{134}$$

and this leads to the following Hamiltonian

$$H = \frac{\Pi^2}{2} + \frac{\omega^2}{2}\Phi^2 \tag{135}$$

The next step is to introduce the creation and annihilation $c^{\dagger}$ and c operators. They are defined by the expressions

$$\Phi = \frac{1}{\sqrt{2\omega}}(c^{\dagger} + c) \tag{136}$$

$$\Pi = i\sqrt{\frac{\omega}{2}}(c^{\dagger} - c) \tag{137}$$

As usual one can check that $[\Phi, \Pi] = i$ implies $[c, c^{\dagger}] = 1$. As it is standard in field theory, we now work in the Heisenberg picture where the operators are time dependent. A straightforward calculation shows that the Hamiltonian can be expressed as

$$H = \frac{\omega}{2}(cc^{\dagger} + c^{\dagger}c) \tag{138}$$

This allows us to calculate the time evolution of the operator c. Indeed, the Heisenberg equation reads

$$i\frac{dc}{dt} = [c, H] = \omega c \tag{139}$$

from which we deduce that

$$c(t) = c(0)e^{-i\omega t} = c_0 e^{-i\omega t} \tag{140}$$

As a consequence, the field operator can be written in terms of the creation and annihilation operators in a way which is very similar to Eq. (60), namely

$$\Phi(t) = \frac{1}{\sqrt{2\omega}}(c_0 e^{-i\omega t} + c_0^{\dagger} e^{i\omega t}) \tag{141}$$

Of course we do not have an integral over the Fourier modes because, in quantum mechanics, we deal with a single harmonic oscillator as opposed to an infinite collection of harmonic oscillators in field theory.

At this level, the theory is known exactly. It is clear that the most interesting part is in fact the treatment of the interactions which, in field theory, is based on the calculation of the Feynman diagrams. Since this calculation is itself based on the calculation of the propagator, one must now evaluate it for our simple system. It is defined by an equation similar to Eq. (100), namely

$$D(t_1, t_2) \equiv \langle 0|T[\Phi(t_1)\Phi(t_2)]|0\rangle \tag{142}$$

where we recall that the time ordered product is defined by

$$T[\Phi(t_1)\Phi(t_2)] = \begin{cases} \Phi(t_1)\Phi(t_2), & t_1 > t_2 \\ \Phi(t_2)\Phi(t_1), & t_2 > t_1 \end{cases}$$

Then, the calculation proceeds straightforwardly and one obtains for $t_1 > t_2$

$$D(t_1, t_2) = \left\langle 0 \left| \frac{1}{\sqrt{2\omega}}(c_0 e^{-i\omega t_1} + c_0^{\dagger} e^{i\omega t_1}) \frac{1}{\sqrt{2\omega}}(c_0 e^{-i\omega t_2} + c_0^{\dagger} e^{i\omega t_2}) \right| 0 \right\rangle \tag{143}$$

$$= \frac{1}{2\omega}e^{-i\omega(t_1 - t_2)} \tag{144}$$

and, in a similar way if $t_2 > t_1$,

$$D(t_1, t_2) = \left\langle 0 \left| \frac{1}{\sqrt{2\omega}}(c_0 e^{-i\omega t_2} + c_0^{\dagger} e^{i\omega t_2}) \frac{1}{\sqrt{2\omega}}(c_0 e^{-i\omega t_1} + c_0^{\dagger} e^{i\omega t_1}) \right| 0 \right\rangle \tag{145}$$
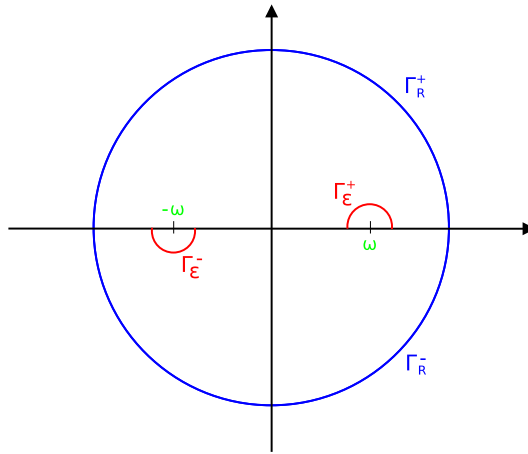
$$= \frac{1}{2\omega}e^{i\omega(t_1 - t_2)} \tag{146}$$

The two previous equations can be summarized into a single one

$$D(t_1, t_2) = \frac{1}{2\omega}e^{-i\omega|t_2 - t_1|} \tag{147}$$

Fig. 3. Contours in the complex time plane used to calculate the Feynman propagators of the one-dimensional harmonic oscillator. As usual, there are two poles on the real axis that can be avoided by going along the two small red circles. The final contour is closed by a large blue circle of radius R.

We see that the system is so simple that we obtain an explicit expression of the propagator in real space (in fact, in field theory, there is also an expression of the propagator in real space but it is not so often used because of its complexity – it is given in terms of ordinary and modified Bessel functions). Nevertheless, in order to compare with the standard approach and Eq. (101), it is interesting to obtain the expression of the propagator in Fourier space. The Fourier transform of a function $f(t)$ is defined by

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(E) \, e^{iEt} \, dE \tag{148}$$

The calculation is easy and well known. Here, for convenience, we quickly remind how it proceeds. For that purpose, let us consider the function

$$f(z) \equiv \frac{e^{-izt}}{z^2 - \omega^2} \tag{149}$$

It has two poles at $z = \pm\omega$ on the real axis, see Fig. 3. We assume that $t > 0$ and consider the following contour: along the real axis with two small circles $\Gamma_\epsilon^\pm$ around the two poles (the circle around $z = -\omega$ goes into the lower part of the complex plane while the circle around $z = +\omega$ goes into the upper part); and we complete by a big circle Γ_R^- going into the lower part of the complex plane. In this case, only the pole $z = +\omega$ is inside the contour $\Gamma = \mathbb{R} \cup \Gamma_\epsilon^\pm \cup \Gamma_R^-$, see Fig. 3. We have

$$\lim_{|z| \to \infty} \left| z f(z) \right| = \lim_{|z| \to \infty} \frac{|z| e^{\Im(z)t}}{z^2 - \omega^2} = 0 \tag{150}$$

since $\Im(z) < 0$ in the lower part of the complex plan. According to the first Jordan lemma, this means that the integral on the big circle vanishes. In the same way, we have $\lim_{z \to 0} |z f(z)| = 0$ and the second Jordan lemma tells that the integral on $\Gamma_\epsilon^\pm$ is also zero. Then, using the fact that $f(z) \equiv P(z)/Q(z)$ and that, in this case, the residue can be written as $\mathrm{Res}(f, z) = P(z)/Q'(z)$, one obtains

$$\mathrm{Res}(f, z = +\omega) = \frac{e^{-i\omega t}}{2\omega} \tag{151}$$

As a consequence

$$\lim_{R \to +\infty} \left[\frac{i}{2\pi} \int_\Gamma f(z) \, dz \right] = \frac{i}{2\pi} \int_{-\infty}^{\infty} f(z) \, dz$$

$$= \frac{i}{2\pi} \times -2i\pi \, \mathrm{Res}(f, z = +\omega)$$

$$= \frac{e^{-i\omega t}}{2\omega} \tag{152}$$

where the minus sign in the second line comes from the fact that the contour is clockwise (i.e. opposite to the direct direction).

Now if $t < 0$, then the big circle is taken in the upper plane. As a consequence, the pole inside the contour is now the one located at $z = -\omega$. Therefore,

$$
\begin{aligned}
\frac{i}{2\pi} \int_{\Gamma} f(z)\,\mathrm{d}z &= \frac{i}{2\pi} \int_{-\infty}^{\infty} f(z)\,\mathrm{d}z \\
&= \frac{i}{2\pi} \times +2i\pi\,\mathrm{Res}(f, z = -\omega) \\
&= \frac{\mathrm{e}^{i\omega t}}{2\omega}
\end{aligned}
\tag{153}
$$

The previous considerations allow us to conclude that

$$
\frac{i}{2\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{e}^{-iEt}}{E^2 - \omega^2}\,\mathrm{d}E = \frac{\mathrm{e}^{-i\omega|t|}}{2\omega} = D(t)
\tag{154}
$$

that is to say we have established the Fourier transform of the propagator. It should be obvious that the above equation is the counterpart of Eq. (101). As in field theory, we will show that it can be used to develop a perturbative method which allows us to take into account interactions.

5.2. Perturbation theory in quantum mechanics

Before discussing our main subject, we need to explain how the Feynman diagrams can be used in quantum mechanics [55,56]. Usually, in quantum mechanics, the perturbation theory is not based on Feynman diagrams [57]. This gives the impression that the techniques of field theory are very different. In fact, as we now demonstrate, these techniques can also be utilized in ordinary quantum mechanics [54] and, obviously (as could have been guessed from the very beginning), they lead to the same physical predictions. In the following, we will discuss this question and compare the two approaches in some details. We believe that this can improve our physical understanding of the vacuum energy problem.

In order to achieve the above mentioned task, we need to recall how interactions are treated in quantum mechanics. For this purpose, let us consider a system obeying the following Schrödinger equation:

$$
i\frac{\mathrm{d}|\Psi(t)\rangle_{\mathrm{S}}}{\mathrm{d}t} = H(p_{\mathrm{S}}, x_{\mathrm{S}})\big|\Psi(t)\big\rangle_{\mathrm{S}}
\tag{155}
$$

where $H(p_{\mathrm{S}}, x_{\mathrm{S}})$ is the Hamiltonian operator. The above equation is written in the Schrödinger representation, hence the subscript "S". Then, one can define the evolution operator $U(t, t_{\mathrm{i}})$ which relates the state vector at the time t to the state vector at some initial time $t_{\mathrm{i}} < t$ by the following expression:

$$
\big|\Psi(t)\big\rangle_{\mathrm{S}} = U(t, t_{\mathrm{i}})\big|\Psi(t_{\mathrm{i}})\big\rangle_{\mathrm{S}}
\tag{156}
$$

This operator obeys

$$
i\frac{\partial U(t, t_{\mathrm{i}})}{\partial t} = H U(t, t_{\mathrm{i}})
\tag{157}
$$

and, provided the Hamiltonian is not explicitly time dependent, it can also be written as

$$
U(t, t_{\mathrm{i}}) = \mathrm{e}^{-iH(t - t_{\mathrm{i}})}
\tag{158}
$$

Of course, in this representation, the operator p_{S} and x_{S} are time independent since only the state can change with time.

One the contrary, in the Heisenberg representation, the state vector does not evolve and can be expressed in terms of the Schrödinger state vector evaluated at some reference time $t = t_0$,

$$
|\Psi\rangle_{\mathrm{H}} \equiv \big|\Psi(t_0)\big\rangle_{\mathrm{S}}
\tag{159}
$$

Notice that the reference time is arbitrary and is not necessarily t_{i}. In this picture, the operators are time-dependent and, for instance, $x_{\mathrm{H}}(t)$ is given by

$$
x_{\mathrm{H}}(t) \equiv U^{\dagger}(t, t_0) x_{\mathrm{S}} U(t, t_0)
\tag{160}
$$

such that $x_{\mathrm{H}}(t_0) = x_{\mathrm{S}}$. In this picture, the time evolution of the operators is computed from the Heisenberg equation which reads

$$i\frac{dx_H}{dt} = i\frac{\partial U^\dagger(t,t_0)}{\partial t}x_S U(t,t_0) + iU^\dagger(t,t_0)x_S\frac{\partial U(t,t_0)}{\partial t} \tag{161}$$

$$= -U^\dagger H x_S U + U^\dagger x_S H U \tag{162}$$

$$= [x_S, H] \tag{163}$$

where in the second line we have used that $-i\partial U^\dagger/\partial t = U^\dagger H$.

Let us now consider the case where interactions are explicitly present. We assume that the corresponding Hamiltonian can be split into two parts, one corresponding to the free theory and one corresponding to the interactions

$$H = H_0 + H_{\text{int}} \tag{164}$$

We also assume that H_{int} is a small perturbation. The free evolution, due to the free Hamiltonian H_0, is described by the evolution operator U_0. This operator is of course different from U which describes the full evolution with the interactions taken into account. Clearly, U_0 can be expressed as

$$U_0(t, t_i) = e^{-iH_0(t-t_i)} \tag{165}$$

Then the idea is to factor out the free evolution (which is supposed to be known) and to focus on the time evolution due to the perturbation on top of the free evolution. For this purpose, we define the interaction representation by (again, the representations are supposed to coincide when $t = t_0$)

$$x_I(t) = U_0^\dagger(t,t_0)x_S U_0(t,t_0) \tag{166}$$

$$= e^{iH_0(t-t_0)}x_S e^{-iH_0(t-t_0)} \tag{167}$$

Clearly $x_I(t)$ is different from $x_H(t)$ because its time evolution is "generated" only by the free Hamiltonian and not by the total one. Let us now calculate the equation of motion of $x_I(t)$. One has

$$\frac{dx_I}{dt} = iH_0 x_I + e^{iH_0(t-t_0)}x_S \times -iH_0 e^{-iH_0(t-t_0)} \tag{168}$$

$$= i[H_0, x_I] \tag{169}$$

Therefore, the time evolution of $x_I(t)$ is controlled by the free Hamiltonian. This evolution is supposed to be known. In the same manner, we define the state vector in the interaction picture by

$$\left|\Psi(t)\right\rangle_I = U_0^\dagger(t,t_0)\left|\Psi(t)\right\rangle_S \tag{170}$$

As a consequence, in the interaction picture, both the state vectors and the operators are time dependent. The equation obeyed by the state vector is given by

$$i\frac{d|\Psi\rangle_I}{dt} = -H_0|\Psi\rangle_I + e^{iH_0(t-t_0)}i\frac{d|\Psi\rangle_S}{dt} \tag{171}$$

$$= -H_0 e^{iH_0(t-t_0)}|\Psi\rangle_S + e^{iH_0(t-t_0)}(H_0 + H_{\text{int}})|\Psi\rangle_S \tag{172}$$

$$= e^{iH_0(t-t_0)}H_{\text{int}}|\Psi\rangle_S \tag{173}$$

$$= e^{iH_0(t-t_0)}H_{\text{int}}e^{-iH_0(t-t_0)}|\Psi\rangle_I \tag{174}$$

$$= H_I|\Psi\rangle_I \tag{175}$$

where $H_I = U_0^\dagger(t,t_0)H_{\text{int}}U_0(t,t_0)$ in agreement with the definition (166). The time evolution of $|\Psi\rangle_I$ is therefore generated only by the interacting Hamiltonian. Then, it is natural to define the evolution operator in the interaction picture. This operator is associated with the Schrödinger equation (175). Its definition reads

$$\left|\Psi(t)\right\rangle_I = U_I(t,t_i)\left|\Psi(t_i)\right\rangle_I \tag{176}$$

Of course U_I is different from U_0 (and from U). A straightforward calculation shows that it can also be expressed as

$$U_I(t,t_i) = T\exp\left[-i\int_{t_i}^{t}d\tau\,H_I(\tau)\right] \tag{177}$$

Another expression for U_I is based on the following considerations. From Eqs. (176) and (170), we have

$$\left|\Psi(t)\right\rangle_I = U_I(t,t_i)\left|\Psi(t_i)\right\rangle_I \tag{178}$$

$$= U_I(t,t_i)U_0^\dagger(t_i,t_0)\left|\Psi(t_i)\right\rangle_S \tag{179}$$

$$= U_0^\dagger(t,t_0)\left|\Psi(t)\right\rangle_S \tag{180}$$

Therefore, one obtains

$$\big|\Psi(t)\big\rangle_S = U_0(t,t_0)U_I(t,t_i)U_0^\dagger(t_i,t_0)\big|\Psi(t_i)\big\rangle_S \tag{181}$$

$$= U(t,t_i)\big|\Psi(t_i)\big\rangle_S \tag{182}$$

from which we deduce that

$$U(t,t_i) = U_0(t,t_0)U_I(t,t_i)U_0^\dagger(t_i,t_0) \tag{183}$$

or

$$U_I(t,t_i) = U_0^\dagger(t,t_0)U(t,t_i)U_0^\dagger(t_i,t_0) \tag{184}$$

$$= e^{iH_0(t-t_0)}e^{-iH(t-t_i)}e^{-iH_0(t_i-t_0)} \tag{185}$$

This expression is very natural since it expresses nothing but the fact that the full evolution of the system is just the combination of the free evolution and of the evolution due to the interacting term. Let us also remark that one can write

$$x_I(t) = U_0^\dagger(t,t_0)x_S U_0(t,t_0) \tag{186}$$

$$= U_0^\dagger(t,t_0)U(t,t_0)x_H U^\dagger(t,t_0)U_0(t,t_0) \tag{187}$$

$$= e^{iH_0(t-t_0)}e^{-iH(t-t_0)}x_H(t)e^{iH(t-t_0)}e^{-iH_0(t-t_0)} \tag{188}$$

$$= U_I(t,t_0)x_H(t)U_I^\dagger(t,t_0) \tag{189}$$

Again, this result is very intuitive and expresses the fact that, in the interacting picture, the free evolution of the operators is factorized out.

After these preliminaries, one can now derive the Gell-Mann Law equation which is at the heart of the perturbative treatment presented in what follows. Let $|n_I\rangle$ be the energy eigenstates of the full Hamiltonian, i.e. the free Hamiltonian plus the Hamiltonian describing the interactions. We denote the true vacuum state by $|0_I\rangle \equiv |\Omega\rangle$, in agreement with the conventions adopted in the previous subsections. Then, one can always expand the vacuum state of the free theory according to

$$|0\rangle = \sum_{n=0}^\infty |n_I\rangle\langle n_I|0\rangle \tag{190}$$

$$= |\Omega\rangle\langle\Omega|0\rangle + \sum_{n\neq 0}^\infty |n_I\rangle\langle n_I|0\rangle \tag{191}$$

As a consequence, applying the operator $e^{-iH(t+t_0)}$ to the above relation, one obtains

$$e^{-iH(t+t_0)}|0\rangle = e^{-iE_0(t+t_0)}|\Omega\rangle\langle\Omega|0\rangle$$
$$+ \sum_{n\neq 0}^\infty e^{-iE_n(t+t_0)}|n_I\rangle\langle n_I|0\rangle \tag{192}$$

Then, we take the limit $t \to \infty(1-i\epsilon)$. This kills all the terms on the right-hand side but the first one. Indeed, we have $E_n > E_0$ ($n \neq 0$) and this means that $e^{-iE_0 t}$ goes to zero less rapidly than all the other exponential terms. Therefore, in this limit, Eq. (192) reads $e^{-iH(t+t_0)}|0\rangle \simeq e^{-iE_0(t+t_0)}|\Omega\rangle\langle\Omega|0\rangle$ and, by inverting this formula, this allows us to express the interacting vacuum as

$$|\Omega\rangle = \lim_{t\to\infty(1-i\epsilon)} \frac{e^{iE_0(t+t_0)}}{\langle\Omega|0\rangle}e^{-iH(t+t_0)}|0\rangle \tag{193}$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{e^{iE_0(t+t_0)}}{\langle\Omega|0\rangle}e^{-iH(t+t_0)}e^{iH_0(t+t_0)}|0\rangle$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{e^{iE_0(t+t_0)}}{\langle\Omega|0\rangle}U_I(t_0,-t)|0\rangle \tag{194}$$

In the second line, we have used the fact that $H_0|0\rangle = 0$ since we can always define the ground state as the state associated to the zero of energy. This means that $e^{iH_0(t+t_0)}|0\rangle = |0\rangle$ and explains the appearance of the new factor $e^{iH_0(t+t_0)}$. In the same fashion, using $e^{iH(t-t_0)}|\Omega\rangle = e^{iE_0(t-t_0)}|\Omega\rangle$, one has

$$\langle 0|e^{-iH(t-t_0)} = \lim_{t\to\infty(1-i\epsilon)} \langle 0|\Omega\rangle\langle\Omega|e^{-iE_0(t-t_0)} \tag{195}$$

from which one deduces that

$$\langle\Omega| = \lim_{t\to\infty(1-i\epsilon)} \frac{e^{iE_0(t-t_0)}}{\langle 0|\Omega\rangle} \langle 0|e^{-iH(t-t_0)} \tag{196}$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{e^{iE_0(t-t_0)}}{\langle 0|\Omega\rangle} \langle 0|e^{iH_0(t-t_0)}e^{-iH(t-t_0)}$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{e^{iE_0(t-t_0)}}{\langle 0|\Omega\rangle} \langle 0|U_{\rm I}(t,t_0) \tag{197}$$

Therefore, the norm of the vacuum state is given by

$$\langle\Omega|\Omega\rangle = \lim_{t\to\infty(1-i\epsilon)} \frac{e^{2iE_0t}}{|\langle 0|\Omega\rangle|^2} \langle 0|U_{\rm I}(t,t_0)U_{\rm I}(t_0,-t)|0\rangle \tag{198}$$

In the following, we take $\langle\Omega|\Omega\rangle = 1$ although this choice is not mandatory (but, then, when we compute a correlation function, it would be necessary to divide the corresponding expression by $\langle\Omega|\Omega\rangle$).

We are now in a position where one can calculate the two-point correlation function of the field operator (recall that $x(t)$ is directly proportional to the field operator, see Eq. (132)). Using the previous results, one obtains

$$\langle\Omega|x_{\rm H}(t_2)x_{\rm H}(t_1)|\Omega\rangle = \lim_{t\to\infty(1-i\epsilon)} \frac{e^{2iE_0t}}{|\langle 0|\Omega\rangle|^2} \langle 0|U_{\rm I}(t,t_0)U_{\rm I}^\dagger(t_2,t_0)x_{\rm I}(t_2)U_{\rm I}(t_2,t_0)U_{\rm I}^\dagger(t_1,t_0)x_{\rm I}(t_1)U_{\rm I}(t_1,t_0)U_{\rm I}(t_0,-t)|0\rangle \tag{199}$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{\langle 0|U_{\rm I}(t,t_0)U_{\rm I}(t_0,t_2)x_{\rm I}(t_2)U_{\rm I}(t_2,t_0)U_{\rm I}(t_0,t_1)x_{\rm I}(t_1)U_{\rm I}(t_1,t_0)U_{\rm I}(t_0,-t)|0\rangle}{\langle 0|U_{\rm I}(t,t_0)U_{\rm I}(t_0,-t)|0\rangle} \tag{200}$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{\langle 0|U_{\rm I}(t,t_2)x_{\rm I}(t_2)U_{\rm I}(t_2,t_1)x_{\rm I}(t_1)U_{\rm I}(t_1,-t)|0\rangle}{\langle 0|U_{\rm I}(t,-t)|0\rangle} \tag{201}$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{\langle 0|{\rm T}[x_{\rm I}(t_2)x_{\rm I}(t_1)U_{\rm I}(t,-t)]|0\rangle}{\langle 0|U_{\rm I}(t,-t)|0\rangle} \tag{202}$$

$$= \lim_{t\to\infty(1-i\epsilon)} \frac{\langle 0|{\rm T}[x_{\rm I}(t_2)x_{\rm I}(t_1)\exp(-i\int_{-t}^{t}{\rm d}\tau\,H_{\rm I})]|0\rangle}{\langle 0|\exp(-i\int_{-t}^{t}{\rm d}\tau\,H_{\rm I})|0\rangle} \tag{203}$$

where, in order to go from the first to the second line, we have used Eq. (198) and the normalization $\langle\Omega|\Omega\rangle = 1$ to express the term $|\langle 0|\Omega\rangle|^2$. Eq. (203) is the Gell-Mann Law equation [16]. It will be the basis for our perturbative treatment of quantum mechanics. It plays the same role as Eq. (97) since it allows us to perturbatively evaluate the various correlation functions in terms of Feynman diagrams (the Gell-Mann Law equation can be generalized to higher correlation function if needed).

5.3. Quantum mechanics and bubble diagrams

In order to mimic the self-interacting field theory studied before, we now consider a quartic oscillator in quantum mechanics [55,56], i.e. a system where the potential is given by $V(x) \sim x^2 + x^4$. More precisely, this means that the interacting Hamiltonian can be expressed as

$$H_{\rm int} = \frac{\lambda}{4!}\Phi^4 \tag{204}$$

where we recall that Φ is given by Eq. (132). Let us then compute the two-point correlation function. For this purpose, as explained before, we use the Gell-Mann Law equation. The numerator reads

$$\left\langle 0\left|{\rm T}\left[\Phi(t_2)\Phi(t_1)\exp\left(-i\int{\rm d}\tau\,H_{\rm I}\right)\right]\right|0\right\rangle \tag{205}$$

where, for simplicity, we do not write the limits (but there will be reestablished when needed). It is also important to remember that the field are written in the interacting picture (and, again for simplicity we do not write the subscript "I", they can be considered as free fields). Obviously, one cannot calculate exactly this quantity so we do it perturbatively by expanding the exponential. At first order in the coupling constant λ, we obtain

$$\left\langle 0\left|{\rm T}\left\{\Phi(t_2)\Phi(t_1)\left[1-i\frac{\lambda}{4!}\int{\rm d}\tau\,\Phi^4(\tau)+\cdots\right]\right\}\right|0\right\rangle = \langle 0|{\rm T}[\Phi(t_2)\Phi(t_1)]|0\rangle$$

$$-i\frac{\lambda}{4!}\int{\rm d}\tau\,\langle 0|{\rm T}[\Phi(t_2)\Phi(t_1)\Phi^4(\tau)]|0\rangle+\cdots \tag{206}$$

$$= D(t_1, t_2) - i\frac{\lambda}{4!} \int d\tau \, \langle 0| \mathrm{T}\big[\Phi(t_2)\Phi(t_1)\Phi^4(\tau)\big]|0\rangle + \cdots \quad (207)$$

It should now be clear that the same expression can also be written in terms on Feynman diagrams. One arrives at the following expression

$$\left\langle 0\left|\mathrm{T}\left[\Phi(t_2)\Phi(t_1)\exp\left(-i\int d\tau\, H_I\right)\right]\right|0\right\rangle = \text{（diagram）} + \frac{3}{4!}\text{（diagram）} + \frac{12}{4!}\text{（diagram）} + \mathcal{O}(\lambda^2) \quad (208)$$

where we have defined the following Feynman rules [55,56]

$$D(t_1, t_2) \equiv \text{（diagram）} \quad (209)$$

$$-i\lambda \int d\tau \equiv \text{（diagram）} \quad (210)$$

Clearly, these Feynman rules are nothing but the one-dimensional version of the standard Feynman rules (105) and (106) of quantum field theory. The only difference is that four-dimensional space–time integrations are simply replaced with one-dimensional time integration. In fact, in terms of diagrams, the expressions in field theory or in quantum mechanics are exactly similar. In particular, we see the appearance of diagrams with loops and external legs (the tadpole) and also of diagrams with loops but no external legs, our bubble diagrams. The advantage of our approach now becomes clear. We are in a position where not only one can compute these diagrams in quantum mechanics but we can also investigate their deep meaning by comparing the result with the standard perturbative theory of quantum mechanics where everything is well under control.

Let us start with the tadpole diagram. Using Eqs. (209) and (210), the corresponding expression can be written as

$$\text{（diagram）} = -i\lambda \int d\tau \, D(t_1, \tau) D(\tau, \tau) D(\tau, t_2) \quad (211)$$

$$= -\frac{i\lambda}{8\omega^3} \int d\tau \, e^{-i\omega|t_1-\tau|-i\omega|\tau-t_2|} \quad (212)$$

$$= -\frac{i\lambda}{4\omega^2} D(t_2, t_1)\left(|t_2 - t_1| - \frac{i}{\omega}\right) \quad (213)$$

The most striking difference with the case of quantum field theory is that the tadpole diagram is now finite. In some sense, this is the signal that ordinary quantum mechanics does not need renormalization: the perturbative approach leads to finite result. Technically, this is due to the fact that we deal with one-dimensional integrals that have better convergence properties. Physically, this is due to the fact that a single quantum-mechanical oscillator leads to finite predictions while an infinite collections of such systems apparently lead to infinite physical quantities.

Now let us calculate the "bubble diagram" present in the second term of the above expansion (208). Having established that the tadpole diagram is now finite, it is particularly interesting to see what happens for this type of diagram. Using again the Feynman rules (209) and (210), it is straightforward to establish that

$$\text{（diagram）} = -i\lambda \int D^2(\tau, \tau)\, d\tau \quad (214)$$

$$= -i\frac{\lambda}{4\omega^2} \int d\tau = \infty \quad (215)$$

Therefore, even in quantum mechanics, the bubble diagram remains divergent. This shows that the divergent nature of the loop and bubble diagrams is, in some sense, different. They are both divergent in quantum field theory but one (loop diagrams with external legs) becomes finite in the limit of quantum mechanics while the other (bubble diagram) remains infinite.

But this now leads to the following question. We know from the standard perturbation theory of quantum mechanics that all the physical predictions are finite, even without renormalization. However, using the other approach based on the Feynman diagrams, we have just seen that it leads to terms that can be divergent. How these two facts can be consistent with each other? The answer is of course that, as in field theory, the bubble diagrams exactly cancel out in the calculation of the correlation functions. This can easily be seen if one notices that

$$\text{（diagram）} + \frac{3}{4!}\text{（diagram）} + \frac{12}{4!}\text{（diagram）} = \left(\text{（diagram）} + \frac{12}{4!}\text{（diagram）} + \cdots\right)$$

$$\times \left(1 + \frac{3}{4!} \, \infty \hspace{-0.5em}\bullet\hspace{-0.5em}\infty \right) + \mathcal{O}(\lambda^2) \tag{216}$$

and that

$$\left\langle 0 \left| \exp\left(-i \int \mathrm{d}\tau \, H_\mathrm{I}\right) \right| 0 \right\rangle = 1 + \frac{3}{4!} \, \infty \hspace{-0.5em}\bullet\hspace{-0.5em}\infty + \mathcal{O}(\lambda^2) \tag{217}$$

Therefore (as expected) the infinities of the bubble diagram never play a role in ordinary quantum mechanics. As a consequence, the two-point correlation function is given by

$$\langle \Omega | \Phi(t_2) \Phi(t_1) | \Omega \rangle = \bullet\!\!-\!\!-\!\!\bullet + \frac{12}{4!} \, \bullet\!\!-\!\!\overset{\bigcirc}{\bullet}\!\!-\!\!\bullet + \mathcal{O}(\lambda^2) \tag{218}$$

which is perfectly finite. Again, in terms of Feynman diagrams the above equation is strictly equivalent to the corresponding one in quantum field theory, see Eq. (124).

5.4. Quantum mechanics and vacuum energy

Let us now discuss how the vacuum energy is evaluated in ordinary quantum mechanics. The calculation is straightforward when there is no interaction and, from Eq. (138), one obtains

$$\langle H \rangle = \frac{\omega}{2} \tag{219}$$

This is of course nothing but the ground state energy of a harmonic oscillator. It cannot be zero because, due the Heisenberg uncertainty principle, the kinetic and potential energy cannot vanish at the same time. It can also be rewritten as

$$\langle H \rangle = \omega^2 D(0) \tag{220}$$

where we have used Eq. (147). This exactly corresponds to Eq. (113) with $m = \omega$ and $4 \to 1$ (since we now work in one dimension rather than in four), except that there is a sign difference which comes from the fact that, in field theory, we have used the signature $(-+++)$. Of course, the big difference is that $D(0)$ is a finite quantity in quantum mechanics. In some sense, the vacuum energy problem appears to be very simple, almost trivial. The ground state of a quantum-mechanical oscillator is finite but a quantum field is a system which can be described by an infinite collection of harmonic oscillators and, as a consequence, its ground state energy is infinite.

Let us now consider a case where an interaction is present as we did in Section 4.3. For simplicity, we consider again the case of a quartic oscillator with $V_\mathrm{int} = \lambda \Phi^4 / 4!$. Upon using the standard perturbation theory in quantum mechanics, one can calculate the displacement of the energy levels. They are given by the standard formula [57]

$$E_n = E_n^0 + \langle n | V_\mathrm{int} | n \rangle \tag{221}$$

where $|n\rangle$ are the unperturbed eigenvectors. It is easy to show that they can be expressed as

$$\Delta E_n \equiv E_n - E_n^0 = \frac{\lambda}{32\omega^2} \left(2n^2 + 2n + 1\right) \tag{222}$$

which means that, for the first two levels, the shift in energy is given by

$$\Delta E_0 = \frac{\lambda}{32\omega^2}, \qquad \Delta E_1 = \frac{5\lambda}{32\omega^2} \tag{223}$$

We can now use the other method, based on the Feynman diagrams. It is easy to obtain that

$$\Delta E_0 = \frac{\lambda}{4!} \langle \Phi^4 \rangle = \frac{3\lambda}{4!} \langle \Phi^2 \rangle^2 = \frac{\lambda}{8} D^2(0)$$
$$= \frac{i}{8 \int \mathrm{d}t} \, \infty \hspace{-0.5em}\bullet\hspace{-0.5em}\infty \tag{224}$$

Clearly, we have just obtained the strict equivalent of Eq. (123), the term $\int \mathrm{d}^4 x$ being replaced by $\int \mathrm{d}t$ as could have been guessed in the context of quantum mechanics. Then, using the fact that $D(0) = 1/(2\omega)$, one obtains that

$$\Delta E_0 = \frac{\lambda}{32\omega^2} \tag{225}$$

that is to say we recover the result established before, see Eq. (223). Therefore, we have shown explicitly that the two approaches lead to the same result. It is also interesting to notice that ΔE_0 is given by a bubble diagram but that this divergent graph is canceled by the infinite term $\int \mathrm{d}t$ so that a finite quantity is left.

The previous considerations also lead to the following question: before we had to deal with the renormalization of the mass. How does this aspect appear in the present context? In this case, we deal with the same diagrams, namely

$$\langle \Omega | T[\Phi(t_1)\Phi(t_2)] | \Omega \rangle = \bullet\!\!-\!\!\!-\!\!\!-\!\!\bullet + \frac{1}{2}\bullet\!\!-\!\!\bigcirc\!\!-\!\!\bullet + \cdots \tag{226}$$

or, in terms of explicit expressions [compare with Eqs. (124) and (125)]

$$\langle \Omega | T[\phi(t_1)\phi(t_2)] | \Omega \rangle = D_{\mathrm{F}}(t_1 - t_2) - \frac{i\lambda}{2} \int \mathrm{d}\tau\, D(t_1 - \tau) D_{\mathrm{F}}(0) D_{\mathrm{F}}(\tau - t_2) \tag{227}$$

$$= D_{\mathrm{F}}(t_1 - t_2) + \frac{i\lambda}{2} D_{\mathrm{F}}(0) \frac{1}{2\pi} \int \mathrm{d}E\, \frac{\mathrm{e}^{-iE(t_1 - t_2)}}{(E^2 - \omega^2)^2} \tag{228}$$

and this leads to

$$\omega_{\mathrm{ren}}^2 = \omega^2 + \frac{\lambda}{2} D(0) \tag{229}$$

which is exactly Eq. (129) (again, there is a sign difference which originates from the sign difference in the propagator, see above). In quantum mechanics, the propagator is finite and therefore

$$\omega_{\mathrm{ren}}^2 = \omega^2 + \frac{\lambda}{4\omega} \tag{230}$$

or

$$\omega_{\mathrm{ren}} = \omega + \frac{\lambda}{8\omega^2} \tag{231}$$

Then, one can express the spectrum in terms of the renormalized "mass" ω_{ren} as we have done in quantum field theory although this is not necessary here since the result is anyway finite. Straightforward calculations lead to

$$E_n = \omega_{\mathrm{ren}}\left(n + \frac{1}{2}\right) - \frac{\lambda}{32\omega_{\mathrm{ren}}^2}\left(2n^2 - 2n - 1\right) \tag{232}$$

It does not come as a surprised that, as a function of the renormalized mass, the spectrum is a different function of n, see Eq. (222). However, from the above expression one also obtains

$$E_1 - E_0 = \omega_{\mathrm{ren}} \tag{233}$$

In presence of an interaction, the difference between the first excited state and the fundamental level has been "modified" (or, rather, is expressed differently). But, observationally, we define or measure the spectrum with respect to $E_1 - E_0$. In other words what we define as ω is in fact $E_1 - E_0$. Therefore, renormalizing ω consists in preserving this definition even in presence of an interaction. Obviously, $E_1 - E_0$ is not the bare ω (the one which appears in the Lagrangian) when the an-harmonic term is present. What we do here is to define ω as the observed one, namely as $E_1 - E_0$, and then calculate the spectrum in terms of that observational quantity. What is done in quantum field theory is clearly exactly similar to the above described procedure except of course that the bare quantities are in fact infinite because $D_{\mathrm{F}}(0)$ is infinite. Therefore, this exercise illustrates nicely the deep meaning of renormalization.

This concludes our discussion of the vacuum energy density in terms of Feynman diagrams. We have shown that the cosmological constant is in fact given in terms of very peculiar graphs, the so-called bubble diagrams. These diagrams are, in a sense, more divergent than the usual loop diagrams because they remain infinite even in the limit of quantum mechanics contrary to the last ones. In non-gravitational physics, this is not a problem because the bubble diagrams always cancel out in the equations describing observable quantities. When the gravitational field is turned on, we face the tasks of renormalizing these diagrams which is more difficult than in the standard situation because of their bad behavior mentioned before. Given that astrophysical observations seem to indicate that the vacuum energy is non-vanishing (see below), the details of the renormalizing procedure become a crucial issue. Not only we have to extract a finite quantity from a divergent graph but this finite quantity must be in agreement with the observations. As argued before, the technique which consists in imposing a spatial cut-off is certainly not consistent (and, as well known, gives a very large contribution to the vacuum energy). A more reasonable approach seems to be dimensional regularization since this satisfies Lorentz invariance and is therefore consistent with the vacuum equation of state. As shown before, this typically leads to Eq. (95). In the next section, we present a new argument, based on the Gaussian effective potential, which seems to support this conclusion.

6. The Gaussian effective potential

The Gaussian effective potential is a non-perturbative approach to quantum field theory [58–66]. Originally developed in the context of quantum mechanics, it is has been generalized to field theory. Much less well known is the fact that the Gaussian effective potential method has something to say about the vacuum energy problem and the aim of this section is to present these considerations [59]. In the next subsection, we quickly recall the main idea and how this approach can be applied to ordinary quantum mechanics. Then, we explain how it can be implemented in quantum field theory and, finally, we discuss its implication for the cosmological constant problem.

6.1. The Gaussian effective potential in quantum mechanics

The main idea underlying the Gaussian effective potential approach is the following one [58]. At the classical level, a system is described by a potential. Once this one is specified, one can compute the evolution of the system. When quantum effects are taken into account, the behavior of the system will be modified and, by means of the formalism of quantum mechanics, one can work out the corresponding physical predictions. The idea of the Gaussian effective potential is to find an effective potential which, at the same time, can be used as if it were a "classical" potential (that is to say can be used in the framework of Newtonian dynamics) and takes into account the quantum effects. At first sight, this approach seems equivalent to what is known in the literature as the effective potential [67,68]. In this method, one requires the wave-function mean value to be centered at a fixed point Φ_0. Then, the effective potential gives the exact ground state energy of the system in presence of quantum corrections. However, in some situation, the result can be quite artificial. Consider for instance a situation where the wave-function possesses two large peaks on each side of Φ_0. In such a case, the effective potential does not necessarily lead to a correct description of what happens in the vicinity of Φ_0 because its shape could just be determined by some average procedure between the two distant peaks. Typically, one would obtain a convex potential centered at Φ_0 while our physical intuition rather indicates that the effective potential should be a double-well potential, the two wells corresponding to the two peaks of the wave-function. Another problem is that the standard effective potential is usually computed perturbatively in $\hbar$. Clearly, when the quantum corrections become important, this approach breaks down. Since the Gaussian effective potential does not suffer from the two above mentioned drawbacks, we will use it in this review. When possible, we will discuss the difference between the two approaches in order to stress their differences and similarities [69].

We now turn to the precise definition of the Gaussian effective potential. This definition is based on a variational principle. Here we consider a one-dimensional system as we have already done in Section 5.1. In particular, this means that the scalar field is given by Eq. (132). Then, the Gaussian effective potential is defined by [58]

$$V_{\mathrm{G}}(\Phi_0) = \min_{\Omega} \langle \Psi | H | \Psi \rangle \tag{234}$$

where H is the Hamiltonian of the system and where the wave-function is taken to be a simple Gaussian (hence the name Gaussian effective potential)

$$\Psi(\Phi) = \left(\frac{\Omega}{\hbar\pi} \right)^{1/4} e^{-\Omega(\Phi-\Phi_0)^2/(2\hbar)} \tag{235}$$

The parameter Ω is related to the width of the wave-function. Here the idea is to find the value of Ω which minimizes the energy of the system. In practice, the mean value of the Hamiltonian that appears in Eq. (234) can be expressed as

$$\langle H \rangle = \int_{-\infty}^{\infty} \mathrm{d}\Phi \, \Psi^*(\Phi) \left[-\frac{\hbar^2}{2} \frac{\mathrm{d}^2}{\mathrm{d}\Phi^2} + V(\Phi) \right] \Psi(\Phi) \tag{236}$$

To illustrate how the calculation of the Gaussian effective potential works on a concrete example, we choose the following potential

$$V(\Phi) = \frac{1}{2} m^2 \Phi^2 + \lambda \Phi^4 \tag{237}$$

that is to say the potential of an an-harmonic oscillator. Another motivation for this choice is that the comparison with field theory is easy since the prototypical model possesses the same potential. Upon using Eq. (235) in Eq. (236), one arrives at

$$\langle H \rangle(\Phi_0, \Omega) = \frac{\hbar\Omega}{4} + \frac{1}{2} m^2 \left(\Phi_0^2 + \frac{\hbar}{2\Omega} \right)$$
$$+ \lambda \left[\Phi_0^4 + 6\Phi_0^2 \frac{\hbar}{2\Omega} + \frac{3\hbar^2}{(2\Omega)^2} \right] \tag{238}$$

We must now determine the optimal Ω. As explained above, this is achieved by requiring $\partial \langle H \rangle / \partial \Omega = 0$, which leads to the equation

$$\Omega^3 - \left(m^2 + 12\lambda\Phi_0^2\right)\Omega - 6\hbar\lambda = 0 \tag{239}$$

This equation allows us to determine Ω in terms of Φ_0. In the following, we write this result as $\Omega_0 \equiv \Omega(\Phi_0)$. Upon using the above relation in the expression of $\langle H \rangle$, one obtains the following expression for the Gaussian effective potential

$$V_{\mathrm{G}}(\Phi_0) = \frac{1}{2}m^2\Phi_0^2 + \lambda\Phi_0^4 + \frac{\hbar\Omega_0}{2} - \frac{3\hbar^2\lambda}{4\Omega_0} \tag{240}$$

Of course, as already emphasized before, one should not forget that Ω_0 is a function of Φ_0.

A complete study of the above potential has been provided in Ref. [58] and we will not repeat this analysis. Here, we just want to illustrate that the Gaussian effective potential is an accurate method to estimate the ground state of a system (even in the strong coupling limit) since this is directly relevant for the cosmological constant problem. Let us consider the case where $m^2 > 0$ (the case of the double-well potential, reminiscent of symmetry breaking in field theory has been studied in Ref. [58]). The ground state of the an-harmonic oscillator is given by Eq. (240)

$$V_{\mathrm{G}}(\Phi_0 = 0) = \frac{\hbar\Omega_0}{2}\left(1 - \frac{3}{2}\frac{\hbar\lambda}{\Omega_0^3}\right) \tag{241}$$

where Ω_0 is solution of

$$\Omega^3 - m^2\Omega - 6\hbar\lambda = 0 \tag{242}$$

The discriminant of this cubic equation can be expressed as

$$\Delta(\Phi_0 = 0) = 4m^6\left(1 - \frac{972}{16}\frac{1}{\xi^6}\right) \tag{243}$$

where the quantity ξ in the above relation is given by

$$\xi^2 \equiv \frac{m^2}{(2\lambda)^{2/3}} \tag{244}$$

The awkward coefficient 972/16 in Eq. (243) originates from our will to work with the definition of ξ used in Ref. [58]. The cubic equation (242) can be explicitly solved. We deal with different branches according to the sign of the discriminant. This sign is determined by $\xi_{\mathrm{lim}} = (972/16)^{1/6} \simeq 1.9827$. If $\xi < \xi_{\mathrm{lim}}$ ($\xi = 0$ represent the quartic oscillator) we are in the strong coupling regime which is of particular interest in order to demonstrate the usefulness of the Gaussian effective potential. Therefore, let us focus on this situation. In this case, the exact solution of Eq. (242) can be written as

$$\Omega_0 = m\left[\frac{3}{2\xi^3} - \frac{1}{27}\sqrt{27\left(\frac{972}{16}\frac{1}{\xi^6} - 1\right)}\right]^{1/3} + m\left[\frac{3}{2\xi^3} + \frac{1}{27}\sqrt{27\left(\frac{972}{16}\frac{1}{\xi^6} - 1\right)}\right]^{1/3} \tag{245}$$

Together with Eq. (241), the above equation gives an excellent approximation for the ground state. For instance, if $m = 1$ and $\lambda = 10$, then $\xi^2 \simeq (1/20)^{2/3} \simeq 0.1357$ and $\Omega_0 = 4$. As a consequence, this gives $V_{\mathrm{G}}(\Phi_0 = 0) \simeq 1.53$, a result that is accurate at $\simeq 1.75\%$ according to Ref. [58].

This simple example has shown that the Gaussian effective potential can be an efficient tool to calculate the ground state of a quantum system even in a non-perturbative regime. Therefore, it appears as an interesting approach for the vacuum energy problem. However, before turning to the calculation of the Gaussian effective potential in field theory, it is interesting to compare it with the standard effective potential in more detail. This is the purpose of the next subsection.

6.2. Comparison with the one loop effective potential

In order to explain how the effective potential is obtained, we need to quickly return to the basics of quantum field theory. In Eq. (96), we have defined the generating functional $Z[J]$. Let us now define another generating functional $W[J]$ by [16,17]

$$Z[J] = e^{iW[J]} \tag{246}$$

Following the procedure in Eq. (97), one can define new correlation functions $G_{\mathrm{c}}^{(n)}$ according to

$$G_{\mathrm{c}}^{(n)} = \left(\frac{1}{i}\right)^{n-1}\frac{\delta^n W[J]}{\delta J(x_1)\cdots\delta J(x_n)}\bigg|_{J=0} \tag{247}$$

As is well known, these functions represent in fact the connected n-points functions. Then, one introduces the classical field Φ_{c}

$$\Phi_c \equiv \frac{\delta W[J]}{\delta J(x)}\bigg|_{J=0} \tag{248}$$

where we remind that $J(x)$ is a source. This definition can be easily justified. Indeed, from Eqs. (96) and (246), one sees that

$$\Phi_c = \frac{1}{iZ[J]}\frac{\delta Z[J]}{\delta J(x)} \tag{249}$$

$$= \frac{\mathcal{N}\int\mathcal{D}\Phi\exp\{i\int\mathrm{d}^4x[\mathcal{L}+J(x)\Phi(x)]\}\Phi(x)}{\mathcal{N}\int\mathcal{D}\Phi\exp\{i\int\mathrm{d}^4x[\mathcal{L}+J(x)\Phi(x)]\}}$$

$$= \frac{\langle\Omega|\Phi(x)|\Omega\rangle}{\langle\Omega|\Omega\rangle} \tag{250}$$

Therefore, we see that the classical field is in fact the vacuum expectation value of the field operator in the vacuum state of the theory. When the source vanishes, the classical field does not necessarily goes to zero, a typical and well-known example being of course symmetry breaking.

The next step consists in introducing a new generating function, the so-called effective action, defined by

$$\Gamma[\Phi_c] = W[J] - \int\mathrm{d}^4x\,J(x)\Phi_c(x) \tag{251}$$

Clearly, $\Gamma[\Phi_c]$ does not depend on the source since $\delta\Gamma/\delta J = 0$. Let us now calculate the functional derivative of Γ with respect with the classical field. This gives

$$\frac{\delta\Gamma[\Phi_c]}{\delta\Phi_c(x)} = \frac{\delta W}{\delta\Phi_c(x)} - \int\mathrm{d}^4y\frac{\delta J(y)}{\delta\Phi_c(x)}\Phi_c(y) - \int\mathrm{d}^4y\,J(y)\frac{\delta\Phi_c(y)}{\delta\Phi_c(x)} \tag{252}$$

$$= \int\mathrm{d}^4y\frac{\delta W}{\delta J(y)}\frac{\delta J(y)}{\delta\Phi_c(x)} - \int\mathrm{d}^4y\frac{\delta J(y)}{\delta\Phi_c(x)}\Phi_c(y) - J(x) = -J(x) \tag{253}$$

In order to interpret these equations and to better understand their meaning, it is interesting to apply them to the case of a free theory where all the calculations can be carried out explicitly. From Eq. (99), we immediately see that (the subscript zero indicates that the calculation is performed for the free theory)

$$iW_0[J] = -\frac{1}{2}\int\mathrm{d}^4x\,\mathrm{d}^4y\,J(x)D_F(x-y)J(y) \tag{254}$$

Upon using Eq. (248), one obtains

$$\Phi_c(z) = -\frac{1}{i}\int\mathrm{d}^4x\,D_F(z-x)J(x) \tag{255}$$

But iD_F is the Green function of the Klein–Gordon operator. As a consequence, the above equation implies that

$$\left(\eta^{\mu\nu}\partial_\mu\partial_\nu - m^2\right)\Phi_c(x) = J(x) \tag{256}$$

i.e. the classical field obeys the classical equation of motion (hence its name). Then, using Eqs. (254) and (255) into the definition (251), the effective action can be expressed as

$$\Gamma_0[\Phi_c] = \frac{1}{2i}\int\mathrm{d}^4x\,\mathrm{d}^4y\,J(x)D_F(x-y)J(y) \tag{257}$$

The next step is to use the equation of motion of the classical field (256) to rewrite the effective action as

$$\Gamma_0[\Phi_c] = \frac{1}{2i}\int\mathrm{d}^4x\,\mathrm{d}^4y\,D_F(x-y)\left(\eta^{\mu\nu}\partial_\mu^x\partial_\nu^x - m^2\right)\Phi_c(x)\left(\eta^{\mu\nu}\partial_\mu^y\partial_\nu^y - m^2\right)\Phi_c(y) \tag{258}$$

where $\partial_\mu^{x,y}$ means a partial derivative with respect to the coordinates at point x and y. Then, integrating by part the x dependent terms in the above integral leads to

$$\Gamma_0[\Phi_c] = \frac{1}{2i}\int\mathrm{d}^4x\,\mathrm{d}^4y\,\Phi_c(x)\left(\eta^{\mu\nu}\partial_\mu^x\partial_\nu^x - m^2\right)D_F(x-y)\left(\eta^{\mu\nu}\partial_\mu^y\partial_\nu^y - m^2\right)\Phi_c(y) \tag{259}$$

Finally, using the fact that D_F/i is the Green function satisfying $-\eta^{\mu\nu}\partial_\mu\partial_\nu D_F + m^2 D_F = \delta$, one obtains

$$\Gamma_0[\Phi_c] = \frac{1}{2}\int\mathrm{d}^4x\,\Phi_c(x)\left(\eta^{\mu\nu}\partial_\mu\partial_\nu - m^2\right)\Phi_c(x) = -\frac{1}{2}\int\mathrm{d}^4x\left(\eta^{\mu\nu}\partial_\mu\Phi_c\partial_\nu\Phi_c + m^2\Phi_c\right) \tag{260}$$

and one recovers the action of a free scalar field theory. In presence of interactions, $\Gamma[\Phi_c]$ will no longer agree with the classical action. The quantum corrections will transform it into a complicated non-local functional. In this situation, one typically expects an expression that can be written as

$$\Gamma[\Phi_c] = -\int d^4x \left[\frac{1}{2} A(\Phi_c)\eta^{\mu\nu}\partial_\mu\Phi_c\partial_\nu\Phi_c + V_{\rm eff}(\Phi_c) + B(\Phi_c)\left(\eta^{\mu\nu}\partial_\mu\Phi_c\partial_\nu\Phi_c\right)^4 + \cdots \right] \tag{261}$$

where $A(\Phi_c)$ and $B(\Phi_c)$ are functions (not functionals) of Φ_c and the dots represent higher derivative terms. Clearly, the previous considerations justify the name "effective action" for $\Gamma[\Phi_c]$.

Another interesting aspect is that the effective action is also the generating functional of the irreducible correlation functions $\Gamma^{(n)}$ (also known as proper vertices). In other words, one has

$$\Gamma^{(n)}(x_1, \ldots, x_n) = \frac{\delta^n \Gamma[\Phi_c]}{\delta\Phi_c(x_1)\cdots\delta\Phi_c(x_n)}\bigg|_{\Phi_c=0} \tag{262}$$

Equivalently, one can express the effective action as a Volterra expansion, namely

$$\Gamma[\Phi_c] = \sum_{n=0}^\infty \frac{1}{n!} \int d^4x_1\cdots d^4x_n\, \Gamma^{(n)}(x_1, \ldots, x_n)\Phi_c(x_1)\cdots\Phi_c(x_n) \tag{263}$$

which is nothing but another way to write the functional (261).

Now, suppose that we are interested in determining the effective action concretely. It is clear that this is a complicated calculation. As a first step, it is worth calculating the effective potential in Eq. (261). It is true that it does not represent all the information contained in the effective action but this would allow us to determine the location of the minimum of the system in presence of quantum corrections. Therefore, this is an interesting quantity. For this purpose, it is sufficient to consider a case where Φ_c is constant. Indeed, in this situation, the effective action reduces to $\Gamma = -V V_{\rm eff}(\Phi_c)$, where V is the space–time volume. On the other hand, the Volterra expansion can be re-expressed as

$$\Gamma[\Phi_c] = \sum_{n=0}^\infty \frac{1}{n!} \int d^4x_1\cdots d^4x_n\, \Gamma^{(n)}(x_1, \ldots, x_n)\Phi_c^n \tag{264}$$

This expression can be further simplified if one takes the Fourier transform of the proper vertex,

$$\Gamma^{(n)}(x_1, \ldots, x_n) = \int \frac{d^4k_1}{(2\pi)^4}\cdots\frac{d^4k_n}{(2\pi)^4}\,\Gamma^{(n)}(k_1, \ldots, k_n)e^{-ik_1\cdot x_1}\cdots e^{-ik_n\cdot x_n} \tag{265}$$

In fact, for convenience, we introduce the coefficients $\overline{\Gamma}^{(n)}(k_1, \ldots, k_n)$ that are easier to manipulate. They are defined by the following expression

$$\Gamma^{(n)}(k_1, \ldots, k_n) \equiv (2\pi)^4\delta(k_1 + \cdots + k_n)\overline{\Gamma}^{(n)}(k_1, \ldots, k_n) \tag{266}$$

Then, we insert the Fourier expansion (265) into the expression (264) of the effective action. The integrals over space of the exponentials lead to Dirac functions in momentum. As a consequence, one arrives at the following expression

$$\Gamma[\Phi_c] = V \sum_{n=0}^\infty \frac{1}{n!} \overline{\Gamma}^{(n)}(\boldsymbol{k}_i = 0)\Phi_c^n \tag{267}$$

where, as already mentioned, $V = (2\pi)^4\delta(0)$ is the space–time volume. One finally reaches the result that

$$V_{\rm eff}(\phi_c) = -\sum_{n=0}^\infty \frac{1}{n!} \overline{\Gamma}^{(n)}(\boldsymbol{k}_i = 0)\Phi_c^n \tag{268}$$

Therefore, in order to calculate the effective potential, one just has to evaluate the proper vertices with vanishing momenta. At one loop, this gives rise to the following sum

$$V_{\rm eff} = \qquad\qquad\qquad\qquad\qquad\qquad\qquad + \cdots \tag{269}$$

It turns out that this series can be summed up as we are now going to show.

In order to see how it works, let us calculate the second diagram in the above expansion. For this purpose, we start with the following diagram

$$\mathcal{D} = \text{} \tag{270}$$

According to the Feynman rules (105) and (106), this diagram is equivalent to the expression

$$\mathcal{D} = (-i\lambda)^2 \int \mathrm{d}u\, \mathrm{d}v\, D(t_1 - u)D(t_2 - u)D^2(u - v)D(v - t_3)D(v - t_4) \tag{271}$$

where u and v denotes the two internal vertices. Upon using Eq. (154), this expression takes the form (here, we use the quantum-mechanical theory introduced in Section 5.1 where the "mass" is denoted ω)

$$\mathcal{D} = (-i\lambda)^2 \int \mathrm{d}u\, \mathrm{d}v \int \frac{\mathrm{d}E_1}{2\pi}\frac{\mathrm{d}E_2}{2\pi}\frac{\mathrm{d}E_3}{2\pi}\frac{\mathrm{d}E_4}{2\pi}\frac{\mathrm{d}E}{2\pi}\frac{\mathrm{d}\tilde{E}}{2\pi} \frac{i e^{-iE_1(t_1-u)}}{E_1^2 - \omega^2}\frac{i e^{-iE_2(t_2-u)}}{E_2^2 - \omega^2}\frac{i e^{-iE_3(v-t_3)}}{E_3^2 - \omega^2}$$
$$\times \frac{i e^{-iE_4(v-t_4)}}{E_4^2 - \omega^2}\frac{i e^{-iE(u-v)}}{E^2 - \omega^2}\frac{i e^{-i\tilde{E}(u-v)}}{\tilde{E}^2 - \omega^2} \tag{272}$$

The integrals over u and v are easily performed and leads to two Dirac functions. Then, a further integration over $\tilde{E}$ can be performed and we are left with the expression

$$\mathcal{D} = (-i\lambda)^2 (2\pi)^2 \delta(E_1 + E_2 - E_3 - E_4) \int \frac{\mathrm{d}E_1}{2\pi}\frac{\mathrm{d}E_2}{2\pi}\frac{\mathrm{d}E_3}{2\pi}\frac{\mathrm{d}E_4}{2\pi}\frac{\mathrm{d}E}{2\pi} \frac{i e^{-iE_1 t_1}}{E_1^2 - \omega^2}\frac{i e^{-iE_2 t_2}}{E_2^2 - \omega^2}\frac{i e^{-iE_3 t_3}}{E_3^2 - \omega^2}\frac{i e^{-iE_4 t_4}}{E_4^2 - \omega^2}$$
$$\times \frac{i}{E^2 - \omega^2}\frac{i}{(E_3 + E_4 - E)^2 - \omega^2} \tag{273}$$

In order to compute the expression (269) (or rather the contribution of the second diagram to this expression), we have seen that we must amputate the diagram by appending the external leg propagators. This amounts to ignore the factors $i/(E_i^2 - \omega^2)$ in the above expression. Moreover, since the graph is given by the Fourier transform of $\mathcal{D}$, this means that only the expression inside the above quadruple integrals over E_i should be considered. Finally, the "momentum" (in the present context the energy) of the external legs should be put to zero (i.e. we require all the E_i's left in the expression of $\mathcal{D}$ to vanish) and we must factorize out a Dirac function expressing energy conservation [times $(2\pi)^2$ since we are considering a $0 + 1$ theory]. This means that Eq. (273) takes the following form

$$\text{} = (-i\lambda)^2 \int \frac{\mathrm{d}E}{2\pi}\frac{i}{E^2 - \omega^2}\frac{i}{E^2 - \omega^2} \tag{274}$$

The fact that the propagator appears squared is of course not fortuitous. It comes from the fact that we have two "double legs" attached to the loop. For n "double legs", it is clear that we would have the propagator to the power n and an overall factor $(-i\lambda)^n$. As a consequence, at one loop, Eq. (268) can be expressed as

$$V_{\mathrm{eff}}(\Phi_c) = -\sum_{p=1}^{\infty} \frac{\Phi_c^{2p}}{2p}\left(\frac{\lambda 4!}{2}\right)^p \int \frac{\mathrm{d}E}{2\pi}\frac{1}{(E^2 - \omega^2)^p} \tag{275}$$

where we have taken into account a symmetry factor and where we have noted $n = 2p$, p being the number of external points. The sum can now be performed and we obtain

$$V_{\mathrm{eff}}(\Phi_c) = -\frac{1}{2}\int_{-\infty}^{\infty} \frac{\mathrm{d}E}{2\pi}\ln\left(1 + \frac{12\lambda\Phi_c^2}{E^2 - \omega^2}\right) \tag{276}$$

$$= \frac{1}{2}\left(\omega - \sqrt{12\lambda\Phi_c^2 - \omega^2}\right) \tag{277}$$

$$= \frac{\hbar}{2}\left(-m + \sqrt{m^2 + 12\lambda\Phi_c^2}\right) \tag{278}$$

where we have re-established $\hbar$ and where we have taken into account the sign difference for ω (or m) due to the fact that we work with a $(- + + +)$ signature, see the remark after Eq. (220). This expression should be compared to Eq. (240). In fact the first term just comes from the fact that the method of the effective potential assumes that $V_{\mathrm{eff}}(0) = 0$. Therefore, in order to compare with the Gaussian effective potential, one must ignore this rescaling. As a consequence, the total potential (including the leading order in $\hbar^0$) reads

$$V_{\mathrm{eff}}(\Phi_c) = \frac{1}{2}m^2\Phi_c^2 + \frac{\hbar}{2}\sqrt{m^2 + 12\lambda\Phi_c^2} \tag{279}$$

One sees that the vacuum energy is predicted to be $V_{\text{eff}}(0) = \hbar m/2$ a result which is not realistic in the non-perturbative regime.

In fact the comparison can be made in an even more explicit manner if one remembers that the one loop Gaussian effective potential consists in retaining the dominant term is a systematic expansion in $\hbar$. As a consequence, one should ignore the $\hbar^2$ term in Eq. (238) and the $\hbar$ term in Eq. (239). In particular, this last approximation leads to the following explicit solution for the gap equation

$$\overline{\Omega}^2 = m^2 + 12\lambda\Phi_0^2 = \frac{\mathrm{d}^2 V}{\mathrm{d}\Phi_0^2} \tag{280}$$

If we then use Eq. (240) and neglects the $\hbar^2$ term in that equation, one obtains

$$V_{\text{G}}(\Phi_0) = V(\Phi_0) + \frac{\hbar}{2}\sqrt{\frac{\mathrm{d}^2 V}{\mathrm{d}\Phi_0^2}} + \mathcal{O}(\hbar^2) \tag{281}$$

$$= \frac{1}{2}m^2\Phi_0^2 + \frac{\hbar}{2}\sqrt{m^2 + 12\lambda\Phi_0^2} + \cdots \tag{282}$$

that is to say exactly Eq. (279). We conclude that the effective potential is in fact contained in the Gaussian effective potential approach. It just consists in neglecting higher power of $\hbar$ since this is a one loop approach. As a consequence, the effective potential cannot be reliable in the non-perturbative regime where the quantum effects are strong.

Having demonstrated that the Gaussian effective potential approach is an efficient tool, we now need to study how it works in quantum field theory. Then, we will able to use it in order to calculate the energy density of the vacuum state.

6.3. The Gaussian effective potential in field theory and the vacuum energy

We now turn to the main question of this section, namely the calculation of the Gaussian effective potential in quantum field theory [59,60,63,65]. For this purpose, we now consider a model similar to the one considered in the previous subsections, namely a free model plus a quartic self-interaction

$$\mathcal{H} = \frac{1}{2}\dot{\Phi}^2 + \frac{1}{2}\delta^{ij}\partial_i\Phi\partial_j\Phi + \frac{1}{2}m_{\text{B}}^2\Phi^2 + \lambda_{\text{B}}\Phi^4 \tag{283}$$

where m_{B} and λ_{B} are the bare mass and coupling constant, respectively. Then, we write the field operator as (we work in d space–time dimensions)

$$\Phi(t, \boldsymbol{x}) = \Phi_0 + \frac{1}{(2\pi)^{(d-1)/2}} \int \frac{\mathrm{d}^{d-1}\boldsymbol{k}}{\sqrt{2\omega_\Omega(k)}} \left(c_{\boldsymbol{k}}\mathrm{e}^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + c_{\boldsymbol{k}}^\dagger \mathrm{e}^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}}\right) \tag{284}$$

with

$$\omega_\Omega(k) \equiv \sqrt{k^2 + \Omega^2} \tag{285}$$

The quantity Φ_0 represents a classical and constant field. The parameter Ω denotes the mass of the excitations around this classical value. The idea at the basis of the calculation is the same as before, namely minimizing the energy with respect to the parameter Ω. Using the properties of the creation and annihilation operators, it is straightforward to calculate the mean value of the Hamiltonian in the vacuum state. We find

$$\langle\mathcal{H}\rangle(\Omega, \Phi_0) = I_1(\Omega) + \frac{1}{2}(m_{\text{B}}^2 - \Omega^2)I_0(\Omega) + \frac{1}{2}m_{\text{B}}^2\Phi_0^2$$
$$+ \lambda_{\text{B}}\Phi_0^4 + 6\lambda_{\text{B}}\Phi_0^2 I_0(\Omega) + 3\lambda_{\text{B}}I_0^2(\Omega) \tag{286}$$

where we have defined the integrals $I_p(\Omega)$ by

$$I_p(\Omega) = \frac{1}{(2\pi)^{d-1}} \int \frac{\mathrm{d}^{d-1}\boldsymbol{k}}{2\omega_\Omega(k)} \omega_\Omega^{2p}(k) \tag{287}$$

Of course, the integrals $I_p(\Omega)$ can be divergent and this signals, as usual in field theory, the need for renormalization. In fact, these integrals can be performed exactly and one is led to

$$I_p(\Omega) = \frac{1}{2(4\pi)^{(d-1)/2}} \frac{\Gamma[(2-d-2p)/2]}{\Gamma[(1-2p)/2]} \Omega^{d+2p-2} \tag{288}$$

It is clear that, for some values of d and p, the above expression does not exist and is only formal.

As already mentioned, in order to calculate the Gaussian effective potential, one must vary $\langle \mathcal{H} \rangle$ with respect to Ω. Upon using Eq. (286) and $dI_p/d\Omega = (2p-1)\Omega I_{p-1}$, one arrives at the following "gap" equation, the field theory counterpart of Eq. (239):

$$\Omega^2 = m_B^2 + 12\lambda_B\big[\Phi_0^2 + I_0(\Omega)\big] \tag{289}$$

This is an algebraic equation for Ω. Its solution can be written as $\overline{\Omega} = \overline{\Omega}(\Phi_0; m_B, \lambda_B)$, that is to say it depends on Φ_0 and on the bare parameters (in the following we no longer write the dependence on the bare parameters). Then, the Gaussian effective potential is defined by inserting this solution into Eq. (286). This leads to

$$V_G(\Phi_0) \equiv \langle \mathcal{H} \rangle\big[\overline{\Omega}(\Phi_0), \Phi_0\big] \tag{290}$$

Upon using the gap equation, one can also find a more compact expression, namely

$$V_G(\Phi_0) = I_1\big[\overline{\Omega}(\Phi_0)\big] - 3\lambda_B I_0^2\big[\overline{\Omega}(\Phi_0)\big] + \frac{1}{2}m_B^2\Phi_0^2 + \lambda_B\Phi_0^4 \tag{291}$$

At this level, this expression is in fact still formal because of the divergences of the integrals I_p. Therefore, we must first tame those divergences. This is done by means of a renormalization of the parameters appearing in the Lagrangian of the model.

Renormalization consists in parametrizing the physical quantities in terms of new, observable parameters, different from the bare parameters that appear in the Lagrangian. We now define the renormalized mass by

$$m_R^2 = \frac{d^2 V_G}{d\Phi_0^2}\bigg|_{\Phi_0=0} \tag{292}$$

Upon using Eq. (291) and the following relation [which comes from differentiating Eq. (289)]

$$\frac{d\overline{\Omega}}{d\Phi_0} = \frac{\Phi_0}{\overline{\Omega}}\frac{12\lambda_B}{1 + 6\lambda_B I_{-1}(\overline{\Omega})} \tag{293}$$

one finds that

$$\frac{d^2 V_G}{d\Phi_0^2} = m_B^2 + 12\lambda_B\big[\Phi_0^2 + I_0(\overline{\Omega})\big] - \frac{(12\lambda_B\Phi_0)^2 I_{-1}(\overline{\Omega})}{1 + 6\lambda_B I_{-1}(\overline{\Omega})} \tag{294}$$

As a consequence, putting $\Phi_0 = 0$ in the above equation, the renormalized mass can be expressed as

$$m_R^2 = m_B^2 + 12\lambda_B I_0(\overline{\Omega}_0) \tag{295}$$

where $\overline{\Omega}_0$ denotes the solution of the gap equation (289) when $\Phi_0 = 0$, i.e. $\Omega^2 = m_B^2 + 12\lambda_B I_0(\Omega)$. It is obvious that $\overline{\Omega}_0$ is not a function of Φ_0 but depends only on m_B and λ_B. In fact, this can also been seen from Eq. (289) since it tells us that $m_R^2 = \overline{\Omega}_0^2$. As a consequence, the above equation can be re-written as

$$m_B^2 = m_R^2 - 12\lambda_B I_0(m_R) \tag{296}$$

We are now in a position where we can calculate the Gaussian effective potential in terms of the renormalized mass. For this purpose, we restart from Eq. (290) and substitute the expression (296). One obtains

$$V_G(\Phi_0) = I_1(\overline{\Omega}) + \frac{1}{2}\big(m_R^2 - \overline{\Omega}^2\big)I_0(\overline{\Omega}) + \frac{1}{2}m_R^2\Phi_0^2 + \lambda_B\Phi_0^4 + 6\lambda_B\big[I_0(\overline{\Omega}) - I_0(m_R)\big]\Phi_0^2$$
$$+ 3\lambda_B I_0^2(\overline{\Omega}) - 6\lambda_B I_0(\overline{\Omega})I_0(m_R) \tag{297}$$

However, at this stage, the expression of the Gaussian effective potential is not yet a function of the renormalized mass only as one still needs to express $I_0(\overline{\Omega})$ and $I_1(\overline{\Omega})$ in terms of $I_0(m_R)$ and $I_1(m_R)$. There are several ways to do it, see Ref. [59]. Here we choose to use a method where a cut-off is temporarily introduced. We work in $d = 4$ dimensions to simplify the calculations. Let us exemplify the method with I_0. We have

$$I_0(\Omega) = \frac{1}{2(2\pi)^3}\int d^3\boldsymbol{k}\,\omega_\Omega^{-1} \tag{298}$$

$$= \frac{1}{4\pi^2}\int_0^M dk\,k^2\big(k^2 + \Omega^2\big)^{-1/2} \tag{299}$$

where we have performed the angular integrals and where we have introduced a cut-off M to regulate the otherwise divergent integral. This gives

$$I_0(\Omega) = \frac{1}{4\pi^2}\left[\frac{M}{2}\sqrt{\Omega^2 + M^2} - \frac{\Omega^2}{2}\ln\!\left(M + \sqrt{\Omega^2 + M^2}\right) + \frac{\Omega^2}{2}\ln\Omega\right] \tag{300}$$

As a consequence, one can write

$$I_0(\Omega) - I_0(m) = \frac{1}{2}(m^2 - \Omega^2)I_{-1}(m) + \frac{1}{4\pi^2}\left[\frac{M}{2}\sqrt{\Omega^2 + M^2} - \frac{M}{2}\sqrt{m^2 + M^2} + \frac{\Omega^2}{2}\ln\!\left(\frac{M + \sqrt{m^2 + M^2}}{M + \sqrt{\Omega^2 + M^2}}\right)\right.$$
$$\left. + \frac{\Omega^2}{2}\ln\Omega - \frac{m^2}{2}\ln m + \frac{1}{2}(m^2 - \Omega^2)\left(\frac{M}{\sqrt{m^2 + M^2}} + \ln m\right)\right] \tag{301}$$

where we have used

$$I_{-1}(\Omega) = \frac{1}{4\pi^2}\left[\frac{-M}{\sqrt{\Omega^2 + M^2}} + \ln\!\left(M + \sqrt{\Omega^2 + M^2}\right) - \ln\Omega\right] \tag{302}$$

an expression which can be obtained with the same method as described above. Then we take the limit $M \to +\infty$ and Eq. (301) becomes

$$I_0(\Omega) - I_0(m) = -\frac{1}{2}(\Omega^2 - m^2)I_{-1}(m) + \frac{m^2}{16\pi^2}\big[x\ln x - (x - 1)\big] \tag{303}$$

where $x \equiv \Omega^2/m^2$. Notice that, at this stage of the calculation, there is no need to write x in terms of $\overline{\Omega}$ and/or m_R. The above expression is valid for any Ω and m. Following Ref. [59], we call the function between the square brackets in the above equation $L_2(x)$. Exactly in the same manner, one can also show that

$$I_1(\Omega) - I_1(m) = \frac{1}{2}(\Omega^2 - m^2)I_0(m) - \frac{1}{8}(\Omega^2 - m^2)^2 I_{-1}(m) + \frac{m^4}{32\pi^2}L_3(x) \tag{304}$$

with $L_3(x) = [2x^2\ln x - 2(x - 1) - 3(x - 1)^2]/4$. Then, upon using Eqs. (303) and (304) in Eq. (297), one obtains

$$V_G(\Phi_0) = I_1(m_R) - 3\lambda_B I_0^2(m_R) + \frac{1}{2}m_R^2\Phi_0^2 + \lambda_B\Phi_0^4 + \frac{m_R^4}{32\pi^2}L_3(\bar{x}) + \frac{3}{4}\lambda_B m_R^4(\bar{x} - 1)^2 I_{-1}^2(m_R)$$
$$+ \frac{1}{8}I_{-1}(m_R)m_R^4(\bar{x} - 1)\left\{(\bar{x} - 1) - \frac{3\lambda_B}{2\pi^2}\left[L_2(\bar{x}) + \frac{16\pi^2}{m_R^2}\Phi_0^2\right]\right\}$$
$$- \frac{m_R^4}{32\pi^2}L_2(\bar{x})\left\{(\bar{x} - 1) - \frac{3\lambda_B}{8\pi^2}\left[L_2(\bar{x}) + 32\pi^2\frac{\Phi_0^2}{m_R^2}\right]\right\} \tag{305}$$

with $\bar{x} \equiv \overline{\Omega}^2/m_R^2$. Finally, the gap equation (289) can also be re-written in terms of the renormalized mass. The corresponding expression reads

$$(\bar{x} - 1)\big[1 + 6\lambda_B I_{-1}(m_R)\big] = \frac{3\lambda_B}{4\pi^2}\left[L_2(\bar{x}) + 16\pi^2\frac{\Phi_0^2}{m_R^2}\right] \tag{306}$$

We see that the dependence on Φ_0 of the Gaussian effective potential is quite complicated. Indeed, solving the gap equation provides us with $\bar{x} = \bar{x}(\Phi_0)$ that we must then insert in Eq. (305).

We shall now discuss the renormalization of the coupling constant λ_B. It is defined by

$$\lambda_R = \frac{1}{4!}\frac{d^4 V_G}{d\Phi_0^4}\bigg|_{\Phi_0=0} \tag{307}$$

Proceeding in the same manner as before, one obtains the following equation for the quantity λ_R

$$\lambda_R = \lambda_B\frac{1 - 12\lambda_B I_{-1}(m_R)}{1 + 6\lambda_B I_{-1}(m_R)} \tag{308}$$

The renormalization of the coupling constant has been discussed at length in Refs. [59,63] and we will not repeat all the details since, here, our main concern is the vacuum energy. In fact we will focus on the case named "precarious" in Ref. [59]. Let us return to Eq. (308); one can invert this expression and express the bare coupling constant in terms of the renormalized one. One obtains

$$\lambda_B = \frac{1 - 6\lambda_R I_{-1}(m_R)}{24 I_{-1}(m_R)}\left\{1 \pm \sqrt{1 - \frac{48\lambda_R I_{-1}(m_R)}{[1 - 6\lambda_R I_{-1}(m_R)]^2}}\right\} \tag{309}$$

Then, from Eq. (302), we see that $I_{-1}(m_R)$ is in fact a logarithmically divergent quantity when the cut-off M is sent to infinity. Therefore, we must treat $I_{-1}(m_R)$ as an arbitrarily large quantity. In this limit, if one chooses the plus sign in Eq. (309), one obtains

$$\lambda_B = -\frac{1}{2}\lambda_R + \mathcal{O}\left[\frac{1}{I_{-1}(m_R)}\right] \tag{310}$$

As shown in Ref. [59], this case is not viable. Intuitively, it is clear that a large negative coupling constant leads to a potential which is not bounded from below. Therefore, we are left with the minus sign for which one finds that

$$\lambda_B = -\frac{1}{6I_{-1}(m_R)}\left\{1 + \frac{1}{2\lambda_R I_{-1}(m_R)} + \mathcal{O}\left[\frac{1}{I_{-1}^2(m_R)}\right]\right\} \tag{311}$$

In passing, it is worth signaling that this equation can be re-expressed in a different manner as follows. Let us first introduce the new characteristic scale M_C (not to be confused with the cut-off M discussed before) defined by

$$\frac{1}{\lambda_R} \equiv -\frac{1}{4\pi^2}\ln\left(\frac{m_R^2}{M_C^2}\right) \tag{312}$$

Then, if we use the fact that

$$I_{-1}(M_C) - I_{-1}(m_R) = -\frac{1}{8\pi^2}L_1(x) \tag{313}$$

where $L_1(x) = \ln x$ (an expression which can be derived with the same method that has led to Eqs. (301) and (304)), the bare coupling constant can be re-written as

$$\lambda_B = -\frac{1}{6I_{-1}(M_C)} \tag{314}$$

Therefore, it has a negative but infinitesimal value which implies that the stability properties of the corresponding scenario are much better compared to the case (310). As argued in Ref. [59], besides simplifying the expression of the renormalized coupling constant, the scale M_C plays an important role in a complete study of the precarious case (even if, in the present context, this is not apparent since we focus on the vacuum energy only). Inserting the expression (311) (or Eq. (314)) into Eq. (305), one can re-write the effective potential as

$$V_G(\Phi_0) = I_1(m_R) - 3\lambda_B I_0^2(m_R) + \frac{1}{2}m_R^2\bar{x}\Phi_0^2 + \frac{m_R^4}{32\pi^2}L_3(\bar{x}) - \frac{m_R^4}{16\lambda_R}(\bar{x}-1)^2$$
$$+ \mathcal{O}\left[\frac{1}{I_{-1}(m_R)}\right] \tag{315}$$

while the gap equation (306) becomes

$$\bar{x} - 1 = \frac{\lambda_R}{4\pi^2}\left[L_2(\bar{x}) + 16\pi^2\frac{\Phi_0^2}{m_R^2}\right] \tag{316}$$

Let us notice that, in Eq. (315), we have left the quantity λ_B into the field independent term $I_1(m_R) - 3\lambda_B I_0^2(m_R)$ since it represents the zero point energy that will be treated in detail in the next Section 6.4. Finally, if one uses the explicit form of the functions L_2 and L_3, then the Gaussian effective potential takes the form

$$V_G(\Phi_0) = I_1(m_R) - 3\lambda_B I_0^2(m_R) + \frac{1}{2}m_R^2\bar{x}\Phi_0^2$$
$$+ \frac{m_R^4}{128\pi^2}\left[2\bar{x}^2\ln\bar{x} - 2(\bar{x}-1) - 3(\bar{x}-1)^2 - 2\frac{4\pi^2}{\lambda_R}(\bar{x}-1)^2\right] \tag{317}$$

and the gap equation can be re-expressed as

$$(\bar{x}-1)\left(1 + \frac{4\pi^2}{\lambda_R}\right) - 16\pi^2\frac{\Phi_0^2}{m_R^2} = \bar{x}\ln\bar{x} \tag{318}$$

The two equations (317) and (318) give a complete expression of the Gaussian effective potential for the renormalized precarious "$\lambda\Phi^4$" theory [59,63].

As we have done for the quantum mechanics case, it is interesting to compare the above result to the standard effective potential approach. We know that this potential can be obtained from the Gaussian effective potential method by neglecting

the terms proportional to $\hbar^2$. Therefore, we restart from Eq. (286) and neglect the term $3\lambda_{\mathrm{B}}I_0^2$ since this corresponds to the standard one loop method (i.e. it is of order $\hbar^2$). One obtains

$$\langle \mathcal{H} \rangle (\Omega, \Phi_0) = I_1(\Omega) + \frac{1}{2}\big(m_{\mathrm{B}}^2 - \Omega^2\big)I_0(\Omega) + \frac{1}{2}m_{\mathrm{B}}^2 \Phi_0^2 + \lambda_{\mathrm{B}}\Phi_0^4 + 6\lambda_{\mathrm{B}}\Phi_0^2 I_0(\Omega) + \mathcal{O}\big(\hbar^2\big) \tag{319}$$

As a consequence, this leads to the new following gap equation

$$\Omega^2 = m_{\mathrm{B}}^2 + 12\lambda_{\mathrm{B}}\Phi_0^2 \tag{320}$$

Clearly, this relation is the exact counter-part of Eq. (280). Then, one has to calculate the renormalized mass and coupling constant. One arrives at

$$m_{\mathrm{R}}^2 = m_{\mathrm{B}}^2 + 12\lambda_{\mathrm{B}}I_0(m_{\mathrm{B}}), \tag{321}$$

$$\lambda_{\mathrm{R}} = \lambda_{\mathrm{B}}\big[1 - 18\lambda_{\mathrm{B}}I_{-1}(m_{\mathrm{B}})\big]. \tag{322}$$

These equations should be compared to Eqs. (296) and (308). We notice that it is now m_{B} that appears in the argument of the integrals rather than m_{R} as in Eqs. (296) and (308). This is because Eq. (320) implies that $\overline{\Omega}(\Phi_0 = 0) = m_{\mathrm{B}}$ and not $\overline{\Omega}(\Phi_0 = 0) = m_{\mathrm{R}}$ as before. The final step consists in using the two previous equation in Eq. (319). One obtains

$$V_{\mathrm{G}}(\Phi_0) = I_1(m_{\mathrm{B}}) + \frac{1}{2}m_{\mathrm{R}}^2 \Phi_0^2 + \lambda_{\mathrm{R}}\Phi_0^4$$
$$+ \frac{m_{\mathrm{B}}^4}{64\pi^2}\left[\left(1 + 12\lambda_{\mathrm{B}}\frac{\Phi_0^2}{m_{\mathrm{B}}^2}\right)^2 \ln\left(1 + 12\lambda_{\mathrm{B}}\frac{\Phi_0^2}{m_{\mathrm{B}}^2}\right) - 12\lambda_{\mathrm{B}}\frac{\Phi_0^2}{m_{\mathrm{B}}^2} - 216\lambda_{\mathrm{B}}^2\frac{\Phi_0^4}{m_{\mathrm{B}}^4}\right] + \cdots \tag{323}$$

the dots corresponding to the higher order terms in $\hbar$. If one neglects those terms, then V_{G} reduces to V_{eff}. The corresponding expression is nothing but the Coleman–Weinberg potential. The previous analysis confirms that the effective potential is indeed a particular case of the Gaussian effective potential approach. An even more detailed comparisons of the two methods can be found in Refs. [59,63]. Here, we do not pursue this issue and now come to the question of the vacuum energy problem.

6.4. Vacuum energy and Gaussian effective potential

We are now in a position where the main problem of this section can be addressed. From Eqs. (305) and (315), we see that the vacuum energy density predicted by the Gaussian effective potential method (i.e. the term left after having taken $\Phi_0 = 0$ in those expressions) can be expressed as

$$\rho_{\mathrm{vac}} = I_1(m_{\mathrm{R}}) - 3\lambda_{\mathrm{B}}I_0(m_{\mathrm{R}}) \tag{324}$$

As we have already stressed many times, it is important to remember that this result is a priori non-perturbative. Using the expression of the precarious coupling constant (311), one obtains

$$\rho_{\mathrm{vac}} = I_1(m_{\mathrm{R}}) + \frac{1}{2}\frac{I_0(m_{\mathrm{R}})}{I_{-1}(m_{\mathrm{R}})}\left[1 + \frac{1}{2\lambda I_{-1}(m_{\mathrm{R}})}\right]I_0(m_{\mathrm{R}}) \tag{325}$$

The expressions of the integrals I_p have been given in Eqs. (287) and (288). The three integrals that appear in the above expression are all divergent. Nevertheless, the ratio I_0/I_{-1} is finite and reads

$$\frac{I_0(m_{\mathrm{R}})}{I_{-1}(m_{\mathrm{R}})} = \frac{m_{\mathrm{R}}^2}{2 - d} \tag{326}$$

Nevertheless, even after having inserted the above ratio into Eq. (325), this expression still contains singular terms. Explicitly, one has

$$\rho_{\mathrm{vac}} = \frac{m_{\mathrm{R}}^4}{4(2-d)}\left[-m_{\mathrm{R}}^{d-4}\frac{4-d}{2\sqrt{\pi}(4\pi)^{(d-1)/2}}\Gamma\left(-\frac{d}{2}\right) + \frac{1}{\lambda_{\mathrm{R}}(2-d)}\right] \tag{327}$$

However, writing this expression for $d = 4 + \epsilon$ and taking the limit $\epsilon \to 0$, we have the remarkable result that a second cancellation occurs and that the final expression is in fact finite. It can be expressed as

$$\rho_{\mathrm{vac}} = \frac{m_{\mathrm{R}}^4}{128\pi^2}\left[1 - 2\ln\left(\frac{m_{\mathrm{R}}^2}{M_{\mathrm{C}}^2}\right)\right] \tag{328}$$

where we have used Eq. (314). Even more remarkable is that the above equation is very similar to Eq. (95). In particular, the vacuum energy is proportional to fourth power, not of the cut-off, but of the mass of the corresponding particle and the

scale M_C is present as a logarithmic correction. Therefore, we conclude that our analysis based on the Gaussian effective potential has totally confirmed the results presented in the previous sections. We end this section by noticing that the vacuum energy predicted by the Coleman–Weinberg potential is, see Eq. (323), $\rho_{\mathrm{vac}} = I_1(m_B)$ and is therefore different from the one obtained before, see Eq. (328). Moreover, it seems that this expression remains divergent.

7. Contribution from other fields

So far, for simplicity, we have only treated the case of a scalar field. However, in order to be more realistic, it is clear that one must also evaluate the vacuum energy for other type of fields. This question is the purpose of the present section. As we will see, the corresponding calculations are in fact very similar. In fact, the important question is the dependence of the overall coefficient which multiplies ρ_{vac} with the spin of the particle. As is well known, this is at the origin of a remarkable cancellation which leads to the concept of super-symmetry. Therefore, in the following, we will pay special attention to this issue.

7.1. Fermion fields

We start with the case of a Dirac spinor field describing a spin 1/2 particle [16–22]. For convenience, we quickly recall basics fact about the quantum field theory of a spin 1/2 field. Then, we turn to the calculation of ρ_{vac}.

The Lagrangian of the free model is [we treated the mass-less case in Eq. (35)]

$$\mathcal{L}_{\mathrm{Dirac}} = -\overline{\Psi}\left(i\gamma^\mu \partial_\mu - m\right)\Psi \tag{329}$$

where γ^μ are matrices satisfying

$$\{\gamma^\mu, \gamma^\nu\} = -2g^{\mu\nu} \tag{330}$$

The above equation is in fact valid for any metric tensor. Here, we just consider the flat space–time case and, therefore, we will always consider that $g_{\mu\nu} = \eta_{\mu\nu}$. The quantity $\Psi(t, \boldsymbol{x})$ represents a four components Dirac spinor and $\overline{\Psi} \equiv \Psi^\dagger \gamma^0$ (this definition was already mentioned just after Eq. (33)). Here, we use the Dirac representation which means that the γ^μ matrices can be taken to be equal to [recall that, in Eq. (31), we used the chiral representation]

$$\gamma^0 = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$$
$$\gamma^5 = \begin{pmatrix} 0 & \mathbb{I}_2 \\ \mathbb{I}_2 & 0 \end{pmatrix} \tag{331}$$

where σ^i are the standard two-dimensional Pauli matrices. The equation of motion for the spinor Ψ is easily obtained from Eq. (329). Varying the Dirac Lagrangian, it is straightforward to show that

$$\left(i\gamma^\mu \partial_\mu - m\right)\Psi = 0 \tag{332}$$

which is of course nothing but the Dirac equation. Then, as we have done for the scalar field case, we now expand the Ψ operator in Fourier modes according to

$$\Psi(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega(k)}} \sum_{r=1}^{2} \left[b_{\boldsymbol{k}}^r u(\boldsymbol{k}, r) \mathrm{e}^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + \left(d_{\boldsymbol{k}}^r\right)^\dagger v(\boldsymbol{k}, r) \mathrm{e}^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}} \right] \tag{333}$$

As appropriate for a spinor field, the operators $b_{\boldsymbol{k}}^r$ and $d_{\boldsymbol{k}}^r$ satisfy anti-commutation rules. The index r represents the two polarization states. Inserting the expansion (333) into the Dirac equation (332) leads to two equations, namely

$$\left(\gamma^0 \omega - \gamma^i k_i - m\right) u(\boldsymbol{k}, r) = 0 \tag{334}$$
$$\left(\gamma^0 \omega - \gamma^i k_i + m\right) v(\boldsymbol{k}, r) = 0 \tag{335}$$

Upon using the explicit form (331), the first equation can be re-written as

$$\begin{pmatrix} (m - \omega)\mathbb{I}_2 & \sigma^i k_i \\ -\sigma^i k_i & (m + \omega)\mathbb{I}_2 \end{pmatrix} u(\boldsymbol{k}, r) = 0 \tag{336}$$

In order to have a non-trivial solution, the determinant of this matrix should vanish and this leads to

$$w(k) = \sqrt{k^2 + m^2} \tag{337}$$

where we have used the following properties of the Pauli matrices, $\sigma^i k_i \sigma^j k_j = k^2$. On the other hand, Eq. (336) also implies that

$$u(\mathbf{k}, r) = \begin{pmatrix} \psi \\ \frac{\sigma^i k_i}{m+\omega} \psi \end{pmatrix} \tag{338}$$

where ψ is an arbitrary two components spinor. Choosing the solution to be an eigenstate of the spin operator along the z-direction (in the rest frame of the particle), the two components of $u(\mathbf{k}, r)$ can be expressed as

$$u(\mathbf{k}, 1) = \sqrt{m+\omega} \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \frac{\sigma^i k_i}{m+\omega} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} \tag{339}$$

and

$$u(\mathbf{k}, 2) = \sqrt{m+\omega} \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{\sigma^i k_i}{m+\omega} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} \tag{340}$$

It is easy to check that this spinor is normalized according to

$$\bar{u}(\mathbf{k}, r)u(\mathbf{k}, s) = 2m\delta_{sr} \tag{341}$$

In the same manner, the solution $v(\mathbf{k}, r)$, corresponding to the other branch of the Fourier expansion, can be written as

$$v(\mathbf{k}, 1) = \sqrt{m+\omega} \begin{pmatrix} \frac{\sigma^i k_i}{m+\omega} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} \tag{342}$$

and

$$v(\mathbf{k}, 2) = \sqrt{m+\omega} \begin{pmatrix} \frac{\sigma^i k_i}{m+\omega} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} \tag{343}$$

The normalization of this spinor satisfies a relation almost identical to that of $u(\mathbf{k}, r)$ except for the presence of a minus sign

$$\bar{v}(\mathbf{k}, r)v(\mathbf{k}, s) = -2m\delta_{sr} \tag{344}$$

Finally, we also have

$$\bar{u}(\mathbf{k}, r)v(\mathbf{k}, s) = 0 \tag{345}$$

which can easily be obtained from the above expressions.

All the previous considerations are standard textbooks calculations. We now turn to our main goal, namely the calculation of the vacuum energy. Therefore, the next step consists in calculating the energy–momentum tensor. Either using the Noether procedure or a variation of the action with respect to vierbeins lead to the equation

$$T_{\mu\nu} = -\frac{i}{4}(\overline{\Psi}\gamma_\mu\partial_\nu\Psi + \overline{\Psi}\gamma_\nu\partial_\mu\Psi - \partial_\mu\overline{\Psi}\gamma_\nu\Psi - \partial_\nu\overline{\Psi}\gamma_\mu\Psi) \tag{346}$$

Of course, this expression is symmetric in μ and ν as it should. Endowed with this formula, one can now evaluate the various components of the stress–energy tensor in the vacuum state. Let us first calculate the energy density. One has

$$\langle\rho\rangle = -\frac{i}{4}\langle 0|(-2\Psi^\dagger\partial_0\Psi + 2\partial_0\Psi^\dagger\Psi)|0\rangle \tag{347}$$

since $\gamma^0 = -\gamma_0$ and $(\gamma^0)^2 = \mathbb{I}_4$. Inserting the Fourier expansion (333) into the above expression, one arrives at

$$\langle\rho\rangle = -\frac{1}{(2\pi)^3}\frac{1}{2}\int d\mathbf{k} \sum_{s=1}^{2} v^\dagger(\mathbf{k}, s)v(\mathbf{k}, s) \tag{348}$$

This equation can be further simplified. Upon using the explicit form of the spinors given above, one can show that

$$\sum_{s=1}^{2} v^\dagger(\mathbf{k}, s)v(\mathbf{k}, s) = 4\omega(k) \tag{349}$$

and, therefore, we obtain the following expression for the vacuum energy

$$\langle\rho\rangle = -\frac{1}{(2\pi)^3}\frac{4}{2}\int d\mathbf{k}\,\omega(k) \tag{350}$$

Several comments are in order at this point. From the above expression, we see that we find minus four times the corresponding result for a scalar field, see Eq. (67). We have a factor four because we have two particles (the particle and its anti-particle) with two polarization state each. Of course, the most striking aspect of the above equation is that $\langle \rho \rangle$ is negative. It is therefore interesting to understand the origin of the minus sign in more details. For this purpose, let us calculate the Hamiltonian operator associated to the Dirac field

$$H = \int \mathrm{d}^3\boldsymbol{x}\, T_{00}. \tag{351}$$

Inserting the Fourier expansion (333) of the spinor Ψ into the above equation, it is straightforward to arrive at[2]

$$H = \int \mathrm{d}\boldsymbol{k} \sum_{s=1}^{2} \omega(k) \big[(b_{\boldsymbol{k}}^s)^\dagger b_{\boldsymbol{k}}^s - d_{\boldsymbol{k}}^s (d_{\boldsymbol{k}}^s)^\dagger \big] \tag{352}$$

However, as we have already recalled, a spinor field is a quantity which anti-commutes and, therefore, one has $\{d_{\boldsymbol{k}}^s, (d_{\boldsymbol{p}}^r)^\dagger\} = \delta(\boldsymbol{k} - \boldsymbol{p})\delta_{rs}$. As a consequence, the Hamiltonian (352) can be re-written as

$$H = \int \mathrm{d}\boldsymbol{k} \sum_{s=1}^{2} \omega(k) \big[N_b + N_d - \delta(0) \big] \tag{353}$$

where $N_b \equiv (b_{\boldsymbol{k}}^s)^\dagger b_{\boldsymbol{k}}^s$ and $N_d \equiv (d_{\boldsymbol{k}}^s)^\dagger d_{\boldsymbol{k}}^s$ are the particle and anti-particle number operators respectively. Of course, their mean value vanishes in the vacuum state which contains no particle. The zero point energy is thus given by the formally infinite term $\delta(0)$. Since this term appears with a minus sign in the Hamiltonian, the corresponding energy density is indeed negative. The origin of this minus sign is the anti-commuting properties of the creation and annihilation operators. We conclude that the vacuum energy density is negative because we deal with fermions which are anti-commuting objects.

Let us now turn to the calculation of the pressure. Upon using Eq. (346), one has

$$\langle p \rangle = \frac{1}{3} \perp^{\mu\nu} T_{\mu\nu} \tag{354}$$

$$= -\frac{i}{6} g^{00} (\overline{\Psi}\gamma_0 \partial_0 \Psi - \partial_0 \overline{\Psi}\gamma_0 \Psi) - \frac{i}{6} g^{ij} (\overline{\Psi}\gamma_i \partial_j \Psi - 2\partial_i \overline{\Psi}\gamma_j \Psi)$$
$$\quad - \frac{i}{6} u^0 u^0 (\overline{\Psi}\gamma_0 \partial_0 \Psi - \partial_0 \overline{\Psi}\gamma_0 \Psi) \tag{355}$$

$$= -\frac{i}{6} \delta^{ij} (\overline{\Psi}\gamma_i \partial_j \Psi - 2\partial_i \overline{\Psi}\gamma_j \Psi) \tag{356}$$

From the above expression, we see that we have two terms to calculate. Using the definition of $\overline{\Psi}$, the first one reads

$$\langle 0|\Psi^\dagger \gamma^0 \gamma^i \partial_i \Psi|0\rangle = -\frac{i}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega(k)}} k_i \sum_{s=1}^{2} v^\dagger(\boldsymbol{k},s)\gamma^0 \gamma^i v(\boldsymbol{k},s) \tag{357}$$

while the second one takes the form

$$\langle 0|\Psi^\dagger \gamma^0 \gamma_i^i \Psi|0\rangle = -\langle 0|\partial_i \Psi^\dagger \gamma^0 \gamma^i \Psi|0\rangle \tag{358}$$

As a consequence, one finds that the vacuum pressure is given by the following expression

$$\langle p \rangle = -\frac{1}{3} \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega(k)}} \sum_{s=1}^{2} v^\dagger(\boldsymbol{k},s)\gamma^0 \gamma^i k_i v(\boldsymbol{k},s) \tag{359}$$

One must now evaluate the sum present in the above integral. We have seen that the spinors v obeys the Dirac equation $(\gamma^0 \omega - \gamma^i k_i + m)v(\boldsymbol{k},s) = 0$. Using this last equation, one deduces that

---

[2] For convenience, let us notice that the following relationships are useful for this calculation

$$u^\dagger(\boldsymbol{k},s)u(\boldsymbol{k},r) = 2\omega\delta_{rs}$$

$$v^\dagger(\boldsymbol{k},s)v(\boldsymbol{k},r) = 2\omega\delta_{rs}$$

$$v^\dagger(\boldsymbol{k},s)v(-\boldsymbol{k},r) = 0$$

These three equations are responsible for the fact that only the quantity $\omega(k)$ appears into the expression of the Hamiltonian (352).

$$v^\dagger(\boldsymbol{k}, s)\gamma^0\gamma^i k_i v(\boldsymbol{k}, s) = \omega v^\dagger(\boldsymbol{k}, s)v(\boldsymbol{k}, s) + v^\dagger(\boldsymbol{k}, s)\gamma^0 m v(\boldsymbol{k}, s) \tag{360}$$

Then, one can sum over the polarization and use the normalization of the spinors established before, see for instance Eqs. (341), (344) and (345), to obtain that

$$\sum_{s=1}^{2} v^\dagger(\boldsymbol{k}, s)\gamma^0\gamma^i k_i v(\boldsymbol{k}, s) = 4\omega^2 + \sum_{s=1}^{2} v^\dagger(\boldsymbol{k}, s)\gamma^0 m v(\boldsymbol{k}, s) \tag{361}$$

On the other hand, with the explicit expressions (342) and (343), it is easy to directly check that $\sum_{s=1}^{2} v^\dagger(\boldsymbol{k}, s)\gamma^0 m v(\boldsymbol{k}, s) = -4m^2$ from which one can establish that

$$\sum_{s=1}^{2} v^\dagger(\boldsymbol{k}, s)\gamma^0\gamma^i k_i v(\boldsymbol{k}, s) = 4\omega^2 - 4m^2 = 4k^2 \tag{362}$$

which is exactly the sum needed to evaluate the pressure. Therefore, our final expression for $\langle p \rangle$ can be written as

$$\langle p \rangle = -\frac{1}{(2\pi)^3}\frac{4}{6}\int \mathrm{d}\boldsymbol{k}\,\frac{k^2}{\omega(k)} \tag{363}$$

We see that, again, we recover exactly the same result as for the scalar field but with the additional multiplicative factor -4, see Eq. (69). This means that all the previous considerations discussed in the scalar field case also apply to the fermion case. In particular, the regularization of the above divergent integrals would proceed in the same way. Introducing a cut-off would break Lorentz invariance and would lead to an incorrect equation of state. On the other hand, using, say, dimensional regularization produces the correct vacuum equation of state and leads to a regularized vacuum energy which is minus four times the regularized scalar field vacuum energy. We do not need to repeat all these calculations here and we now move to the case of a vector field.

7.2. Proca fields

Having treated the case of a spin 1/2 field in the previous subsection, we now turn to the case of a Proca field [16–22]. A Proca field $A_\mu(t, \boldsymbol{x})$ is a massive vector, spin 1, field. The corresponding action can be written as

$$S[A_\mu] \equiv -\int \mathrm{d}^4 x\,\sqrt{-g}\left(\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu\right) \tag{364}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ and m is the mass of the particle. Again, we will restrict our considerations to flat space–time and, therefore, we will take for the metric determinant $g = \eta = 1$.

Before turning to the calculation of the vacuum stress–energy momentum, we quickly remind some basics fact about the quantization of this type of fields. Varying with respect to A_μ, one obtains the equation of motion

$$\partial_\mu F^{\mu\nu} - m^2 A^\nu = 0 \tag{365}$$

Taking the derivative of this expressions and noting that $\partial_\nu\partial_\mu F^{\mu\nu} = 0$ (since this is the contraction of a symmetric expression with an anti-symmetric one), one arrives at

$$\partial_\nu A^\nu = 0, \tag{366}$$

provided that $m \neq 0$, which we assume in this section. As a consequence, the equation of motion of the vector field is given by

$$\partial_\nu\partial^\nu A_\mu - m^2 A_\mu = 0 \tag{367}$$

which is nothing but a Klein–Gordon equation for each component of A_μ. As usual, we Fourier expand the field in terms of creation and annihilation operators. This leads to the following expression

$$A_\mu(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega(k)}}\sum_{\alpha=0}^{4}\epsilon_\mu^\alpha(\boldsymbol{k})\big[a_{\boldsymbol{k}}^\alpha \mathrm{e}^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + \big(a_{\boldsymbol{k}}^\alpha\big)^\dagger \mathrm{e}^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}}\big] \tag{368}$$

where, as usual, one has $\omega(k) = \sqrt{k^2 + m^2}$. Of course, the operators $a_{\boldsymbol{k}}^\alpha$ satisfy commutation rules since a vector field is a bosonic field. The sum over α is a sum over four vectors that span space–time and form a basis. At this point, A_μ satisfies the equation of motion but not yet the constraint $\partial_\nu A^\nu = 0$. Working out this constraint, it is easy to show that this amounts to $\epsilon_\mu^\alpha k^\mu a_{\boldsymbol{k}}^\alpha = 0$. If we now choose a frame where the wave-vector is aligned along the z-direction, then the basis can be taken to be

$$\epsilon_\mu^0 = \begin{pmatrix} -\omega/m \\ 0 \\ 0 \\ k/m \end{pmatrix}, \qquad \epsilon_\mu^1 = \begin{pmatrix} 0 \\ \hat{\epsilon}_x^1 \\ \hat{\epsilon}_y^1 \\ 0 \end{pmatrix} \tag{369}$$

$$\epsilon_\mu^2 = \begin{pmatrix} 0 \\ \hat{\epsilon}_x^2 \\ \hat{\epsilon}_y^2 \\ 0 \end{pmatrix}, \qquad \epsilon_\mu^3 = \begin{pmatrix} -k/m \\ 0 \\ 0 \\ \omega/m \end{pmatrix} \tag{370}$$

where $\hat{\epsilon}^{1,2}$ are vectors perpendicular to the wave-vector (the time components of ϵ^0 and ϵ^3 carry a minus sign because we show here the covariant vectors). But we have $\epsilon_\mu^0 k^\mu \neq 0$ despite the fact that this is the only choice such that $\epsilon_\mu^0 \epsilon^{\mu 3} = 0$. Therefore, the only possibility left in order to satisfy the constraint is to take $a_{\boldsymbol{k}}^0 = 0$. As a consequence, the Fourier expansion of the vector potential is now given by

$$A_\mu(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega(k)}} \sum_{\alpha=1}^3 \epsilon_\mu^\alpha(\boldsymbol{k}) \left[a_{\boldsymbol{k}}^\alpha e^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + \left(a_{\boldsymbol{k}}^\alpha\right)^\dagger e^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}} \right] \tag{371}$$

where the three vectors appearing in this decomposition are still given in Eqs. (369) and (370).

Let us now come to our main subject and calculate the stress–energy tensor. There are several ways to carry out this calculation. Here, we choose to return to Eq. (364) and to vary this action with respect to the space–time metric tensor. Then, in order to work with the stress–energy tensor in flat space–time, we will take $g_{\mu\nu} = \eta_{\mu\nu}$ but, of course, only after having performed the variation. Straightforward manipulations lead to the following expression

$$T_{\mu\nu} \equiv -\frac{2}{\sqrt{-g}} \frac{\delta S[A_\mu]}{\delta g^{\mu\nu}} \tag{372}$$

$$= -g_{\mu\nu}\left(\frac{1}{4} g^{\lambda\alpha} g^{\kappa\beta} F_{\lambda\kappa} F_{\alpha\beta} + \frac{1}{2} m^2 g^{\alpha\beta} A_\alpha A_\beta\right) + g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} + m^2 A_\mu A_\nu \tag{373}$$

We are now in a position where we can compute the various components of $T_{\mu\nu}$. Let us start with the energy density. From the above expression, one deduces that

$$T_{00} = \frac{1}{2}\delta^{ij} F_{0i} F_{0j} + \frac{1}{4}\delta^{ij}\delta^{kl} F_{ik} F_{jl} + \frac{1}{2}m^2 A_0^2 + \frac{1}{2}m^2 \delta^{ij} A_i A_j \tag{374}$$

Then, in order to evaluate the above equation, we have to separately calculate the four terms that appear in this expression. Let us first give the relevant results needed to calculate the first term. Upon using Eq. (371), it is easy to establish that

$$\delta^{ij}\langle 0|\partial_0 A_i \partial_0 A_j|0\rangle = \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k)} \delta^{ij} \sum_{\alpha=1}^3 \epsilon_i^\alpha \epsilon_j^\alpha \omega^2 \tag{375}$$

$$\delta^{ij}\langle 0|\partial_0 A_i \partial_j A_0|0\rangle = -\frac{1}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k)} \delta^{ij} \sum_{\alpha=1}^3 \epsilon_i^\alpha \epsilon_0^\alpha \omega k_j \tag{376}$$

and,

$$\delta^{ij}\langle 0|\partial_i A_0 \partial_j A_0|0\rangle = \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k)} \delta^{ij} \sum_{\alpha=1}^3 \epsilon_0^\alpha \epsilon_0^\alpha k_i k_j \tag{377}$$

In the same manner, the second term, containing the space component of the field strength, can be expressed as

$$\delta^{ij}\delta^{kl}\langle 0|F_{ik} F_{jl}|0\rangle = \frac{2}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k)} \delta^{ij}\delta^{kl} \sum_{\alpha=1}^3 \left(\epsilon_k^\alpha \epsilon_l^\alpha k_i k_j - \epsilon_k^\alpha \epsilon_j^\alpha k_i k_l\right) \tag{378}$$

The third term can be written as

$$\langle 0|A_0^2|0\rangle = \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k)} \sum_{\alpha=1}^3 \epsilon_0^\alpha \epsilon_0^\alpha \tag{379}$$

and, finally, the fourth and last term takes the form

$$\delta^{ij}\langle 0|A_i A_j|0\rangle = \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k)} \sum_{\alpha=1}^3 \delta^{ij} \epsilon_i^\alpha \epsilon_j^\alpha \tag{380}$$

Having calculated the four terms appearing in the expression of T_{00}, one must now combine them. Using the relationships satisfied by the vector basis,[3] one obtains the following expression for the vacuum energy density:

$$\langle \rho \rangle = \frac{1}{(2\pi)^3} \frac{3}{2} \int d\boldsymbol{k}\, \omega(k) \tag{381}$$

Based on our experience with the calculation of the vacuum energy of a spin 1/2 field, this expression is totally expected. As before, it is equal to the scalar field energy density times a multiplicative factor which takes into account the spin of the particle. Here, for a massive vector field, there are three polarization states and, as a consequence, the multiplicative factor is three.

In order to be exhaustive, let us now evaluate the pressure. The calculation proceeds exactly as before and it is easy to show that

$$\perp^{\mu\nu} T_{\mu\nu} = T_{00} + m^2 A_0^2 - m^2 \delta^{ij} A_i A_j \tag{382}$$

from which, using the relations already established before, we immediately deduce that

$$\langle p \rangle = \frac{1}{(2\pi)^3} \frac{3}{6} \int d\boldsymbol{k}\, \frac{k^2}{\omega(k)} \tag{383}$$

Of course, this is again the expected formula, namely the scalar field pressure times a multiplicative factor equals to three. Clearly, there is no need to repeat the discussion about the regularization of the above divergent integrals. It proceeds exactly as explained before.

7.3. Mass-less vector field

Finally, we consider the case where the vector field is mass-less, $m = 0$ [16–22]. This is of course necessary if we want to realistically evaluate the vacuum energy density since we know that this type of field exists in Nature. If the mass vanishes, this means that the dispersion relation is now given by

$$\omega(k) = k \tag{384}$$

as appropriate for photons. In that case, contrary to the massive case, we can no longer establish that $\partial_\nu A^\nu = 0$. Therefore, the equation of motion takes the form

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu (\partial_\mu A^\mu) = 0 \tag{385}$$

At this stage, we need to fix a gauge. As is well known, it is convenient to work in the Coulomb gauge defined by the following conditions:

$$A_0 = 0, \qquad \partial_i A^i = 0 \tag{386}$$

In particular, the last condition implies that $A_i(t, \boldsymbol{x})$ is a transverse field. As a consequence, its Fourier expansion can be written as

$$A_i(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d\boldsymbol{k}}{\sqrt{2\omega(k)}} \sum_{\alpha=1}^{2} \epsilon_i^\alpha(\boldsymbol{k}) \left[a_{\boldsymbol{k}}^\alpha e^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + (a_{\boldsymbol{k}}^\alpha)^\dagger e^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}} \right] \tag{387}$$

The two three-vectors ϵ^1 and ϵ^2 are perpendicular to the wave-vectors $\boldsymbol{k}$. In the mass-less case, only two degrees of freedom remain.

---

[3] For convenience, let us give the two equations needed to complete the calculation. It involves simple combination of the polarization vectors. Using the explicit representation given by Eqs. (369) and (370), one arrives at

$$\delta^{ij} \sum_{\alpha=1}^{3} \epsilon_i^\alpha \epsilon_j^\alpha = 2 + \frac{\omega^2}{m^2}$$

$$\sum_{\alpha=1}^{3} \epsilon_0^\alpha \epsilon_0^\alpha = \frac{k^2}{m^2}$$

Then, these two expressions can be combined to lead to the following relation which is what is needed in order to simplify the expression of the vacuum energy density

$$\sum_{\alpha=1}^{3} \left(\epsilon_k^\alpha \epsilon_l^\alpha k_i k_j - \epsilon_k^\alpha \epsilon_j^\alpha k_i k_l \right) = k^2 \left(2 + \frac{\omega^2}{m^2} \right) - \frac{k^2 \omega^2}{m^2} = 2$$

The calculation of the stress–energy tensor is now straightforward. Using the calculation of the previous section, in particular the expression of the time–time component of the stress–energy tensor (374), one finds that

$$T_{00} = \frac{1}{2}\delta^{ij}\partial_0 A_i \partial_0 A_j + \frac{1}{4}\delta^{ij}\delta^{kl}F_{ik}F_{jl} \tag{388}$$

and, since we now have,

$$\delta^{ij}\sum_{\alpha=1}^{2}\epsilon_i^\alpha \epsilon_j^\alpha = 2 \tag{389}$$

$$\sum_{\alpha=1}^{2}\left(\epsilon_k^\alpha \epsilon_l^\alpha k_i k_j - \epsilon_k^\alpha \epsilon_j^\alpha k_i k_l\right) = 2k^2 \tag{390}$$

one finally arrives at the following expression for the vacuum energy density:

$$\langle\rho\rangle = \frac{1}{(2\pi)^3}\frac{2}{2}\int \mathrm{d}\boldsymbol{k}\,\omega(k) \tag{391}$$

that is to say the scalar field energy density times two, as expected since there are two polarization states. For the pressure, using the same approach, one immediately obtains

$$\langle p\rangle = \frac{1}{(2\pi)^3}\frac{2}{6}\int \mathrm{d}\boldsymbol{k}\,\frac{k^2}{\omega(k)} \tag{392}$$

where, as usual, the factor two in front of the whole expression originates from the two polarization states of a mass-less vector particle.

We conclude by stating the main result established in this section. For any type of fields, the vacuum energy density can be written as

$$\langle\rho\rangle = \frac{1}{(2\pi)^3}\frac{s}{2}\int \mathrm{d}\boldsymbol{k}\,\omega(k) \tag{393}$$

where s represents the number of polarization states ($s = 1$ for a scalar field, $s = 4$ for a spinor field, $s = 3$ for a massive vector field and $s = 2$ for a mass-less vector field). The pressure is given by a similar expression. When these divergent integrals are regularized, one can show that, again for any type of field, this leads to the Lorentz invariant form of the vacuum equation of state, namely $\langle p\rangle = -\langle\rho\rangle$. In principle, the results of this section makes the accurate evaluation of the vacuum energy density in the standard model of particle physics possible. However, before discussing this point, we need to say something about the fact that ρ_{vac} is positive for bosons and negative for fermions.

7.4. Application: super-symmetry

7.4.1. Motivation

The fact that the sign of the zero-point energy density is different for fermions and bosons, see Eqs. (67) and (350), suggests a very simple way of solving the cosmological constant problem. It is clear that one can design a theory where the two contributions are equal in absolute value such that the final result is exactly zero. Let us see how it works in practice. For this purpose, let us consider again the Lagrangian of a massive real scalar field, see also Eq. (6)

$$\mathcal{L}_\phi = -\frac{\alpha}{2}\eta^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi - \frac{1}{2}m^2\Phi^2 \tag{394}$$

where we have introduced a constant α in front of the kinetic term. Before we took $\alpha = 1$ and it is interesting to discuss why we made this choice (after all, how do we know that $\alpha = 1$?). As usual, see Eq. (60), one Fourier expands the scalar field as

$$\Phi(t,\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}}\int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega(k,\alpha)}}\left(c_{\boldsymbol{k}}\mathrm{e}^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + c_{\boldsymbol{k}}^\dagger \mathrm{e}^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}}\right) \tag{395}$$

with a new expression for the quantity ω, compare to Eq. (61), namely

$$\omega(k,\alpha) = \sqrt{k^2 + \frac{m^2}{\alpha^2}} \tag{396}$$

since this is the condition for Eq. (395) to be a solution of the "new" equation of motion (see also Eq. (59))

$$-\alpha\ddot{\phi} + \alpha\delta^{ij}\partial_i\phi\partial_j\phi - m^2\phi = 0 \tag{397}$$

The new dispersion relation can easily be understood. We see that the quantity α simply renormalizes the value of the mass m. This makes sense since, if $m = 0$, we clearly see that the relativistic relation $\omega = k$ is not affected by the presence of the constant α.

The next step is to derive the commutation relations for the creation and annihilation operators. Let us assume for the moment that we have the standard relation (62)

$$\left[c_{\boldsymbol{k}}, c_{\boldsymbol{p}}^{\dagger}\right] = \delta(\boldsymbol{k} - \boldsymbol{p}) \tag{398}$$

Then, from the expression of the Lagrangian, one can calculate the conjugate momentum. It reads $\Pi(t, \boldsymbol{x}) = -\alpha \dot{\Phi}(t, \boldsymbol{x})$. This implies that the commutation relation between the field and its conjugate momentum now reads

$$\left[\phi(t, \boldsymbol{x}), \pi(t, \boldsymbol{y})\right] = i\alpha \delta(\boldsymbol{x} - \boldsymbol{y}) \tag{399}$$

an expression which reduces to the standard commutation relation only if $\alpha = 1$. Therefore, if one wants to have $[\Phi, \Pi] = i\delta(\boldsymbol{x} - \boldsymbol{y})$ and, at the same time, $[c_{\boldsymbol{k}}, c_{\boldsymbol{p}}^{\dagger}] = \delta(\boldsymbol{k} - \boldsymbol{p})$ there is no other choice that taking $\alpha = 1$. It is worth noticing that working with the two previous commutation relations (398), (399) and $\alpha \neq 1$ would lead to a different prediction for the energy density of the vacuum. Indeed, the stress–energy tensor is now given by

$$T_{\mu\nu} = \alpha \partial_{\mu}\Phi \partial_{\nu}\Phi - g_{\mu\nu}\left(\frac{\alpha}{2}g^{\alpha\beta}\partial_{\alpha}\Phi \partial_{\beta}\Phi + \frac{1}{2}m^2\Phi^2\right) \tag{400}$$

which implies that

$$T_{00} = \frac{\alpha}{2}\dot{\Phi}^2 + \frac{\alpha}{2}\delta^{ij}\partial_i\Phi \partial_j\Phi + \frac{1}{2}m^2\Phi^2. \tag{401}$$

From this expression, it is easy to show that

$$\langle\rho\rangle = \frac{\alpha}{(2\pi)^3}\frac{1}{2}\int \mathrm{d}\boldsymbol{k}\,\omega(k, \alpha) \tag{402}$$

an expression that should be compared to Eq. (67).

Of course, one can also demonstrate that the consistency can be re-establish if the commutation relations for the creation and annihilation operators are taken to be $[c_{\boldsymbol{k}}, c_{\boldsymbol{p}}^{\dagger}] = \delta(\boldsymbol{k} - \boldsymbol{p})/\alpha$. In this case the commutation relation of the field and its conjugate momentum have the correct form. Moreover, one has

$$\langle 0|\dot{\Phi}^2|0\rangle = \frac{1}{(2\pi)^3}\int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k, \alpha)}\frac{\omega^2(k, \alpha)}{\alpha} \tag{403}$$

$$\langle 0|\delta^{ij}\partial_i\Phi \partial_j\Phi|0\rangle = \frac{1}{(2\pi)^3}\int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k, \alpha)}\frac{k^2}{\alpha} \tag{404}$$

$$\langle 0|\Phi^2|0\rangle = \frac{1}{(2\pi)^3}\int \frac{\mathrm{d}\boldsymbol{k}}{2\omega(k, \alpha)}\frac{1}{\alpha} \tag{405}$$

these formulas should be compared to Eqs. (63), (64) and (65), which implies that

$$\langle\rho\rangle = \frac{1}{(2\pi)^3}\frac{1}{2}\int \mathrm{d}\boldsymbol{k}\,\omega(k, \alpha) \tag{406}$$

that is to say the "standard" result.

Therefore, in conclusion, the "1/2" coefficient in the expression of the vacuum energy density is only obtained if the commutation relations and the normalization of the field are properly and consistently chosen. In particular, if one chooses to work with a coefficient "1/2" in front of the kinetic term (canonical normalization), then there is no other choice than working with $[c_{\boldsymbol{k}}, c_{\boldsymbol{p}}^{\dagger}] = \delta(\boldsymbol{k} - \boldsymbol{p})$. Or, in other words, the previous commutation relation implies that a "1/2" in front of the kinetic term must be present.

We now apply the previous discussion to the cases of a complex scalar field and to a Majorana spinor. The Lagrangian for a complex scalar field should be written as

$$\mathcal{L} = -g^{\mu\nu}\partial_{\mu}\Phi^*\partial_{\nu}\Phi - m^2\Phi^*\Phi \tag{407}$$

with, this time, no "1/2" in front of the whole expression. Indeed, if one writes that

$$\phi = \frac{1}{\sqrt{2}}(A_1 + iA_2) \tag{408}$$

then the Lagrangian takes the form of the sum of two Lagrangian for the two real canonically normalized fields A_1 and A_2.

The previous discussion has also important implications for the Lagrangian of a Majorana spinor. A Majorana spinor is a spinor satisfying

$$\Psi_M^c = \Psi_M \tag{409}$$

where the charge conjugate spinor Ψ^c is defined by

$$\Psi^c \equiv C\overline{\Psi}^T \tag{410}$$

In this expression the symbol "T" denotes the transpose matrix. The matrix C is the charge conjugation operator that can be represented by $i\gamma^2\gamma^0$ and satisfies $C\gamma_\mu^T C^{-1} = -\gamma_\mu$. We have of course $(\Psi^c)^c = \Psi$. From a Dirac spinor Ψ one can always write two Majorana spinors

$$\Psi_{M1} = \frac{1}{\sqrt{2}}(\Psi + \Psi^c) \tag{411}$$

$$\Psi_{M2} = \frac{1}{i\sqrt{2}}(\Psi - \Psi^c) \tag{412}$$

Of course, the above expressions can always be inverted and re-written as

$$\Psi = \frac{1}{\sqrt{2}}(\Psi_{M1} + i\Psi_{M2}) \tag{413}$$

$$\Psi^c = \frac{1}{\sqrt{2}}(\Psi_{M1} - i\Psi_{M2}) \tag{414}$$

The two previous expression can be inserted in the Dirac Lagrangian (329) (taking $m = 0$). This leads to the following expression

$$\begin{aligned}
\mathcal{L}_{\text{Dirac}} &= -\frac{1}{\sqrt{2}}(\overline{\Psi}_{M1} - i\overline{\Psi}_{M2})i\gamma^\mu\partial_\mu\frac{1}{\sqrt{2}}(\overline{\Psi}_{M1} + i\overline{\Psi}_{M2}) \\
&= -\frac{i}{2}\overline{\Psi}_{M1}\gamma^\mu\partial_\mu\Psi_{M1} - \frac{i}{2}\overline{\Psi}_{M2}\gamma^\mu\partial_\mu\Psi_{M2} - \frac{i}{2}\left(-i\overline{\Psi}_{M2}\gamma^\mu\partial_\mu\Psi_{M1} + i\overline{\Psi}_{M1}\gamma^\mu\partial_\mu\Psi_{M2}\right)
\end{aligned} \tag{415}$$

But, using partial integrations, one easily shows that

$$\overline{\Psi}_{M2}\gamma^\mu\partial_\mu\Psi_{M1} = -(\partial_\mu\overline{\Psi}_{M2})\gamma^\mu\Psi_{M1} = +\overline{\Psi}_{M1}\gamma^\mu\partial_\mu\Psi_{M2} \tag{416}$$

so that the last term vanishes. Therefore, one conclude that the Lagrangian for a Majorana spinor is given by

$$\mathcal{L}_{\text{Majorana}} = -\frac{1}{2}\overline{\Psi}_M\left(i\gamma^\mu\partial_\mu - m\right)\Psi_M \tag{417}$$

and that the expansion of a Majorana spinor operator in Fourier can be expressed as, see also Eq. (333)

$$\Psi_M(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}}\int\frac{d\boldsymbol{k}}{\sqrt{2\omega(k)}}\sum_{r=1}^{2}\left[b_{\boldsymbol{k}}^r u(\boldsymbol{k}, r)e^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + \left(b_{\boldsymbol{k}}^r\right)^\dagger v(\boldsymbol{k}, r)e^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}}\right] \tag{418}$$

where we notice that only the operators $b_{\boldsymbol{k}}^r$ are present which reflects the fact that, for a Majorana spinor, the particle and the anti-particle are the same. But the most important aspect is the coefficient "1/2" in Eq. (417). Clearly, it means that if we perform the calculation of $\langle\rho\rangle$ again, we are going to find

$$\langle\rho\rangle = -\frac{1}{(2\pi)^3}\frac{2}{2}\int d\boldsymbol{k}\,\omega(k) \tag{419}$$

that is to say an expression with an overall factor 2 instead of 4 for a Dirac fermion, see Eq. (350).

We are now in a position to address the main point of this section. Let us consider the following model described by a complex scalar field and a Majorana spinor:

$$\mathcal{L} = -g^{\mu\nu}\partial_\mu\Phi^*\partial_\nu\Phi - m_B^2\phi^*\phi - \frac{1}{2}\overline{\Psi}_M\left(i\gamma^\mu\partial_\mu - m_F\right)\Psi_M \tag{420}$$

where m_B is the mass of the bosonic field and m_F is the mass of the fermionic field. From the above considerations, it is obvious to calculate the vacuum energy density. It reads

$$\langle\rho\rangle = \frac{1}{(2\pi)^3}\frac{2}{2}\int d\boldsymbol{k}\sqrt{k^2 + m_B^2} - \frac{1}{(2\pi)^3}\frac{2}{2}\int d\boldsymbol{k}\sqrt{k^2 + m_F^2} \tag{421}$$

We have a factor two for the complex scalar field since we have seen that a complex scalar field is equivalent to two real scalar fields and a factor -2 for the Majorana spinor as shown before. In general, the quantity $\langle \rho \rangle$ is not zero because $m_{\rm B} \neq m_{\rm F}$. But if we now assume that

$$m_{\rm B} = m_{\rm F} \tag{422}$$

then the zero-point energy density exactly cancels out and we have solved the cosmological constant problem. The above model is in fact nothing but the simplest super-symmetric model, the so-called Wess–Zumino model [19,70–74].

7.4.2. The vacuum energy in super-symmetry

Having motivated the idea of super-symmetry, we now study in more detail its implications. In particular, we would like to check that the type of cancellation studied before is generic in super-symmetry. In order to simplify the discussion, we start with considering the Lagrangian (420) with $m_{\rm B} = m_{\rm F} = 0$. We will now use the standard super-symmetric notation with the dotted two-components Weyl spinors [19,74]. In this formalism, the Majorana spinor can be written as

$$\Psi_{\rm M} = \begin{pmatrix} \Psi_\alpha \\ \overline{\Psi}^{\dot\alpha} \end{pmatrix} \tag{423}$$

Then, using that the Dirac matrices can be expressed as

$$\gamma^\mu = \begin{pmatrix} 0 & (\sigma^\mu)_{\alpha\dot\alpha} \\ (\overline{\sigma}^\mu)^{\dot\alpha\alpha} & 0 \end{pmatrix} \tag{424}$$

we find that the Lagrangian (417) can be re-written as

$$\mathcal{L}_{\rm Majorana} = -\frac{i}{2}\overline{\Psi}_{\dot\alpha}(\overline{\sigma}^\mu)^{\dot\alpha\alpha}\partial_\mu\Psi_\alpha - \frac{i}{2}\Psi^\alpha(\sigma^\mu)_{\alpha\dot\alpha}\partial_\mu\overline{\Psi}^{\dot\alpha} \tag{425}$$

where we used that (by definition) $\overline{\Psi}_{\dot\alpha} = (\Psi_\alpha)^*$ and $\Psi^\alpha = (\overline{\Psi}^{\dot\alpha})^*$. The next step consists in using the relation $\Psi^\alpha(\sigma^\mu)_{\alpha\dot\alpha}\partial_\mu\overline{\Psi}^{\dot\alpha} = -\partial_\mu\overline{\Psi}_{\dot\alpha}(\overline{\sigma}^\mu)^{\dot\alpha\alpha}\Psi_\alpha$ (the spinor components are Grassmann variables and, therefore, anti-commute, see below) and, then, to transform the resulting expression by integration by part. In this way, one shows that the Lagrangian (425) is in fact twice the first term. As a consequence, the Wess–Zumino model is described by

$$\mathcal{L}_{\rm W-Z} = -\eta^{\mu\nu}\partial_\mu\Phi^\dagger\partial_\nu\Phi - i\overline{\Psi}_{\dot\alpha}(\overline{\sigma}^\mu)^{\dot\alpha\alpha}\partial_\mu\Psi_\alpha \tag{426}$$

Let us now consider a super-symmetric transformation. In order to understand what it really means, let us use the analogy with a gauge transformation [19,74]. For definiteness, we consider a SU(2) transformation. Let v be a SU(2) doublet of spinor fields. After an infinitesimal transformation, this doublet v becomes a new doublet v' such that

$$v' = e^{i\epsilon^j\sigma_j/2}v \simeq \left(\mathbb{I}_2 + i\epsilon^j\frac{\sigma_j}{2}\right)v \tag{427}$$

or

$$\delta v = i\epsilon^j\frac{\sigma_j}{2}v \tag{428}$$

where σ_j are the Pauli matrices, the generators of the group. The quantity ϵ^j is a vector (or a collection of three parameters) specifying the transformation. In super-symmetry, the transformation is characterized by a Weyl spinor ξ^α which, therefore, plays the role of the quantity ϵ^j. A super-symmetric transformation transforms a boson into a fermion and vice-versa. As a consequence, one can write

$$\delta\Phi = \sqrt{2}\xi^\alpha\Psi_\alpha \tag{429}$$

$$\delta\Psi_\alpha = -i\sqrt{2}(\sigma^\mu)_{\alpha\dot\beta}\overline{\xi}^{\dot\beta}\partial_\mu\Phi \tag{430}$$

where the factors i or $\sqrt{2}$ are introduced for future convenience. Before showing that the Wess–Zumino Lagrangian is indeed invariant under a super-symmetric transformation, one needs to introduce a last result. We would like the Lagrangian to be invariant without using the equation of motion (in other words, we would like the Lagrangian to be invariant "off-shell"). In this case, one needs to add a term such that the new Lagrangian reads

$$\mathcal{L}_{\rm W-Z} = -\eta^{\mu\nu}\partial_\mu\Phi^\dagger\partial_\nu\Phi - i\overline{\Psi}_{\dot\alpha}(\overline{\sigma}^\mu)^{\dot\alpha\alpha}\partial_\mu\Psi_\alpha - F^\dagger F \tag{431}$$

where F is a complex scalar field. It is easy to see that this term is harmless since the corresponding equation of motion is just $F = 0$. But this term is necessary if one wants to match the bosonic and fermionic degrees of freedom. A Weyl spinor has two-complex components and we now have two complex scalar fields. Obviously, we also need to modify our super-symmetric transformation and they now take the form

$$\delta\Phi = \sqrt{2}\xi^\alpha\Psi_\alpha \tag{432}$$

$$\delta\Psi_\alpha = \sqrt{2}\xi_\alpha F - i\sqrt{2}(\sigma^\mu)_{\alpha\dot\beta}\bar\xi^{\dot\beta}\partial_\mu\Phi \tag{433}$$

$$\delta F = i\sqrt{2}(\partial_\mu\Psi^\alpha)(\sigma^\mu)_{\alpha\dot\beta}\bar\xi^{\dot\beta} \tag{434}$$

Our goal is now to show that the Lagrangian (431) is indeed invariant under the three transformations (432), (433) and (434).

In fact, the key to show that the Lagrangian is invariant is to be able to evaluate the above transformations for the complex conjugate quantities. We have

$$\begin{aligned}
\delta\Phi^\dagger &= \sqrt{2}(\xi^\alpha\Psi_\alpha)^\dagger = \sqrt{2}(\Psi_\alpha)^\dagger(\xi^\alpha)^\dagger \\
&= \sqrt{2}(\Psi_\alpha)^*(\xi^\alpha)^* = \sqrt{2}\,\overline\Psi_{\dot\alpha}\bar\xi^{\dot\alpha} \\
&= \sqrt{2}\,\bar\xi_{\dot\alpha}\overline\Psi^{\dot\alpha} = -\sqrt{2}\,\bar\xi^{\dot\alpha}\overline\Psi_{\dot\alpha}
\end{aligned} \tag{435}$$

We cannot review here the complete dotted spinors formalism in detail, see Ref. [74], but in order to make the calculations reasonably self-consistent, it is nevertheless interesting to recall a few things. The first point to discuss is the definition of the dagger symbol for Grassmann variables. In fact, a very clear discussion can be found in p. 35 of Ref. [74] and there is no need here to entirely repeat it. In brief, one needs to remember that we deal with quantum fields. Usually, the dagger symbols means complex conjugation and transposition. For a scalar field, it simply amounts to take the complex conjugate of the mode function and put (or remove) a dagger to the creation and annihilation operators. For the spinor component Ψ_α, this is the same (here it is important to remember that one discusses one component only and not the spinor viewed as a column and, therefore, we "cannot" take the transpose of the matrix – there is no matrix!) and one defines $\Psi_\alpha^\dagger$ as the complex conjugate of the corresponding spinor component times the creation and annihilation operators with or without the dagger symbol. However, the question arises as how to evaluate the Hermitian conjugate of a product of Grassmann variables. Here, Ref. [74] shows that the most convenient way is to define $(\xi^\alpha\Psi_\alpha)^\dagger = \Psi_\alpha^\dagger\xi^{\alpha\dagger}$, that is to say something similar to what we are used to in the case of two matrices even if, again, we do not deal with matrices here but with components of Weyl spinors. This definition justifies the first step in the previous equation.

Then, it is also interesting to discuss how the two last equations are obtained. Let us write $\overline\Psi_{\dot 1} = a$, $\overline\Psi_{\dot 2} = b$ and $\bar\xi^{\dot 1} = c$, $\bar\xi^{\dot 2} = d$, where a, b, c and d are (complex) Grassmann variables. As a consequence, one has $\overline\Psi_{\dot\alpha}\bar\xi^{\dot\alpha} = ac + bd$. On the other hand, $\bar\xi_{\dot\alpha} \equiv \epsilon_{\dot\alpha\dot\beta}\bar\xi^{\dot\beta}$, where

$$\epsilon_{\dot\alpha\dot\beta} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{436}$$

from which we deduce that $\bar\xi_{\dot 1} = d$ and $\bar\xi_{\dot 2} = -c$. In the same manner, one has $\overline\Psi^{\dot\alpha} = \epsilon^{\dot\alpha\dot\beta}\overline\Psi_{\dot\beta}$ with

$$\epsilon^{\dot\alpha\dot\beta} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{437}$$

from which one obtains $\overline\psi^{\dot 1} = -b$ and $\overline\psi^{\dot 2} = a$. Using these results, one has $\bar\xi_{\dot\alpha}\overline\Psi^{\dot\alpha} = -db - ca = ac + bd = \overline\Psi_{\dot\alpha}\bar\xi^{\dot\alpha}$, which is the wanted result. But, using again the expressions of the components obtained above, one also notices that $\bar\xi_{\dot\alpha}\overline\Psi^{\dot\alpha} = -\bar\xi^{\dot\alpha}\overline\Psi_{\dot\alpha}$ which is the last result used to establish the expression (435) of $\delta\Phi^\dagger$. In a more compact notation, one has

$$\bar\xi_{\dot\alpha}\overline\Psi^{\dot\alpha} = \epsilon_{\dot\alpha\dot\beta}\bar\xi^{\dot\beta}\overline\Psi^{\dot\alpha} = \bar\xi^{\dot\beta}\epsilon_{\dot\alpha\dot\beta}\overline\Psi^{\dot\alpha} = -\bar\xi^{\dot\beta}\epsilon_{\dot\beta\dot\alpha}\overline\Psi^{\dot\alpha} = -\bar\xi^{\dot\beta}\overline\Psi_{\dot\beta} \tag{438}$$

the key point being the fact that the matrix $\epsilon_{\dot\alpha\dot\beta}$ is anti-symmetric.

We now need the transformation of $(\delta\Psi_\alpha)^* = \delta\overline\Psi_{\dot\alpha}$. One has

$$\delta\overline\Psi_{\dot\alpha} = \sqrt{2}\,\bar\xi_{\dot\alpha}F^\dagger + i\sqrt{2}\big[(\sigma^\mu)_{\alpha\dot\beta}\big]^*\xi^\beta\partial_\mu\phi^\dagger \tag{439}$$

$$= \sqrt{2}\,\bar\xi_{\dot\alpha}F^\dagger + i\sqrt{2}(\sigma^\mu)_{\beta\dot\alpha}\xi^\beta\partial_\mu\Phi^\dagger \tag{440}$$

Finally, following the same lines, the transformation of $\delta F^\dagger$ can be expressed as

$$\delta F^\dagger = -i\sqrt{2}(\bar\xi^{\dot\beta})^\dagger(\partial_\mu\Psi^\alpha)^\dagger\big[(\sigma^\mu)_{\alpha\dot\beta}\big]^* \tag{441}$$

$$= -i\sqrt{2}\xi^\beta(\partial_\mu\overline\Psi^{\dot\alpha})(\sigma^\mu)_{\beta\dot\alpha} \tag{442}$$

We are now in a position where one can compute the variation of the super-symmetric Lagrangian. Collecting the above results, one obtains

$$\delta\mathcal{L}_{\text{W-Z}} = -\eta^{\mu\nu}\partial_\mu\delta\Phi^\dagger\partial_\nu\Phi - \eta^{\mu\nu}\partial_\mu\Phi^\dagger\partial_\nu\delta\Phi - i\delta\overline{\Psi}_{\dot\alpha}(\overline\sigma^\mu)^{\dot\alpha\alpha}\partial_\mu\Psi_\alpha - i\overline{\Psi}_{\dot\alpha}(\overline\sigma^\mu)^{\dot\alpha\alpha}\partial_\mu\delta\Psi_\alpha - \delta F^\dagger F - F^\dagger\delta F \tag{443}$$

$$= -i\sqrt{2}F^\dagger\overline\xi_{\dot\alpha}(\overline\sigma^\mu)^{\dot\alpha\alpha}\partial_\mu\Psi_\alpha - i\sqrt{2}F^\dagger(\partial_\mu\Psi^\alpha)(\sigma^\mu)_{\alpha\dot\beta}\overline\xi^{\dot\beta} - i\sqrt{2}\,\overline\Psi_{\dot\alpha}(\overline\sigma^\mu)^{\dot\alpha\alpha}\xi_\alpha\partial_\mu F$$

$$+ i\sqrt{2}\xi^\beta(\sigma^\mu)_{\beta\dot\alpha}(\partial_\mu\overline\Psi^{\dot\alpha})F + \sqrt{2}\eta^{\mu\nu}\overline\xi^{\dot\alpha}\partial_\mu\overline\Psi_{\dot\alpha}\partial_\nu\Phi - \sqrt{2}(\overline\sigma^\mu)^{\dot\alpha\alpha}(\sigma^\nu)_{\alpha\dot\beta}\overline\Psi_{\dot\alpha}\overline\xi^{\dot\beta}\partial_\mu\partial_\nu\Phi$$

$$- \sqrt{2}\eta^{\mu\nu}\xi^\alpha\partial_\nu\Psi_\alpha\partial_\mu\Phi^\dagger + \sqrt{2}(\sigma^\nu)_{\beta\dot\alpha}(\overline\sigma^\mu)^{\dot\alpha\alpha}\xi^\beta\partial_\mu\Psi_\alpha\partial_\nu\Phi^\dagger \tag{444}$$

Each term in this expression needs to be studied in detail. The first term is also equal to $i\sqrt{2}F^\dagger\partial_\mu\Psi^\alpha(\sigma^\mu)_{\alpha\dot\alpha}\overline\xi^{\dot\alpha}$ and, therefore, exactly cancels the second one. The third one is, after integration by parts, $i\sqrt{2}(\partial_\mu\overline\Psi_{\dot\alpha})(\overline\sigma)^{\dot\alpha\alpha}\xi_\alpha F$, which is also equal to $-i\sqrt{2}\xi^\alpha(\sigma^\mu)_{\alpha\dot\alpha}(\partial_\mu\overline\Psi^{\dot\alpha})F$. This cancels exactly the fourth term. As a consequence, all the term involving the field F give a total derivative contribution only. It remains the four terms involving the scalar field Φ and $\Phi^\dagger$. We first focus on the terms involving Φ only. For this purpose, we use the following relation [19,74]:

$$(\overline\sigma^\mu)^{\dot\alpha\alpha}(\sigma^\nu)_{\alpha\dot\beta} = \eta^{\mu\nu}\delta^{\dot\alpha}{}_{\dot\beta} + 2(\overline\sigma^{\mu\nu})^{\dot\alpha}{}_{\dot\beta} \tag{445}$$

where

$$\overline\sigma^{\mu\nu} \equiv \frac{1}{4}(\overline\sigma^\mu\sigma^\nu - \overline\sigma^\nu\sigma^\mu) = -\overline\sigma^{\nu\mu} \tag{446}$$

Therefore, the second term involving Φ [i.e. the sixth term in the above formula (444)] can be re-written as

$$-\sqrt{2}(\overline\sigma^\mu)^{\dot\alpha\alpha}(\sigma^\nu)_{\alpha\dot\beta}\overline\Psi_{\dot\alpha}\overline\xi^{\dot\beta}\partial_\mu\partial_\nu\Phi = -\sqrt{2}[\eta^{\mu\nu}\delta^{\dot\alpha}{}_{\dot\beta} + 2(\overline\sigma^{\mu\nu})^{\dot\alpha}{}_{\dot\beta}]\overline\Psi_{\dot\alpha}\overline\xi^{\dot\beta}\partial_\mu\partial_\nu\Phi$$

$$= -\sqrt{2}\eta^{\mu\nu}\overline\Psi_{\dot\alpha}\overline\xi^{\dot\alpha}\partial_\mu\partial_\nu\Phi \tag{447}$$

the second term in this equation leading to a vanishing contribution since $\overline\sigma^{\mu\nu}$ is anti-symmetric. The next manipulation is to integrate by part. This gives

$$-\sqrt{2}(\overline\sigma^\mu)^{\dot\alpha\alpha}(\sigma^\nu)_{\alpha\dot\beta}\overline\Psi_{\dot\alpha}\overline\xi^{\dot\beta}\partial_\mu\partial_\nu\Phi = -\sqrt{2}\eta^{\mu\nu}\overline\Psi_{\dot\alpha}\overline\xi^{\dot\alpha}\partial_\mu\partial_\nu\Phi = \sqrt{2}\eta^{\mu\nu}\partial_\mu\overline\Psi_{\dot\alpha}\overline\xi^{\dot\alpha}\partial_\nu\Phi$$

$$= \sqrt{2}\eta^{\mu\nu}\overline\xi_{\dot\alpha}\partial_\mu\overline\Psi^{\dot\alpha}\partial_\nu\Phi = -\sqrt{2}\eta^{\mu\nu}\overline\xi^{\dot\alpha}\partial_\mu\overline\Psi_{\dot\alpha}\partial_\nu\Phi \tag{448}$$

Therefore, this term exactly cancels the first term involving Φ [or the fifth term in the formula (444)]. Clearly, this also works for the two terms involving $\Phi^\dagger$. We have thus shown, without using the equation of motion, that the variation of the Wess–Zumino Lagrangian (431) is a total derivative. As a consequence, the Wess–Zumino model is indeed super-symmetric, that is to say invariant under the super-symmetric transformations (432), (433) and (434).

We have seen before that the constant spinor ξ^α was the equivalent of the vector ϵ^j in the case of the SU(2) transformation considered at the beginning of this section. But what is the equivalent of the generator $\sigma_j/2$ for a super-symmetric transformation? Moreover, we know that the generators obey the formula $[\sigma_i/2, \sigma_j/2] = i\epsilon_{ijk}\sigma_k/2$. What is the corresponding equation satisfied by the super-symmetric generators? We now turn to these questions.

Using again the analogy with the SU(2) transformation, we expect that an infinitesimal super-symmetric transformation to have the form

$$\delta_\xi\Phi \simeq i(\xi^\alpha Q_\alpha + \overline\xi_{\dot\alpha}\overline{Q}^{\dot\alpha})\Phi \tag{449}$$

Here Q_α is the super-symmetric generator, the equivalent of $\sigma_j/2$. Clearly, since $\xi^\alpha Q_\alpha$ must be a scalar, Q_α is a Weyl spinor as the notation indicates.

In order to find the super-symmetric algebra, we first compute the quantity $(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)\Phi$ using Eqs. (432), (433) and (434). It is straightforward to show that

$$(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)\Phi = 2\xi^\alpha\eta_\alpha F - 2i\xi^\alpha(\sigma^\mu)_{\alpha\dot\beta}\overline\eta^{\dot\beta}\partial_\mu\Phi$$

$$-2\eta^\alpha\xi_\alpha F + 2i\eta^\alpha(\sigma^\mu)_{\alpha\dot\beta}\overline\xi^{\dot\beta}\partial_\mu\Phi \tag{450}$$

$$= -2i\xi^\alpha(\sigma^\mu)_{\alpha\dot\beta}\overline\eta^{\dot\beta}\partial_\mu\Phi + 2i\eta^\alpha(\sigma^\mu)_{\alpha\dot\beta}\overline\xi^{\dot\beta}\partial_\mu\Phi \tag{451}$$

since $\xi^\alpha\eta_\alpha = \eta^\alpha\xi_\alpha$. But, of course, we can evaluate the same commutator using the expression (449) for the infinitesimal super-symmetric transformation. This leads to

$$(\delta_\eta\delta_\xi - \delta_\xi\delta_\eta)\Phi = -(\eta^\alpha Q_\alpha + \overline\eta_{\dot\alpha}\overline{Q}^{\dot\alpha})(\xi^\beta Q_\beta + \overline\xi_{\dot\beta}\overline{Q}^{\dot\beta})\Phi$$

$$+ (\xi^\alpha Q_\alpha + \overline\xi_{\dot\alpha}\overline{Q}^{\dot\alpha})(\eta^\beta Q_\beta + \overline\eta_{\dot\beta}\overline{Q}^{\dot\beta})\Phi \tag{452}$$

Then, expanding this expression (carefully taking into account the fact that the spinor components are Grassmann variables), we arrive at

$$(\delta_\eta \delta_\xi - \delta_\xi \delta_\eta)\Phi = \eta^\alpha (Q_\alpha Q_\beta + Q_\beta Q_\alpha)\xi^\beta \Phi$$
$$+ \bar{\eta}_{\dot{\alpha}}\left(\overline{Q}^{\dot{\alpha}}\overline{Q}^{\dot{\beta}} + \overline{Q}^{\dot{\beta}}\overline{Q}^{\dot{\alpha}}\right)\bar{\xi}_{\dot{\beta}}\Phi + \eta^\alpha\left(Q_\alpha \overline{Q}^{\dot{\beta}} + \overline{Q}^{\dot{\beta}}Q_\alpha\right)\bar{\xi}_{\dot{\beta}}\Phi - \xi^\alpha\left(Q_\alpha \overline{Q}^{\dot{\beta}} + \overline{Q}^{\dot{\beta}}Q_\alpha\right)\bar{\eta}_{\dot{\beta}}\Phi \quad (453)$$

One must now compare Eq. (451) with Eq. (453). First of all, we notice that, in Eq. (451), there is no term with two undotted or two dotted spinors ξ and η. This implies that

$$\{Q_\alpha, Q_\beta\} = \left\{Q^{\dot{\alpha}}, Q^{\dot{\beta}}\right\} = 0 \tag{454}$$

Then, a comparison of the first "cross-terms" leads to

$$\{Q_\alpha, \overline{Q}_{\dot{\beta}}\}\Phi = -2i\left(\sigma^\mu\right)_{\alpha\dot{\beta}}\partial_\mu \Phi \tag{455}$$

and a similar relation for the second cross-term. Remembering that $P_\mu = -i\partial_\mu$ (the minus sign is in fact just a convention; the problem with super-symmetry is that the literature contains many different conventions; a nice treatment of this question can be found in the first problem of Ref. [75], Appendix C), one arrives at

$$\{Q_\alpha, \overline{Q}_{\dot{\beta}}\} = 2\left(\sigma^\mu\right)_{\alpha\dot{\beta}}P_\mu \tag{456}$$

This equation shows that two successive super-symmetric transformations are in fact equivalent to a space–time translation. As we will see, this has far-reaching consequences. However, for the subject discussed here, the above formula has also very important implications. Indeed, working it out explicitly, one has

$$Q_1 \overline{Q}_{\dot{1}} + \overline{Q}_{\dot{1}} Q_1 = 2P_3 + 2P_0 \tag{457}$$
$$Q_2 \overline{Q}_{\dot{2}} + \overline{Q}_{\dot{2}} Q_2 = -2P_3 + 2P_0 \tag{458}$$

from which we deduce

$$4P_0 = Q_1 \overline{Q}_{\dot{1}} + \overline{Q}_{\dot{1}} Q_1 + Q_2 \overline{Q}_{\dot{2}} + \overline{Q}_{\dot{2}} Q_2 \tag{459}$$

Therefore, since the Hamiltonian is nothing but the time translation generator P_0, this implies that

$$\langle 0|H|0\rangle = \frac{1}{4}\langle 0|Q_1 \overline{Q}_{\dot{1}} + \overline{Q}_{\dot{1}} Q_1 + Q_2 \overline{Q}_{\dot{2}} + \overline{Q}_{\dot{2}} Q_2|0\rangle \geqslant 0 \tag{460}$$

We have learned two things: firstly, if the vacuum state is not super-symmetric, i.e. if super-symmetry is spontaneously broken (the underlying theory is super-symmetric but the solution or the state in which the system is placed is not), then the vacuum energy is necessarily positive. Secondly, another remarkable consequence of the above results is that, when the vacuum state is super-symmetric, that is to say when

$$Q_\alpha|0\rangle = 0 \tag{461}$$

one has $\langle 0|H|0\rangle = 0$, i.e. the vacuum energy is automatically zero. We have thus proven that the cancellation discussed at the beginning of this section is in fact not a coincidence. Moreover, we have identified under which conditions it is valid: it is always true for a super-symmetric system. This is the reason why super-symmetry has the potential to solve the cosmological constant problem. It gives a concrete example of a symmetry which forces the vacuum energy to vanish.

Unfortunately, we know that super-symmetry must be broken in the real world. This comes from the fact that each boson (fermion) of the standard model is not observed to have a super-symmetric partner of the same mass. This means that vacuum energy is now given by Eq. (421). Let us evaluate it with our "incorrect" method which consists in regularizing the divergent integrals with a cut-off M, see Eq. (74). We find

$$\langle\rho\rangle_{\text{SUSY}} = \frac{M^4}{8\pi^2}\left(1 + \frac{m_\text{B}^2}{M^2} + \cdots\right) - \frac{M^4}{8\pi^2}\left(1 + \frac{m_\text{F}^2}{M^2} + \cdots\right) \tag{462}$$

$$= \frac{M^2}{8\pi^2}\left(m_\text{B}^2 - m_\text{F}^2\right) \tag{463}$$

We see that, even if super-symmetry is broken, the most divergent part $\propto M^4$ of the energy density cancel out. If we now properly estimate the energy density by means of dimensional regularization (95), one obtains that

$$\langle\rho\rangle_{\text{SUSY}} = \frac{m_\text{B}^4}{32\pi^2}\ln\left(\frac{m_\text{B}^4}{\mu^2}\right) - \frac{m_\text{F}^4}{32\pi^2}\ln\left(\frac{m_\text{F}^4}{\mu^2}\right) \tag{464}$$

Therefore, roughly speaking, one can estimate that vacuum energy is now given by $\simeq M_{\text{SUSY}}^4$ where M_{SUSY} is the typical difference between the mass of the super-partners, which is also the super-symmetric breaking mass.

A last remark is in order. In this section we have considered super-symmetry in flat space–time while the cosmological constant problem is formulated in curved space–time. In order to be consistent, we should therefore treat super-symmetry

in this last context. As is well known, this leads to super-gravity [19,75]. Regarding the vacuum energy question, super-gravity presents important differences compared to super-symmetry. In particular, one loses the result that vacuum energy is always positive. But, this also shows that discussing the zero-point energy density issue in curved space–time can bring new aspects to the problem. This is the reason why, in the next section, we address this question.

8. The vacuum energy density in curved space–time

In the previous sections, we have computed the vacuum energy density in flat space–time. This may seem inappropriate since we have argued that the problem occurs only in curved space–time. Here we show how to justify the previous approach. Intuitively, this is justified because we deal with ultra-violet divergences for which the large scale structure of the curved manifold should not play a too important role. In other words, in the ultra-violet regime, one only probes the local properties of space–time and, locally, one cannot distinguish a curved manifold from the Minkowski space–time. However, we need to put this intuitive reasoning on solid grounds. For this purpose, let us consider our toy model again where matter is simply represented by a scalar field. The only difference with the previous sections is that, now, this quantum field lives in a curved space–time. Therefore, in order to have meaningful Einstein equations, one must assume that the quantum average of the stress–energy tensor sources the Einstein equations. In other words, we start from [53]

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda_{\mathrm{B}}g_{\mu\nu} = \kappa\langle T_{\mu\nu}\rangle \tag{465}$$

Then, by analogy with Eq. (3), we define the effective action W to be

$$\langle T_{\mu\nu}\rangle = -\frac{2}{\sqrt{-g}}\frac{\delta W}{\delta g^{\mu\nu}} \tag{466}$$

The effective action can be written in term of the sourceless generating functional $Z[0]$, namely

$$W = -i\ln Z[0] \tag{467}$$

where we recall that the explicit expression of $Z[0]$ is given by, see also Eq. (96),

$$Z[0] = \int \mathcal{D}\phi\, e^{iS_{\mathrm{matter}}[\phi]} \tag{468}$$

the quantity S_{matter} being the action of the scalar field. The above result can be proven in the following way. Varying $Z[0]$, one obtains

$$\delta Z[0] = \int \mathcal{D}\phi\, i\delta S_{\mathrm{matter}}\, e^{iS_{\mathrm{matter}}[\phi]} \tag{469}$$

$$= \langle 0|i\delta S_{\mathrm{matter}}|0\rangle \tag{470}$$

As a consequence,

$$-\frac{2}{\sqrt{-g}}\frac{\delta W}{\delta g^{\mu\nu}} = -\frac{2}{\sqrt{-g}}\frac{-i}{Z[0]}\frac{\delta Z[0]}{\delta g^{\mu\nu}} \tag{471}$$

$$= \frac{1}{Z[0]}\left\langle 0\left|-\frac{2}{\sqrt{-g}}\frac{\delta S_{\mathrm{matter}}}{\delta g^{\mu\nu}}\right|0\right\rangle \tag{472}$$

$$= \frac{1}{Z[0]}\langle 0|T_{\mu\nu}|0\rangle = \langle T_{\mu\nu}\rangle \tag{473}$$

since the mean value is defined by $\langle\cdots\rangle \equiv \langle 0|\cdots|0\rangle/\langle 0|0\rangle$.

The next step is to calculate W explicitly in our case, i.e. for a scalar field. We now show that the effective action can be expressed in terms of the Green function. The Green function is defined by the following expression [53]

$$\left(-g^{\mu\nu}\nabla_\mu\nabla_\nu + m^2\right)G_{\mathrm{F}}(x, y) = \frac{1}{\sqrt{-g}}\delta(x - y) \tag{474}$$

It is easy to show that $G_{\mathrm{F}}(x, y) = iD_{\mathrm{F}}(x - y)$, where the propagator has been defined in Eq. (100). Integrating by part the action of the scalar field $\Phi(x)$, the generating function can be written as

$$Z[J] = \int \mathcal{D}\Phi \exp\left[-\frac{i}{2}\int \mathrm{d}^d x\sqrt{-g}\,\Phi(x)\left(-g^{\mu\nu}\nabla_\mu\nabla_\nu + m^2\right)\Phi(x) + i\int \mathrm{d}^d x\sqrt{-g}\,J(x)\Phi(x)\right] \tag{475}$$

where we have considered the d-dimensional case for future convenience. This path integral is a Gaussian integral and, therefore, can be evaluated using the standard techniques. For $J = 0$ one has
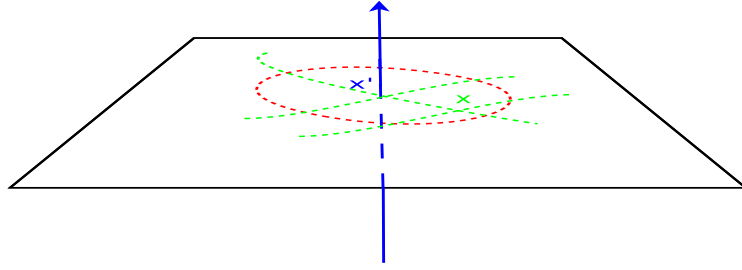
Fig. 4. Sketch of the Riemann coordinates used in order to derive an approximate expression for the Green function, see Eq. (495). The plot aims at illustrating the fact that the Riemann coordinates are a local concept. Since it is sufficient to establish the Green function locally (and not in the entire curved manifold), one can always endow the neighborhood of a point x' (denoted here by the region surrounded by the closed dotted line) with special coordinates x^μ such that the calculation is simplified. It is important to notice that this does not restrict the generality of the obtained results.

$$Z[0] = (2\pi)^{\infty/2}(\det K)^{-1/2} \tag{476}$$

$$= (2\pi)^{\infty/2}e^{\frac{1}{2}\operatorname{tr}(\ln K^{-1})} \tag{477}$$

where K is the following operator, appearing in Eq. (475),

$$K(x, y) \equiv i\sqrt{-g}\big(-g^{\mu\nu}\nabla_\mu\nabla_\nu + m^2\big)\delta^d(x - y) \tag{478}$$

But the inverse of the operator K is in fact the propagator $D_{\rm F}(x, y)$. A pedagogical, but simplistic, demonstration of this fact is the following one. The generating functional can be written as

$$Z[0] = \int \Pi_x \, d\Phi_i \, e^{-\frac{1}{2}\sum_{x,y}\Phi_x K_{xy}\Phi_y} \tag{479}$$

where, in order to reproduce Eq. (478) (or, rather, to mimic this formula), the matrix K_{xy} is taken to be $K_{xy} = K\delta_{xy}$ and the "number K is defined" by $K \equiv i\sqrt{-g}(-g^{\mu\nu}\nabla_\mu\nabla_\nu + m^2)$. Then, the inverse of the matrix K must obey

$$\sum_z K_{xz}K_{zy}^{-1} = \delta_{xy} \tag{480}$$

that is to say, using the definition of K_{xy},

$$K K_{xy}^{-1} = \delta_{xy} \tag{481}$$

But this equation is nothing but Eq. (474). This shows that, as announced, $K_{xy}^{-1} = (D_{\rm F})_{xy}$. As a consequence, one can write

$$Z[0] = (2\pi)^{\infty/2}e^{\frac{1}{2}\operatorname{tr}(\ln D_{\rm F})} \tag{482}$$

or,

$$W = -\frac{i}{2}\operatorname{tr}(\ln D_{\rm F}) - i\ln\big[(2\pi)^{\infty/2}\big] \tag{483}$$

We have reduced the evaluation of the effective action to the evaluation of the Green function (or the propagator) in curved space–time. We now turn to this question.

For this purpose, we introduce the Riemann normal coordinates x^μ with origin at the point x', see Fig. 4 and Refs. [76–78]. As it will become gradually clear, the motivation for such a choice is as follows. Since we want to see how the Einstein equations are modified, one only needs to evaluate the Green function locally. In this case, one can introduce coordinates that drastically simplify the problem, i.e. the calculation. This is exactly the role played by the Riemann normal coordinates. In these coordinates, the metric can be written as [76,77]

$$g_{\mu\nu} = \eta_{\mu\nu} - \frac{1}{3}R_{\mu\alpha\nu\beta}x^\alpha x^\beta - \frac{1}{6}R_{\mu\alpha\nu\beta;\gamma}x^\alpha x^\beta x^\gamma + \left(-\frac{1}{20}R_{\mu\alpha\nu\beta;\gamma\delta} + \frac{2}{45}R_{\alpha\mu\beta\lambda}R^\lambda{}_{\gamma\nu\delta}\right)x^\alpha x^\beta x^\gamma x^\delta + \cdots \tag{484}$$

Then, we insert this metric into the Green function equation (474) and retain terms with coefficients involving four derivatives of the metric or fewer. A lengthy calculation leads to

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\overline{G} - \left(m^2 - \frac{1}{6}\right)\overline{G} - \frac{1}{3}R_\alpha{}^\nu\partial_\nu\overline{G} + \frac{1}{3}R^\mu{}_\alpha{}^\nu{}_\beta x^\alpha x^\beta \partial_\mu\partial_\nu\overline{G} + \frac{1}{6}R_{;\alpha}x^\alpha\overline{G} + \left(-\frac{1}{3}R_\alpha{}^\nu{}_{;\beta} + \frac{1}{6}R_{\alpha\beta}{}^{;\nu}\right)x^\alpha x^\beta\partial_\nu\overline{G}$$

$$+ \frac{1}{6}R^\mu{}_\alpha{}^\nu{}_{\beta;\gamma}x^\alpha x^\beta x^\gamma \partial_\mu\partial_\nu\overline{G} + \frac{1}{12}R_{;\alpha\beta}x^\alpha x^\beta\overline{G} + \left(-\frac{1}{30}R_\alpha{}^\lambda R_{\lambda\beta} + \frac{1}{60}R^\kappa{}_\alpha{}^\lambda{}_\beta R_{\kappa\lambda} + \frac{1}{60}R^{\lambda\mu\kappa}{}_\alpha R_{\lambda\mu\kappa\beta} - \frac{1}{120}R_{;\alpha\beta}\right)$$

$$+ \frac{1}{40}\eta^{\mu\nu}\partial_\mu\partial_\nu R_{\alpha\beta}\bigg)x^\alpha x^\beta\overline{G} + \bigg(-\frac{3}{20}R^\nu{}_{\alpha;\beta\gamma} + \frac{1}{10}R_{\alpha\beta}{}^{;\nu}{}_\gamma - \frac{1}{60}R^\kappa{}_\alpha{}^\nu{}_\beta R_{\kappa\lambda} + \frac{1}{15}R^\kappa{}_{\alpha\lambda\beta}R_\kappa{}^\nu{}_\gamma{}^\lambda\bigg)x^\alpha x^\beta x^\gamma\partial_\nu\overline{G}$$

$$+ \bigg(\frac{1}{20}R^\mu{}_\alpha{}^\nu{}_{\beta;\gamma\delta} + \frac{1}{15}R^\mu{}_{\alpha\lambda\beta}R^\lambda{}_\gamma{}^\nu{}_\delta\bigg)x^\alpha x^\beta x^\gamma x^\delta\partial_\mu\partial_\nu\overline{G} = -\delta^d(x) \tag{485}$$

where all the indices are raised or lowered by the Minkowski metric and where we have defined a rescaled Green function according to $G_F(x,y) \equiv g^{-1/4}\overline{G}(x,y)$. The next step is to Fourier transform the above equation. In this way the local expansion which, in real space, looks like an expansion in the coordinates x^μ is transformed, in momentum space, as an expansion in terms of the inverse of k^μ, as appropriate for a local expansion. More precisely, defining

$$\overline{G}(x,x') = \frac{1}{(2\pi)^d}\int d^d k\, e^{i\eta^{\alpha\beta}k_\alpha x_\beta}\overline{G}(k) \tag{486}$$

we can expand the Green function as $\overline{G}(k) = \overline{G}_0(k) + \overline{G}_1(k) + \cdots$ where $\overline{G}_i(k)$ is the quantity appearing in front of a coefficient involving the i-th derivatives of the metric tensor [76,77]. Let us notice that the Green function is a function of x' in the sense that x^μ represents the coordinates of a point x defined such x' is at the origin. Another remark is that we define the Fourier transform of the Green function with an overall factor $(2\pi)^{-d}$ such that the coefficient of the Dirac function in the right-hand side of Eq. (485) is minus one and does not contain factors of π, thanks to the formula $\delta^d(x) = 1/(2\pi)^d\int d^d k\, e^{ikx}$.

Introducing this expansion into Eq. (485) and consistently identifying the terms of a given order leads to the desired result. This iterative procedure gives

$$\overline{G}(k) = \frac{1}{k^2+m^2} + \frac{R}{6}\frac{1}{(k^2+m^2)^2} + \frac{i}{6}R_{;\alpha}\frac{1}{k^2+m^2}\frac{\partial}{\partial k_\alpha}\big[(k^2+m^2)^{-1}\big] + \frac{R^2}{36}\frac{1}{(k^2+m^2)^3}$$

$$+ a_{\alpha\beta}\frac{1}{k^2+m^2}\frac{\partial^2}{\partial k_\alpha\partial k_\beta}\big[(k^2+m^2)^{-1}\big] + \cdots \tag{487}$$

$$= \frac{1}{k^2+m^2} + \frac{R}{6}\frac{1}{(k^2+m^2)^2} + \frac{i}{12}R_{;\alpha}\frac{\partial}{\partial k_\alpha}\big[(k^2+m^2)^{-2}\big] + \frac{1}{3}a_{\alpha\beta}\partial^\alpha\partial^\beta\big[(k^2+m^2)^{-2}\big]$$

$$+ \bigg(\frac{R^2}{36} - \frac{2}{3}a^\lambda{}_\lambda\bigg)\frac{1}{(k^2+m^2)^3} + \cdots \tag{488}$$

where the coefficient $a_{\alpha\beta}$ can be expressed as

$$a_{\alpha\beta} = -\frac{3}{40}R_{;\alpha\beta} - \frac{1}{40}\eta^{\mu\nu}\partial_\mu\partial_\nu R_{\alpha\beta} + \frac{1}{30}R_\alpha{}^\lambda R_{\lambda\beta} - \frac{1}{60}R^\kappa{}_\alpha{}^\lambda{}_\beta R_{\kappa\lambda} + \frac{1}{60}R^{\lambda\mu\kappa}{}_\alpha R_{\lambda\mu\kappa\beta} \tag{489}$$

Inserting this last expression into Eq. (486) leads to

$$\overline{G}(x,x') = \frac{1}{(2\pi)^d}\int d^n k\, e^{i\eta^{\alpha\beta}k_\alpha x_\beta}\bigg[1 + f_1(x,x')\bigg(-\frac{\partial}{\partial m^2}\bigg) + f_2(x,x')\bigg(-\frac{\partial}{\partial m^2}\bigg)^2 + \cdots\bigg]\frac{1}{k^2+m^2} \tag{490}$$

with

$$f_1(x,x') = \frac{R}{6} + \frac{1}{12}R_{;\alpha}x^\alpha - \frac{1}{3}a_{\alpha\beta}x^{\alpha\beta} \tag{491}$$

$$f_2(x,x') = \frac{R^2}{72} - \frac{1}{3}a^\lambda{}_\lambda \tag{492}$$

The final step of the calculation consists in expressing the Green function in the so-called proper time formalism [76,77]. Firstly, we replace the quantity $(k^2+m^2)^{-1}$ with

$$\frac{1}{k^2+m^2} = \int\limits_0^\infty i\, ds\, e^{-is(k^2+m^2)} \tag{493}$$

and, secondly, we make use of the result

$$\frac{1}{(2\pi)^d}\int d^n k\, e^{i\eta^{\alpha\beta}k_\alpha x_\beta - is(k^2+m^2)} = \frac{i}{(4\pi)^{d/2}}(is)^{-d/2}e^{-ism^2+i\sigma/(2s)} \tag{494}$$

where $\sigma = \eta^{\alpha\beta}x_\alpha x_\beta/2$. This leads to our final expression for the Green function, namely

$$\overline{G}(x,x') = \frac{i}{(4\pi)^{d/2}}\int\limits_0^\infty \frac{i\, ds}{(is)^{d/2}}e^{-im^2 s - \sigma/(2is)}\big[1 + is f_1 + (is)^2 f_2 + \cdots\big] \tag{495}$$

Of course, one could always push the expansion to higher orders in inverse of the momenta in order to obtain higher orders in s. But, in the present context, this will not be necessary.

We now have all the tools to evaluate the effective action (483). We now seek an explicit expression for the quantity $\ln(D_F)$. As already mentioned, D_F can be viewed as the inverse of the "continuous matrix" K which means that, formally, one can write

$$D_F = \int_0^\infty e^{-iKs} i \, ds \tag{496}$$

where we have used again Eq. (493). This expression allows us to express the matrix element of e^{-iKs}. Indeed, given that $D_F = -ig^{-1/4}\overline{G}$, a direct comparison of Eq. (495) with the previous relation leads to

$$\left(e^{-iKs}\right)_{xx'} = \frac{g^{-1/4}}{(4\pi)^{d/2}} \frac{e^{-im^2s-\sigma/(2is)}}{(is)^{d/2}}\left[1 + isf_1 + (is)^2 f_2 + \cdots\right] \tag{497}$$

Now, let us consider the following integral

$$\int_\lambda^\infty e^{-iKs}\frac{i\,ds}{is} = -\int_{i\lambda K}^\infty \frac{e^{-t}}{t}\,dt = -\gamma - \ln(i\lambda K) - \sum_{k=1}^\infty \frac{(-1)^{k+1}(i\lambda K)^k}{kk!} \tag{498}$$

In the limit $\lambda \to 0$, up to an infinite constant that can be ignored, only the logarithm survives. Therefore, we can define the logarithm of the operator K as

$$-\ln K = \ln D_F = \int_0^\infty e^{-iKs}\frac{i\,ds}{is} \tag{499}$$

Then, using Eq. (497), one obtains

$$(\ln D_F)_{xx'} = \frac{g^{-1/4}(x)}{(4\pi)^{d/2}}\int_0^\infty \frac{e^{-im^2s-\sigma(x,x')/(2is)}}{(is)^{d/2}}\left[1 + isf_1(x,x') + (is)^2 f_2(x,x') + \cdots\right]\frac{i\,ds}{is} \tag{500}$$

As a consequence, we have now an explicit expression for the effective action, namely

$$W = -\frac{i}{2}\int d^dx\sqrt{-g}\lim_{x'\to x}(\ln D_F)_{xx'} \tag{501}$$

or

$$W = \int d^dx\sqrt{-g}\left\{-\frac{i}{2}\frac{1}{(4\pi)^{d/2}}\int_0^\infty \frac{e^{-im^2s}}{(is)^{d/2}}\left[1 + isf_1 + (is)^2 f_2 + \cdots\right]\frac{i\,ds}{is}\right\} \tag{502}$$

In the coincident limit $x' \to x$, the determinant of the metric tensor has been taken to one since it is calculated in x', that is to say for $x = 0$. For the same reason, we also have considered that $\sigma = 0$. The above expression can be evaluated explicitly in terms of Euler function and one obtains

$$W = \int d^dx\sqrt{-g}\left\{-\frac{i}{2}\frac{1}{(4\pi)^{d/2}}\left(\frac{m}{\mu}\right)^{d-4}\left[m^4\Gamma\left(-\frac{d}{2}\right) + m^2 f_1\Gamma\left(1-\frac{d}{2}\right) + f_2\Gamma\left(2-\frac{d}{2}\right) + \cdots\right]\right\} \tag{503}$$

where we have introduced the scale μ in order to maintain the correct dimension of the effective action. The above expression can be analyzed in a dimensional regularization scheme. In this case, one is led to the following expression:

$$iW = \int d^dx\sqrt{-g}\left\{\left[\frac{2}{\epsilon} - \gamma - \ln\left(\frac{m^2}{4\pi\mu^2}\right)\right]\left(\frac{m^4}{64\pi^2} - \frac{m^2 f_1}{32\pi^2} + \frac{f_2}{32\pi^2}\right) + \cdots\right\} \tag{504}$$

which is our final result. The i factor can be absorbed into a constant shift so we do not need to worry about it. If we compare the effective action with the original Lagrangian (1), we see that the first divergent term in the above expression can be absorbed into a redefinition of the cosmological constant (the second term can be viewed as a redefinition of the Newton constant and the third term leads to the appearance of new terms in the action, see Ref. [53]). Explicitly, one has

$$\Lambda_{\text{eff}} = \Lambda_B + \kappa\frac{m^4}{64\pi^2}\left[\frac{2}{\epsilon} - \gamma - \ln\left(\frac{m^2}{4\pi\mu^2}\right)\right] \tag{505}$$

that is to say exactly the same expression as before (except an unimportant factor 3/2), see for instance Eq. (94). As a consequence in a $\overline{\text{MS}}$ scheme [16], we can remove the divergence and obtain exactly the same expression derived before, see Eq. (95), namely

$$\Lambda_{\text{eff}} = \Lambda_{\text{B}} + \kappa \frac{m^4}{64\pi^2} \ln\left(\frac{m^2}{\mu^2}\right) \tag{506}$$

The calculations of this section justify our previous approach in flat space–time. Indeed, we have just proven that all the results derived before in flat space–time are in fact valid in the more rigorous approach where the curvature of space–time is properly taken into account. This confirms that the vacuum energy does not scale as the cut-off to the power four, see Eqs. (74) and (77) but is in fact proportional to the mass of the particle to the power four times a logarithmic factor depending on the renormalization scale. In some sense, the calculation in flat space–time can be considered as a computational trick: it is sufficient to calculate the vacuum energy in this simple framework since, from the previous considerations, we know that this is also the result that consistently emerges from an approach where the curvature of space–time is properly included. Of course, physically, the reason for this success is not so surprising and was already mentioned before. Since we deal with the ultraviolet behavior of the theory, a local analysis should necessarily lead to a correct and consistent result.

It is also worth mentioning that Eqs. (95) and (506) recently attracted lot of attention, see Refs. [47,51]. In particular, as we discussed at length in Section 4.1 and as was emphasized in these articles, these equations are the consequence of the fact that the regularization scheme used to tame the divergence of the vacuum energy must respect Lorentz invariance. In fact, these equations have been known for a long time, see in particular Eq. (6.50) of the standard textbook [53]. Therefore, the fact that the cosmological constant is proportional to the mass of the particle to the power four, and not to the cut-off to the power four as usually claimed, see again Eqs. (74) and (77), is in fact not a new result. Of course, it modifies a lot our estimate of vacuum energy. For instance, the correct result gives zero for the photons while, for the same situation, the result based on the wrong regularization scheme gives infinity. Clearly, this means that the correct regularization scheme leads to a cosmological constant much smaller than the one obtained from the wrong approach. As a consequence, one could hope that the cosmological constant problem is in fact just an artifact due to the use of an incorrect regularization method. We turn to this question in the next section but, unfortunately, we will see that this is not the case.

9. The value of the cosmological constant

We are now in a position to conclude the first part of this review. The considerations of the previous sections lead to the following expression for the vacuum energy [51]

$$\rho_{\text{vac}} = \sum_i n_i \frac{m_i^4}{64\pi^2} \ln\left(\frac{m_i^2}{\mu^2}\right) + \rho_{\text{B}} + \rho_{\text{vac}}^{\text{EW}} + \rho_{\text{vac}}^{\text{QCD}} + \cdots \tag{507}$$

where ρ_{B} indicates the contribution coming from the bared cosmological constant and where the dots means that other phase transition could contribute. Let us now evaluate the first term more precisely for the standard model of particles physics [51]. In this case, one has one scalar field, the Higgs boson with $n_{\text{H}} = 1$, $m_{\text{H}} \simeq 125$ GeV, the six quarks (fermions) for which $n_{\text{quarks}} = -4$ and $m_{\text{t}} \simeq 171.2$ GeV, $m_{\text{b}} \simeq 4.2$ GeV, $m_{\text{c}} \simeq 1.27$ GeV, $m_{\text{s}} \simeq 0.104$ GeV, $m_{\text{u}} \simeq 0.24$ GeV and $m_{\text{d}} \simeq 0.48$ GeV, the leptons (fermions) with $n_{\text{leptons}} = 4$ and $m_{\text{e}} \simeq 0.511$ MeV, $m_{\mu} \simeq 105$ MeV, $m_{\tau} \simeq 1.77$ MeV, the neutrinos (fermions) the mass of which is so small that we can ignore them and, finally, the gauge bosons (massive vector fields), $n_{\text{Z}} = 3$, $m_{\text{Z}} \simeq 91$ GeV and $n_{\text{W}^{\pm}} = 3$, $m_{\text{W}^{\pm}} \simeq 80$ GeV. Of course, we also have the photon but since $m_{\gamma} = 0$, as already mentioned before, it does not contribute to the vacuum energy. The only piece missing in order to calculate ρ_{vac} in Eq. (507) is the renormalization scale μ. It was argued in Ref. [51] that, because we use the photons coming from the supernovae to determine the cosmological constant and because these photons couple to the metric the expansion rate of which is characterized by the Hubble constant, one should take $\mu \sim \sqrt{E_{\gamma} E_{\text{grav}}}$ with $E_{\text{grav}} \simeq H_0 \simeq 3.7 \times 10^{-41}$ GeV and with the energy of the photons corresponding to the wavelength $\lambda \simeq 500$ nm. This leads to $\mu \simeq 3 \times 10^{-25}$ GeV and implies that Eq. (507) reads

$$\rho_{\text{vac}} \simeq -2 \times 10^8 \text{ GeV}^4 + \rho_{\text{B}} + \rho_{\text{vac}}^{\text{EW}} + \rho_{\text{vac}}^{\text{QCD}} + \cdots \tag{508}$$

If one does not want to estimate μ, one can simply plot ρ_{vac} as a function of μ as done in Fig. 5. Clearly, regardless of the precise value of μ, we are very far from $\rho_{\text{vac}} \simeq 10^{72}$ GeV^4 always mentioned in the literature.

At this stage, Eq. (508) can be seen as the "prediction" of the standard model for the vacuum energy. At first sight, the calculation is straightforward and, in order to obtain the number in Eq. (508), we have used techniques of regularization that are known to work in similar contexts (i.e. when one computes a cross-section in particle physics) and to lead to a very good agreement with experiments. Therefore, we now turn to the question of what is known experimentally about the vacuum energy.
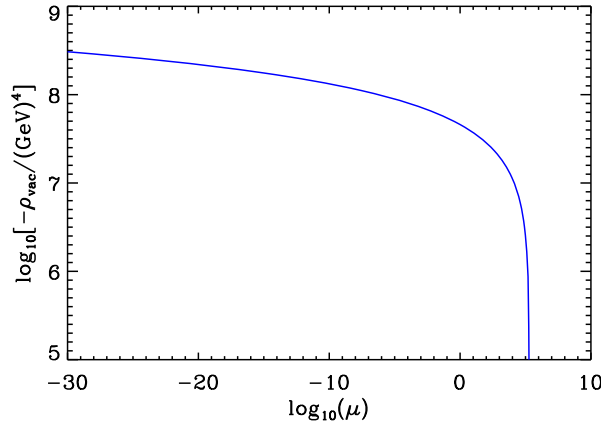
Fig. 5. Evolution of the vacuum energy density [more precisely the first term in Eq. (507) versus the renormalization scale μ]. In the range considered here, the vacuum energy density is negative. The "divergence" observed around $\log_{10} \mu \simeq 5$ does not correspond to a new physical effect but just signal that ρ_{vac} becomes positive.

10. Measuring the cosmological constant in cosmology

Having discussed how the vacuum energy can be calculated, we now address the question of how it is determined and/or constrained experimentally/observationally. Here, we should distinguish the case of cosmology, treated in this section, and the other experiments, treated in the following sections. The difference is that, in cosmology, one claims a detection of Λ while, as we will see, the other experiments can only put an upper limit on its value. Moreover, in cosmology, there are in fact many different and independent observables that can be used to measure the cosmological constant [23,24,79–82]. All these techniques are reviewed in other articles [40,41] and, therefore, it is not our purpose here to discuss this question in details. However, for completeness and also because this has some impact on the theoretical status of the cosmological constant problem, we now discuss how Λ can be determined. We will do so from a theoretical perspective leaving, as mentioned before, the technical aspects to other articles [40,41].

10.1. The accelerating Universe

As is well known, the claim that vacuum energy has been measured (and not only constrained) was first established from the discovery that the expansion of the Universe is accelerated (although one can find many claims about the value of the cosmological constant in the history of cosmology) [23,24]. Since the above mentioned claim has clearly very important consequences for physics, one needs to be very precise at this stage and this is why, in what follows, we will review the exact origin of this result.

In order to construct a cosmological model we apply general relativity to the Universe as a whole [83]. The cosmological principle implies that the Universe is, on large scales, homogeneous and isotropic. As a consequence, the metric is given by the following expression

$$ds^2 = -dt^2 + a^2(t)\gamma^{(3)}_{ij}\, dx^i\, dx^j \tag{509}$$

where $\gamma^{(3)}_{ij}$ is the metric of the three-dimensional space-like sections and where $a(t)$ is the scale factor. The expansion is characterized by the Hubble parameter, $H = \dot{a}/a$ where a dot means a derivative with respect to cosmic time. In the standard model of cosmology, matter is assumed to be a collection of N perfect fluids (at least in the simplest version) and, as a consequence, its stress–energy tensor is given by the following expression

$$T_{\mu\nu} = \sum_{i=1}^{N} T^{(i)}_{\mu\nu} = (\rho_T + p_T)u_\mu u_\nu + p_T g_{\mu\nu} \tag{510}$$

where ρ_T is the (total) energy density and p_T the (total) pressure. These two quantities are related by the equation of state

$$p_T = w(\rho_T) \tag{511}$$

[in general, there is an equation of state per fluid considered i.e. $p_i = w_i(\rho_i)$]. The vector u_μ is the four velocity and satisfies the relation $u_\mu u^\mu = -1$. In terms of cosmic time this means that $u^\mu = (1, \mathbf{0})$. The fact that the stress–energy tensor is conserved, $\nabla^\alpha T_{\alpha\mu} = 0$, amounts to

$$\dot{\rho}_T + 3\frac{\dot{a}}{a}(\rho_T + p_T) = 0 \tag{512}$$

This expression is obtained when one takes $\mu = 0$. The case $\mu = i$ does not lead to an interesting equation for the background. Moreover, if we assume that the fluids are separately conserved, then the above relation is in fact true for each fluid, i.e. $\dot{\rho}_i + 3H(\rho_i + p_i) = 0$.

We are now in a position to write down the Einstein equations (10) for the Universe. We arrive at the following nonlinear, second order, differential equations

$$\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \frac{\kappa}{3} \sum_{i=1}^{N} \rho_i + \frac{\Lambda_{\text{eff}}}{3} \tag{513}$$

$$-\left(2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) = \kappa \sum_{i=1}^{N} p_i - \Lambda_{\text{eff}} \tag{514}$$

From the previous formulas, as already signaled in Eq. (11), we see that we can also define $\rho_{\Lambda_{\text{B}}}$ by $\rho_{\Lambda_{\text{B}}} = \Lambda_{\text{B}}/\kappa$ and $p_{\Lambda_{\text{B}}}$ by $p_{\Lambda_{\text{B}}} = -\rho_{\Lambda_{\text{B}}} = -\Lambda_{\text{B}}/\kappa$. In this case, the cosmological constant is described as fluid with a constant energy density and pressure and with an equation of state $p_{\Lambda_{\text{B}}} + \rho_{\Lambda_{\text{B}}} = 0$, i.e.

$$w_{\text{vac}} \equiv \frac{p_{\Lambda_{\text{B}}}}{\rho_{\Lambda_{\text{B}}}} = -1 \tag{515}$$

This is of course consistent with Eq. (5). It is worth noticing that the vacuum energy is described by a fluid with a negative effective pressure.

Combining the two Einstein equations, one gets an equation which permits to calculate the acceleration of the scale factor

$$\frac{\ddot{a}}{a} = -\frac{\kappa}{6} \sum_{i=1}^{N} (\rho_i + 3p_i) + \frac{2}{3} \Lambda_{\text{eff}} \tag{516}$$

Let us notice that, in order to obtain the above formula, we have not assumed anything about the curvature k. This equation is especially interesting because it gives the condition leading to an accelerated expansion, namely

$$\rho_{\text{T}} + 3p_{\text{T}} < 0 \tag{517}$$

Since the energy density of matter must be positive, we see that the above condition requires a negative pressure, i.e. some exotic form of matter. But, this is of course exactly what happens when the vacuum energy is the dominant fluid in the Universe. Then, Eq. (516) can be written as $\ddot{a}/a \simeq 2\Lambda_{\text{eff}}/3$ which is positive for a positive effective cosmological constant. Therefore, in the framework of the standard model described before, the discovery that $\ddot{a} > 0$ means that a fluid with a negative pressure is driving the expansion.

10.2. The Hubble diagram

In this section, we explain how the measurement of the Hubble diagram can be used to infer the value of the cosmological constant.

Let us first consider a source in flat space–time whose absolute luminosity is given by L (energy emitted per unit of time). This quantity can be written as

$$L = \frac{n_{\text{E}} \hbar \nu}{\Delta t} \tag{518}$$

where n_{E} is the number of photons emitted during the time interval Δt. Suppose that we observe the source from a distance d_{L} using a telescope, the surface of the mirror being ΔS. The number of photons received is given by

$$n_{\text{R}} = n_{\text{E}} \left(\frac{\Delta S}{4\pi d_{\text{L}}^2}\right) \tag{519}$$

On the other hand, the flux of photons (or apparent luminosity i.e. energy per unit of time and surface) through the telescope is, by definition,

$$\Phi_{\text{R}} = \frac{n_{\text{R}} \hbar \nu}{\Delta S \Delta t} \tag{520}$$

from which we deduce that

$$\Phi_{\text{R}} = \frac{L}{4\pi d_{\text{L}}^2} \tag{521}$$

We have just recovered the well-known fact that the apparent luminosity of a source decreases as the inverse squared of the distance. On the other hand, we can use this relation as an operational definition of the distance, namely $d_L = \sqrt{L/(4\pi \Phi_R)}$.

The previous considerations can be used as a definition of distance in cosmology. For this purpose, let us now consider the same situation but in a curved space–time with the following metric (a form more explicit than the one given in Eq. (509))

$$ds^2 = -dt^2 + a^2(t)\big(dr^2 + r^2\,d\Omega_2^2\big) \tag{522}$$

We assume that the source is located at the co-moving coordinate r_1 while the observer is located at the origin of the coordinates. In order to reproduce the above flat space–time calculation in this new situation, we first need to calculate the surface of the sphere $t = t_R = $ const., $r = r_1 = $ const. surrounding the source. The surface is just given by the integral of the covariant volume element, namely

$$S = \int \sqrt{g}\,d^2x = \int\limits_0^{\pi} \int\limits_0^{2\pi} \sqrt{a_R^4 r_1^4}\,\sin\theta\,d\theta\,d\varphi = 4\pi a_R^2 r_1^2 \tag{523}$$

This means that the number of photons received in the telescope can now be expressed as

$$n_R = n_E \left(\frac{\Delta S}{4\pi a_R^2 r_1^2} \right) \tag{524}$$

But we have also to take into account the two following phenomena. Firstly, the energy of the photons has changed during the propagation from the source to the observer. If λ is the wavelength, we have $\lambda = (2\pi/k)a(t) \propto \nu^{-1}$, from which we deduce that $E_R/E_E = \nu_R/\nu_E = a_E/a_R$, that is to say we have the redshift due to the expansion. Secondly, if the photons are emitted as a burst of duration Δt_E, they will be received as burst of duration $\Delta t_R = \Delta t_E a_R/a_E$. We are now in a position where we can calculate the apparent luminosity. We obtain

$$\Phi_R = \frac{L}{4\pi a_R^2 r_1^2} \left(\frac{a_E}{a_R} \right)^2 = \frac{L}{4\pi a_R^2 r_1^2} \frac{1}{(1+z)^2} \tag{525}$$

where z is the redshift of the source. Now if we use the operational definition of the luminosity distance (521), we find

$$d_L(z) = a_R r_1 (1+z) = a_R (1+z) \int\limits_{t_E}^{t_R} \frac{c\,d\tau}{a(\tau)} \tag{526}$$

We see from the last expression that the distance of a source depends on the behavior of the scale factor between the time of emission and the time of reception. On the contrary, if we measure the luminosity distance versus the redshift, we can learn about this behavior. Measuring the apparent luminosity is quite easy and, clearly, the difficulty lies in estimating the absolute luminosity of the source. It is also convenient to work in terms of the distance modulus, $m - M$, since the data are often presented in this way. The quantity m (or sometimes m_{bol}) is the apparent bolometric magnitude. It is related to the apparent bolometric luminosity ℓ by

$$\ell = 10^{-2m/5} \times 2.52 \times 10^{-5}\ \text{erg/cm}^2 \times \text{sec} \tag{527}$$

For instance the Sun has an apparent bolometric magnitude of -26.85. On the other hand, the absolute bolometric magnitude M is the apparent bolometric magnitude that the source would have at a distance of 10 pc and is given by

$$L = 10^{-2M/5} \times 3.02 \times 10^{35}\ \text{erg/sec} \tag{528}$$

For instance, the absolute bolometric magnitude of the Sun is 4.72. As a consequence, using these two definitions, the distance modulus is defined by

$$d_L(z) \equiv \frac{c}{H_0} \bar{d}_L(z) \equiv 10^{1+(m-M)/5} \times 10^{-6}\ \text{Mpc} \tag{529}$$

such that $\bar{d}_L$ is dimensionless and $c/H_0 = 3000\,h^{-1}$ Mpc is of course nothing but the Hubble length (we have re-established the speed of light for convenience). From the above expression, one immediately deduces that

$$\mu(z) \equiv m - M = 5\log_{10}\bar{d}_L(z) + 25 + 5\log_{10}\left(\frac{c}{H_0} \right) \tag{530}$$

where c is expressed in $\text{km} \times \text{s}^{-1}$, H_0 in $\text{km} \times \text{s}^{-1} \times \text{Mpc}^{-1}$ and d_L in Mpc.

In practice we only have access to d_L over a limited range of redshifts. Therefore, one can derive a general expression by Taylor expanding the scale factor. For small redshifts, we have
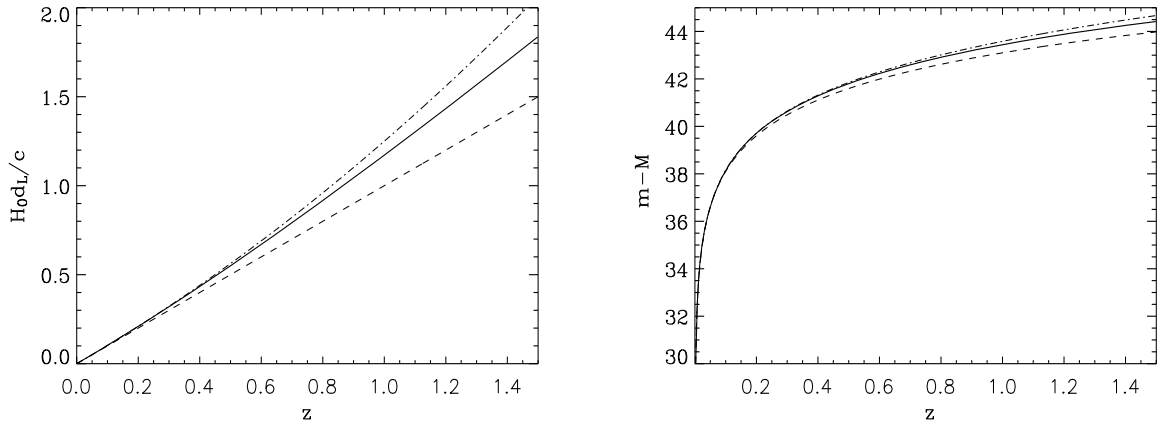
Fig. 6. Luminosity distance $d_L(z)$ (left panel) and distance modulus $\mu(z)$ (right panel) in a matter dominated Universe. The solid line corresponds to the exact result. The dashed line is the linear law $d_L \simeq cz/H_0$ and the dotted dashed line represents the approximated result at second order, see Eq. (536). We have used $h = 0.72$.

$$a(t) = a(t_R) + \dot{a}(t_R)(t - t_R) + \frac{1}{2}\ddot{a}(t_R)(t - t_R)^2 + \cdots = a(t_R)\left[1 + H_R(t - t_R) - \frac{1}{2}H_R^2 q_R(t - t_R)^2 + \cdots\right] \qquad (531)$$

where $q \equiv -a\ddot{a}/(\dot{a})^2$ is the acceleration parameter. A positive acceleration parameter corresponds to a decelerating Universe whether a negative one corresponds to an accelerating Universe. Using the previous expression, the luminosity distance becomes

$$d_L(z) = c(1 + z)\int\limits_{t_E}^{t_R} d\tau\left[1 - H_R(\tau - t_R) + \frac{1}{2}H_R^2 q_R(\tau - t_R)^2 + H_R^2 q_R(\tau - t_R)^2 + \cdots\right] \qquad (532)$$

$$= c(1 + z)\left[-(t_E - t_R) + \frac{1}{2}H_R(t_E - t_R)^2 - \frac{1}{6}H_R^2 q_R(t_E - t_R)^3 - \frac{1}{3}H_R^2 q_R(t_E - t_R)^3 + \cdots\right] \qquad (533)$$

As expected the result only depends on the time of flight $t_E - t_R$. In order to obtain our final result, we must express this time of flight in terms of the redshift. The redshift of the source is defined by $z = a(t_R)/a(t_E) - 1$. Using the Taylor expansion of the scale factor, we end up with

$$z = -H_R(t_E - t_R) + H_R^2\left(1 + \frac{1}{2}q_R\right)(t_E - t_R)^2 + \cdots \qquad (534)$$

which can be easily inverted to give the expression of the time of flight, namely

$$t_E - t_R = -\frac{z}{H_R} + \frac{1}{H_R}\left(1 + \frac{1}{2}q_R\right)z^2 + \cdots \qquad (535)$$

The last step is to replace the above expression into Eq. (533). This gives

$$d_L(z) = \frac{c}{H_0}\left[z + \frac{1}{2}(1 - q_0)z^2 + \cdots\right] \qquad (536)$$

where we have taken into account that the time of reception is the present time (denoted with the subscript "0"). We see that the measurement of the luminosity distance versus the redshift allows us to determine the Hubble parameter today. If one goes sufficiently far in redshifts, one can determine whether the Universe is accelerating or not via the accelerating parameter q_0. It is worth noticing that this last conclusion is independent of the underlying gravity theory (provided this is a metric theory). The typical behavior of the luminosity distance and of the modulus is illustrated in Fig. 6 for a matter dominated Universe. Let us also notice that the method that we have just presented for pedagogical reasons is, of course, not used in practice because we have now measured up to redshifts of $z \simeq 1.4$ for which the Taylor expansion is not valid (and, of course, in the real world the analysis is always more complicated that the simplified version discussed above). In this case, one must use an exact expression, that is to say one must compute the integral in Eq. (526) exactly.

If we now use general relativity (i.e. a specific theory for gravity), one can go further and relates the value of the acceleration parameter to the matter content of the Universe. The critical energy density ρ_{cri}, defined by $\rho_{cri} \equiv 3H^2/\kappa$, is used to evaluate the contribution of each species through the parameter $\Omega_i \equiv \rho_i/\rho_{cri}$ where ρ_i is the energy density of the species i at present time. Then, the Friedman equation can be re-written as

$$\sum_{i=1}^{N} \Omega_i = 1 + \frac{k}{a^2 H^2} \tag{537}$$

and the acceleration parameter can be re-expressed according to

$$q_0 = \frac{1}{2} \sum_{i=1}^{N} \Omega_i (1 + 3w_i) \tag{538}$$

The Universe contains pressure-less gas (cold dark matter and ordinary baryonic matter) for which $w = 0$ but also radiation ($w = 1/3$) and, finally, vacuum energy ($w = -1$). In this case, one obtains

$$q_0 = \frac{1}{2} \Omega_{\mathrm{m}} + \Omega_\gamma - \Omega_{\mathrm{vac}} \tag{539}$$

Today, the contribution of radiation is negligible since we know from the measurement of the CMB that $\Omega_\gamma h^2 \simeq 2.48 \times 10^{-5}$ [81]. Therefore, the acceleration parameter is in fact given by $q_0 \simeq \Omega_{\mathrm{m}}/2 - \Omega_{\mathrm{vac}}$. On the other, we also know that the Universe is spatially flat which means that $\Omega_{\mathrm{m}} + \Omega_{\mathrm{vac}} \simeq 1$ [81]. As a consequence, from the measurement of $q_0 \simeq -0.67 \pm 0.25$ [23,24], one can deduce the two quantities Ω_{m} and Ω_{vac}. On finds that $\Omega_{\mathrm{m}} \simeq 0.3$ and $\Omega_{\mathrm{vac}} \simeq 0.7$. This is how one reaches the conclusion that our Universe is accelerating and that vacuum energy dominates the present energy content. Remarkably, as announced at the beginning of this section, it is a detection and a measurement of ρ_{vac}. One finds

$$\rho_{\mathrm{vac}} \simeq \Omega_{\mathrm{vac}} \rho_{\mathrm{cri}} \simeq 10^{-47} \text{ GeV}^4 \tag{540}$$

a number that should be compared with Eq. (508).

At this stage, several remarks are in order. Firstly, we notice the large mismatch between the theoretical expectation (508) and the above, observationally determined, number (540), something like 54 orders of magnitude (but much less than the 122 orders of magnitude often quoted in the literature). This is, of course, nothing but the cosmological constant problem although we will be more accurate with regards to its definition in the subsequent sections. Something is clearly wrong but, as will be discussed, it is difficult to identify where the mistake is. Secondly, it is clear that the above conclusion rests on many assumptions. For instance, we have assumed that the Universe is homogeneous and isotropic [42] and that gravity is well described by general relativity. This means that, a priori, the cosmological constant problem is present in this context only. Thirdly, we have also implicitly assumed that the reason for the accelerated expansion is vacuum energy. This is clearly a natural idea since there is a term in our theory that must be present since it satisfies all the required properties (covariance, energy conservation) and that precisely leads to the observed phenomenon. A priori, nothing more can be asked to a new theoretical framework. But, given the mismatch discussed before, one can also doubt the identification of the source of the accelerated expansion with vacuum energy. As a matter of fact, it is possible to construct models where the Universe accelerates because of some new source of matter (often named "dark energy"). This is for instance the case of quintessence and/or galileons [43] models where a scalar field is responsible for the acceleration. In this case, one can no longer claim that we have detected the vacuum energy and that its value is given by Eq. (540). However, this does not solve the cosmological constant problem because, even if the source of the acceleration is another fluid, we still have the constraint $\rho_{\mathrm{vac}} < \rho_{\mathrm{cri}}$. Clearly, the number (508) does not satisfy this inequality which means that the problem is still present. In this sense, the constraint deduced from cosmology is very important: even if dark energy is not the cosmological constant, the observation that the energy density today is the critical energy density severely limits the value of Λ_{eff}. Fourthly, in principle, it is possible to check experimentally whether the reason for the accelerated expansion is vacuum energy or something else like quintessence [84–99] (more recently, in Ref. [100], the current acceleration was also explained by a scalar field experiencing quantum fluctuations during inflation). Measuring the equation of state parameter w is the clue since, if this is a constant equals to -1, then one can be sure that the source of the acceleration is the cosmological constant. Unfortunately, this measurement is difficult [101].

One more general grounds, it is clear that it would be better to have constraints, or even a measurement, of vacuum energy in an experimental context which is not cosmology where, as reminded before, measurements are difficult and always subject to many biases and systematic errors. In the next sections, we explore whether this is possible.

11. Measuring the cosmological constant elsewhere than in cosmology?

In the previous section, we have written the Einstein equation in a homogeneous and isotropic Universe. Here, we investigate the influence of the cosmological term in a spherically symmetric situation. This will allow us to study how the presence of Λ_{eff} affects the motion of the planets in our solar system and how it modifies the energy levels of the Hydrogen atom.

11.1. The static and spherically symmetric gravitational field in presence of a cosmological constant

Let us analyze the influence of the cosmological constant in a static and spherically symmetric situation [102]. This means that, in spherical coordinates, one can write the metric tensor as [102]

$$ds^2 = -B(r)\,dt^2 + A(r)\,dr^2 + r^2\big(d\theta^2 + \sin^2\theta\,d\varphi^2\big) \tag{541}$$

where $A(r)$ and $B(r)$ are two free functions. We seek solutions of the Einstein equations (10) with a vanishing stress–energy tensor. It is easy to show that these equations can be expressed as

$$R_{\mu\nu} - \Lambda_{\text{eff}}g_{\mu\nu} = 0 \tag{542}$$

The next step is to calculate the components of the Ricci tensor for the metric tensor given by Eq. (541). Straightforward calculations lead to the following expressions

$$R_{tt} = \frac{B}{A}\left(\frac{1}{2}\frac{B''}{B} - \frac{1}{4}\frac{B'^2}{B^2} - \frac{1}{4}\frac{A'B'}{AB} + \frac{1}{r}\frac{B'}{B}\right) \tag{543}$$

$$R_{rr} = -\frac{1}{2}\frac{B''}{B} + \frac{1}{4}\frac{B'^2}{B^2} + \frac{1}{4}\frac{A'B'}{AB} + \frac{1}{r}\frac{B'}{B} \tag{544}$$

$$R_{\theta\theta} = \frac{1}{A}\left(\frac{r}{2}\frac{A'}{A} - \frac{r}{2}\frac{B'}{B} - 1\right) + 1 \tag{545}$$

$$R_{\varphi\varphi} = R_{\theta\theta}\sin^2\theta \tag{546}$$

the other components being zero. In the above equations, a prime denotes a derivative with respect to the radial coordinate r.

We are now in a position where one can determine the solution of the field equations. If we take the combination $R_{tt}/B + R_{rr}/A = -\Lambda_{\text{eff}}B/B + \Lambda_{\text{eff}}A/A = 0$ [from the Einstein equations (542)], one obtains

$$\frac{1}{rA}\left(\frac{B'}{B} + \frac{A'}{A}\right) = 0 \tag{547}$$

which implies that $B(r) = C/A(r)$, where C is a constant. Using this expression, one can re-express $R_{\theta\theta}$ as

$$R_{\theta\theta} = \frac{B(r)}{C}\left(-r\frac{B'}{B} - 1\right) + 1 \tag{548}$$

and, as a consequence, the corresponding Einstein equation now reads

$$rB'(r) + B(r) = C - C\Lambda_{\text{eff}}r^2 \tag{549}$$

This equation can easily be integrated and one obtains

$$B(r) = D + \frac{C}{r} - D\Lambda_{\text{eff}}\frac{r^2}{3} \tag{550}$$

where D is another integration constant. The constants C and D are fixed in such a way that, in absence of a cosmological constant, one recovers the standard Schwarschild solution. This amounts to take $D = 1$ and $C = 2GM$, where M is the mass of the central body and G the Newton constant. This completes our calculation since the two functions $A(r)$ and $B(r)$ are now completely specified, namely

$$A(r) = \frac{1}{B(r)}, \qquad B(r) = 1 - \frac{2GM}{r} - \frac{\Lambda_{\text{eff}}r^2}{3} \tag{551}$$

We can now compare this solution to the standard Schwarzschild solution. The fact that the time–time component of the metric tensor is inversely proportional to the r–r component is still true but the radial dependence of the function $B(r)$ is modified by the presence of a cosmological constant. In the Newtonian limit, we have $g_{tt} \simeq 1 + 2V(r)$, where $V(r)$ is the Newtonian potential. In the present context, this implies that

$$V(r) \simeq -\frac{GM}{r} - \frac{\Lambda_{\text{eff}}r^2}{6} \tag{552}$$

Therefore, we obtain the usual gravitational potential corrected by a term proportional to the cosmological constant. This term looks like an inverted harmonic oscillator. This corresponds to a force felt by a test body of mass m given by

$$\boldsymbol{F} = \frac{GMm}{r^2}\boldsymbol{u} - \frac{m\Lambda_{\text{eff}}r}{3}\boldsymbol{u} \tag{553}$$

where $\boldsymbol{u}$ is a vector directed towards the central body of mass M. The cosmological constant force is directed in the other direction and can therefore be viewed as a kind of "anti-gravity force". We also notice that it is proportional to r which means that its effect will be more important on large scales. This last property is the clue to understand why measuring the cosmological constant is more efficient on large scales (the best example being of course cosmology). This means that the modification of the planet orbits and/or the energy levels of the atoms will certainly be a small effect. On the other hand, we know that measurements in, say, atomic physics are extremely accurate and, a priori, there is the hope that this accuracy could compensate the smallness of the effect. Unfortunately, as we are now going to study, this will not be the case.

11.2. Planet orbits

Our goal in this section is to study how the planet orbits are modified by the presence of vacuum energy [103–105]. We will follow the treatment of Ref. [103]. We assume that the gravitational field created by the Sun and the cosmological constant is given by Eqs. (541) and (551) with $M = M_\odot$. The motion of a planet is represented by the motion of a test body in this background gravitational field. The corresponding geodesic equation can be written as

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda} \frac{dx^\nu}{d\tau} \frac{dx^\lambda}{d\tau} = 0 \tag{554}$$

where τ is an affine parameter (not necessarily the proper time). Then, after having evaluated the Christoffel symbols for the metric (541), one can express these equations explicitly. One obtains

$$\frac{d^2 t}{d\tau^2} + \frac{B'}{B} \frac{dt}{d\tau} \frac{dr}{d\tau} = 0 \tag{555}$$

$$\frac{d^2 r}{d\tau^2} + \frac{B'}{2A} \left(\frac{dt}{d\tau}\right)^2 + \frac{A'}{2A} \left(\frac{dr}{d\tau}\right)^2 - \frac{r}{A} \left(\frac{d\theta}{d\tau}\right)^2 - \frac{r \sin^2\theta}{A} \left(\frac{d\varphi}{d\tau}\right)^2 = 0 \tag{556}$$

$$\frac{d^2\theta}{d\tau^2} + \frac{2}{r} \frac{dr}{d\tau} \frac{d\theta}{d\tau} - \cos\theta \sin\theta \left(\frac{d\varphi}{d\tau}\right)^2 = 0 \tag{557}$$

$$\frac{d^2\varphi}{d\tau^2} + \frac{2}{r} \frac{dr}{d\tau} \frac{d\varphi}{d\tau} + 2 \frac{\cos\theta}{\sin\theta} \frac{d\theta}{d\tau} \frac{d\varphi}{d\tau} = 0 \tag{558}$$

Let us now solve these equations. Firstly, we can assume that the orbit stays in the $\theta = \pi/2$ plane. This automatically solves Eq. (557). Then, if we divide Eq. (555) by $dt/d\tau$ and Eq. (558) by $d\varphi/d\tau$, we obtain two new equations, namely

$$\frac{d}{d\tau} \left[\ln\left(\frac{dt}{d\tau}\right) + \ln B\right] = 0 \tag{559}$$

$$\frac{d}{d\tau} \left[\ln\left(\frac{d\varphi}{d\tau}\right) + \ln r^2\right] = 0 \tag{560}$$

The first equation can be used to define the affine parameter. We choose it such that it satisfies $dt/d\tau = 1/B$. Then, the second equation tells us that

$$r^2 \frac{d\varphi}{d\tau} = J \tag{561}$$

where J is a constant. Finally, Eq. (556) can be re-written as

$$\frac{d^2 r}{d\tau^2} + \frac{A'}{2A} \left(\frac{dr}{d\tau}\right)^2 - \frac{J^2}{Ar^3} + \frac{B'}{2AB^2} = 0 \tag{562}$$

If we multiply this last equation by $2A \, dr/d\tau$, one arrives at

$$\frac{d}{d\tau} \left[A \left(\frac{dr}{d\tau}\right)^2 + \frac{J^2}{r^2} - \frac{1}{B}\right] = 0 \tag{563}$$

This means that the quantity inside the square brackets is a constant. Let us call this constant $-E$. Then, this gives an equation expressing how the radial coordinate r varies with the angle φ,
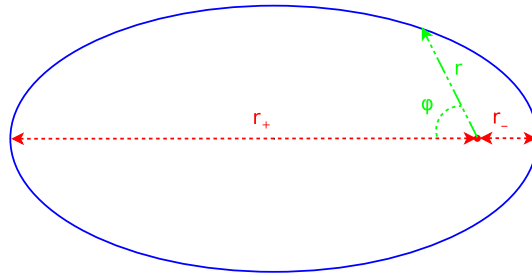
$$\left(\frac{dr}{d\varphi}\right)^2 = \frac{r^4}{A(r)} \left[\frac{1}{J^2 B(r)} - \frac{1}{r^2} - \frac{E}{J^2}\right] \tag{564}$$

It is easy to express the constants E and J in terms of the aphelion and perihelion r_+ and r_-, which are the two points on the orbit such that $dr/d\varphi = 0$, see Fig. 7 (where the elements of the orbits are represented). This gives

$$\frac{1}{J^2} = \frac{r_+^{-2} - r_-^{-2}}{B_+^{-1} - B_-^{-1}} \tag{565}$$

$$E = \frac{r_+^2/B_+ - r_-^2/B_-}{r_+^2 - r_-^2} \tag{566}$$

with $B_\pm$ defined by $B_\pm \equiv B(r_\pm)$.

Fig. 7. Sketch of the elliptic planet orbit with the aphelion and perihelion. The orbit is planar and the position of the planet can be characterized by the coordinates (r, φ).

So far, we have been very general and we have never specified the function $B(r)$. From now on, we are going to use the explicit form of this function, see Eq. (551). Let us notice in particular that, in absence of a cosmological constant and in the limit where $GM/r \ll 1$, the two above expressions reduce to

$$\frac{E-1}{2} \simeq \frac{GM}{r_+ + r_-}, \quad J^2 \simeq 2GMD \tag{567}$$

where $D \equiv r_+ r_- / (r_+ + r_-)$ (of course, not to be confused with the integration constant introduced in the previous subsection). As is shown in the following footnote, these are exactly the relations obtained in the Newtonian case. More generally, Eq. (564) takes the form

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\varphi}\right)^2 = \frac{1-E}{J^2}r^4 + \frac{2GME}{J^2}r^3 - r^2 + 2GMr + \frac{\Lambda_{\mathrm{eff}}r^4}{3} + \frac{E\Lambda_{\mathrm{eff}}r^6}{3J^2} \tag{568}$$

Performing the standard change of variable, $u \equiv 1/r$ and differentiating with respect to φ, we obtain[4]

---

[4] It is interesting to compare this result with the Newtonian analysis. In this case, the Lagrangian is given by

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) + \frac{GMm}{r}$$

where a dot denotes a derivative with respect to time. Again we have chosen the orbit to be in the plane $\theta = \pi/2$. This leads to the following equations of motion

$$m\ddot{r} - mr\dot{\varphi}^2 + \frac{GMm}{r^2} = 0$$

$$mr^2\ddot{\varphi} + 2mr\dot{r}\dot{\varphi} = 0$$

The second equation gives $\dot{\varphi} = J/r^2$ while the first one can be re-written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{r}^2}{2} + \frac{J^2}{2r^2} - \frac{GM}{r}\right) = 0$$

The quantity between the parenthesis is a constant and we call it $-E$. Then, this equation can be re-expressed as

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\varphi}\right)^2 = -\frac{2E}{J^2}r^4 + \frac{2GM}{J^2}r^3 - r^2$$

From this equation one can deduce an expression for E and J in terms of the aphelion and the perihelion. One gets

$$E = \frac{GM}{r_+ + r_-}, \quad J^2 = 2GMD$$

Introducing the variable $u = 1/r$ and differentiating once more with respect to φ, one arrives at

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\varphi^2} + u = \frac{GM}{J^2} \tag{569}$$

This equation can easily be integrated. The solution reads

$$u(\varphi) = F\cos(\varphi - \varphi_0) + \frac{GM}{J^2} = \frac{GM}{J^2}(1 - e\cos\varphi)$$

where F and φ_0 are two arbitrary constants. Clearly, one can always take $\varphi_0 = 0$. The above trajectory represents, as is well known, an ellipse with eccentricity e.

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\varphi^2} + u = \frac{E}{J^2} GM + 3GMu^2 - \frac{E}{J^2}\frac{\Lambda_{\mathrm{eff}}}{3u^3} \tag{570}$$

This is the differential equation that needs to be solved in order to find the trajectory of the planet. It should be compared with its Newtonian counterpart, see Eq. (569). One can check that, in this regime (without a cosmological constant), the two equations are indeed identical since according to Eq. (567), $E \to 1$, $\Lambda_{\mathrm{eff}} = 0$ and the term $3GMu^2$ can be neglected. The relativistic term $3GMu^2$ is responsible for the precession of the perihelion and the last term represents the correction due to the cosmological constant.

We now turn to the solution of Eq. (570). Let us first review how one calculates the precession of the perihelion without a cosmological constant [102]. The idea is to restart from Eq. (568). For $\Lambda_{\mathrm{eff}} = 0$, the polynomial on the right-hand side can be written as

$$\frac{1-E}{J^2} r(r - r_+)(r - r_-)(r - \epsilon) \tag{571}$$

since we know it has to vanish at $r = r_\pm$. The only quantity that remains to be determined in this expression is ϵ. Comparing the above relation with Eq. (568), we deduce that

$$\frac{1-E}{J^2} r_+ r_- \epsilon = 2GM \tag{572}$$

Using the expressions derived before, it is easy to show that $\epsilon = 2GM/(1 - 2GM/D)$. As a consequence, one can write

$$\varphi_+ - \varphi_- = \frac{J}{\sqrt{1-E}} \int_{r_-}^{r_+} \frac{1}{\sqrt{(r - r_+)(r - r_-)}} \frac{1}{\sqrt{1 - \epsilon/r}} \frac{\mathrm{d}r}{r} \tag{573}$$

There are several methods to evaluate this integral. As a matter of fact, it can be integrated exactly in terms of Elliptic functions [48,49]. But, being given that ϵ is a small quantity, it is better to expand the result in this parameter. This leads to easier calculations that are, moreover, more explicit. In addition, the method can be generalized to the case of a non-vanishing cosmological constant, which is of course our main goal. In fact, if we restart from Eq. (564) and use that $B^{-1} \simeq 1 + 2GM/r + \cdots$, we see that the term in the squared bracket in Eq. (564) is a quadratic function of $1/r$ which vanishes at $r = r_\pm$. This means that one necessarily has

$$\frac{1}{J^2 B(r)} - \frac{E}{J^2} - \frac{1}{r^2} \simeq C \left(\frac{1}{r} - \frac{1}{r_-} \right) \left(\frac{1}{r} - \frac{1}{r_+} \right) \tag{574}$$

Then, writing Eq. (564) in terms of u rather than r, one arrives at

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\varphi} \right)^2 \simeq \frac{C}{A(r)} (u - u_-)(u - u_+) \tag{575}$$

Now, when $\Lambda_{\mathrm{eff}} \neq 0$, as is clear from Eq. (568), the term in the square bracket of Eq. (564) is no longer a quadratic function of $1/r$ even if the expansion of the function $B^{-1}(r)$ is used. However, this term still vanishes at $r = r_\pm$ and the integral will be dominated by the contributions coming from the vicinity of these points, see also Eq. (573). The idea is then to continue to work with Eq. (574) and to take into account the cosmological constant in the definition of C [103]. This should provide a reasonable approximation of $\varphi_+ - \varphi_-$. Using Eqs. (565) and (566), one can write

$$\frac{1}{J^2 B(r)} - \frac{E}{J^2} - \frac{1}{r^2} = \frac{r_-^2 [B^{-1}(r) - B_-^{-1}] - r_+^2 [B^{-1}(r) - B_+^{-1}]}{r_+^2 r_-^2 (B_+^{-1} - B_-^{-1})} - \frac{1}{r^2}$$

$$\simeq C \left(\frac{1}{r} - \frac{1}{r_-} \right) \left(\frac{1}{r} - \frac{1}{r_+} \right) \tag{576}$$

Then, if we differentiate both sides twice with respect to u, this gives

$$C = \frac{(u_+ - u_-)(u_+ - u_-)A''(u)}{2(u_+ - u_-)A'(u)} - 1 \tag{577}$$

$$\simeq -1 + \frac{u A''}{A'} \bigg|_{u=L^{-1}} \tag{578}$$

where $L \equiv 2D$ is the semilatus rectum. In the above expression, we have used $A_+ - A_- \simeq (u_+ - u_-)A'(L^{-1})$. So far, the method is completely general provided that $A^{-1} = B$. We now use the explicit form of A^{-1}. Expanding everything in GMu and in Λ_{eff}, one obtains [103]

$$C \simeq -1 + \frac{4GM}{L} + \frac{\Lambda_{\text{eff}}L^3}{GM} + \frac{4\Lambda_{\text{eff}}L^3}{3} + \cdots \tag{579}$$

The final step consists in using the above expression in Eq. (575). One finds [103]

$$\varphi_+ - \varphi_- \simeq \int\limits_{u_-}^{u_+} \frac{A^{1/2}(u)\,\mathrm{d}u}{[C(u-u_+)(u-u_-)]^{1/2}} = \pi\left(1 + \frac{3}{2}\frac{2GM}{L} + \frac{\Lambda_{\text{eff}}L^3}{2GM} + \cdots\right) \tag{580}$$

where we have expanded $A^{1/2}C^{-1/2}$ and where we have used the fact that $\int_{u_-}^{u_+} \mathrm{d}u/[(u-u_+)(u-u_-)]^{1/2} = \ln(u_+ - u_-) - \ln(u_+ - u_-) + 2i\pi/2$, the complex i being exactly cancel by the i coming from the fact that C, which appears in a square root, is negative [its expansion starts with -1, see Eq. (579)]. As a consequence, we obtain

$$\Delta\varphi \equiv 2(\varphi_+ - \varphi_-) - 2\pi \simeq 6\pi\frac{GM}{L} + \Delta\varphi_\Lambda \tag{581}$$

where the first term is the standard relativistic perihelion precession while [103]

$$\Delta\varphi_\Lambda \simeq \frac{\pi\Lambda_{\text{eff}}L^3}{GM} \tag{582}$$

is the contribution due to the cosmological constant. The above equation is the main result of this subsection. It gives the perihelion precession due to the presence of vacuum energy. Let us notice that the scaling $\Delta\varphi_\Lambda \propto L^3$ indicates that the precession is larger for planets that are far from the Sun. It is of course totally consistent with the fact that the effect of the cosmological constant is more important on large scales.

We now use the above result to constrain the amount of vacuum energy present in our Universe. For this purpose, let us now calculate this perihelion shift for, say, Mercury (as mentioned before, the effect is larger for other planets but this does not modify the conclusions obtained below; estimates for Venus, Earth and Mars can be found in Ref. [105]). In this case, we have $r_+ \simeq 68.8 \times 10^6$ km and $r_- \simeq 46 \times 10^6$ km. This implies that $D \simeq 27.7 \times 10^6$ km. We also have $GM \simeq 2.95$ km. Finally, it is interesting to calculate in terms of the vacuum energy density observed in cosmology, see Eq. (540). For the cosmological constant, this leads to the following number $\Lambda_{\text{obs}} \simeq 8.81h^2 \times 10^{-48}$ km^{-2}. This leads to the following expression:

$$\Delta\varphi_\Lambda \simeq 1.6h^2 \times 10^{-24}\left(\frac{\Lambda}{\Lambda_{\text{obs}}}\right) \text{ rad/revolution} \tag{583}$$

or, since Mercury completes 415 revolutions each century and there are $360 \times 60 \times 60/(2\pi)$ arc-seconds per radian,

$$\Delta\varphi_\Lambda \simeq 2.7h^2 \times 10^{-16}\left(\frac{\Lambda}{\Lambda_{\text{obs}}}\right) \text{ arc-second/century} \tag{584}$$

which is, of course, completely unobservable. The reason for such a result has already been mentioned. The solar system has a characteristic scale (say a few astronomical units) which is too small for the cosmological constant to leave a sizable imprint. However, even if one cannot measure the vacuum energy density with this method, it can nevertheless be used to constrain its value. Indeed, the uncertainty in the precession of the perihelion of Mercury is about $0.1''$ par century, which means that

$$\Delta\varphi_\Lambda \simeq 2.7h^2 \times 10^{-16}\left(\frac{\rho_\Lambda}{\rho_{\text{cri}}}\right) < 0.1 \tag{585}$$

implying that

$$\rho_{\text{vac}} \lesssim 10^{15}\rho_{\text{cri}} \simeq 3 \times 10^{-32} \text{ GeV}^4 \tag{586}$$

Of course, this limit is not competitive with what one has obtained in cosmology. However, when compared to Eq. (508), it leads to an interesting piece of information which allows us to define what the cosmological constant problem is more accurately. Indeed, we see that there is a contradiction between the theoretical expectation (508) and the observations of the planet trajectories in our solar system. In other words, the predictions of Eq. (508), if true, would lead to orbits drastically different from what we observe. Therefore, even without the cosmological observations and only from the motion of the planets in the solar system, one sees that our calculation of vacuum energy must be incorrect. Put it differently, it would be incorrect to say that the cosmological constant problem originates only from the recent astrophysical observations. In some sense, the problem is much worse. Only by observing the planets in the solar system, a class of observations which seems to be much more straightforward than the cosmological observations, we know that our regularization method of the vacuum energy must be flawed.

11.3. Modifications of atomic levels

We have just seen that observing the motion of planets in our solar system can tell us something about the value of the vacuum energy. A priori, another possibility to observe the influence of the cosmological constant is through the modification of the atomic energy levels that its presence would cause. In this section, we investigate this question in the case of the Hydrogen atom since this is technically simple [106–111]. We will follow the treatment of Ref. [106]. In particular, since the problem is still spherically symmetric, one can use the metric tensor (541) obtained before.

The motion of an electron around the nucleus, living in a static and spherically symmetric manifold, is described by the following Dirac equation

$$(\Gamma^\mu D_\mu + m)\Psi = 0 \tag{587}$$

where Ψ is a Dirac spinor, see Section 7.1 and m the mass of the electron. In the above equation, Γ_μ are the Dirac matrices satisfying the relation

$$\{\Gamma^\mu, \Gamma^\nu\} = 2g^{\mu\nu} \tag{588}$$

Let us notice that we have already presented this formula, see Eq. (330). Here, we slightly change the notation and now denote the Dirac matrices in curved space–time by Γ^μ (notice that we have also included a factor i in the definition of Γ^μ). The matrices Γ^μ are related to the Dirac matrices in flat space–time, γ^μ, by $\Gamma_\mu = e^\alpha{}_\mu \gamma_\alpha$, where $e^\alpha{}_\mu$ are the vierbein fields such that $g_{\mu\nu} = e^\alpha{}_\mu e^\beta{}_\nu \eta_{\alpha\beta}$. The symbol D_μ denotes the covariant derivative acting on the spinor, namely

$$D_\mu \Psi = (\partial_\mu - \omega_\mu)\Psi \tag{589}$$

where the spinorial connection can be expressed as

$$\omega_\mu \equiv -\frac{1}{4}\gamma_\alpha\gamma_\beta e^\alpha{}_\nu g^{\nu\lambda}(\nabla_\mu e^\beta{}_\lambda) + iqA_\mu \tag{590}$$

with $\nabla_\mu e^\beta{}_\lambda \equiv \partial_\mu e^\beta{}_\lambda - \Gamma^\sigma_{\mu\lambda} e^\beta{}_\sigma$ and A_μ is the vector field describing the electromagnetic field acting on the electron (q is the charge). The Dirac equation can also be re-written as

$$i\partial_0 \Psi = H\Psi \tag{591}$$

where the effective Hamiltonian is given by

$$H = -i(g^{00})^{-1}\Gamma^0\Gamma^i(\partial_i\Psi - \omega_i\Psi) + i\Gamma_0\Psi - i(g^{00})^{-1}\Gamma^0 m\Psi \tag{592}$$

In this case, the Dirac equation looks like an effective Schrödinger equation.

It is clear that the typical dimension of an atom is very small in comparison to the typical curvature scale of the manifold. Therefore, as was done for the calculation of the Green function in Section 8, one can perform a local analysis. For this purpose, we now use the Fermi normal coordinates [78]. In these coordinates, the metric tensor takes the form

$$g_{00} = -1 - R_{0\ell 0m}x^\ell x^m + \cdots \tag{593}$$

$$g_{0i} = -\frac{2}{3}R_{0\ell im}x^\ell x^m + \cdots \tag{594}$$

$$g_{ij} = \delta_{ij} - \frac{1}{3}R_{i\ell jm}x^\ell x^m + \cdots \tag{595}$$

where the dots indicate higher order terms. This implies that the vierbein fields can be expressed as

$$e^\alpha{}_0 = \delta^\alpha{}_0 - \frac{1}{2}R^\alpha{}_{\ell 0m}x^\ell x^m \tag{596}$$

$$e^\alpha{}_i = \delta^\alpha{}_i - \frac{1}{2}R^\alpha{}_{\ell im}x^\ell x^m \tag{597}$$

and the Christoffel symbols can be written as

$$\Gamma^0_{ij} = \frac{1}{3}(R_{0ijm} + R_{0jim})x^m \tag{598}$$

$$\Gamma^0_{0i} = R_{0iom}x^m \tag{599}$$

$$\Gamma^i_{jk} = \frac{1}{3}(R_{jikm} + R_{kijm})x^m, \qquad \Gamma^0_{00} = 0 \tag{600}$$

$$\Gamma^i_{0j} = R_{0mji}x^m, \qquad \Gamma^i_{00} = R_{0i0m}x^m \tag{601}$$

Using the above results, this finally leads to the following expression for the spinorial connection [106]

$$\omega_0 = \frac{1}{2}\gamma_0\gamma_j R^j{}_{00m}x^m + \frac{1}{4}\gamma_k\gamma_j R^{kj}{}_{0m}x^m + iqA_0 \tag{602}$$

$$\omega_i = \frac{1}{4}\gamma_0\gamma_j R^{0j}{}_{im}x^m + \frac{1}{8}\gamma_k\gamma_j R^{kj}{}_{im}x^m + iqA_i \tag{603}$$

The above formulas will allow us to calculate the effective Hamiltonian (592) explicitly. But before reaching this stage, we must also determine vector field A_μ. The Lagrangian of the corresponding gauge field is given by, see also Eq. (364)

$$S = -\int d^4x\sqrt{-g}\left(\frac{1}{4}g^{\alpha\beta}g^{\mu\nu}F_{\mu\alpha}F_{\nu\beta} - j^\mu A_\mu\right) \tag{604}$$

from which we deduce that the equations of motion read

$$g^{\mu\nu}\nabla_\mu\nabla_\nu A_\gamma - R^\sigma{}_\gamma A_\sigma = -j_\gamma \tag{605}$$

At first order in the curvature, these equations can be expressed as [106]

$$\delta^{ij}\partial_i\partial_j A_0 + \frac{1}{3}R_{i\ell jm}x^\ell x^m\partial^i\partial^j A_0 + \frac{5}{3}R_{i00m}x^m\partial^i A_0 + 2R^k{}_{i0m}x^m\partial^i A_k - \frac{2}{3}R_{im}x^m\partial^i A_0 = -j_0 \tag{606}$$

$$\delta^{ij}\partial_i\partial_j A_k + \frac{1}{3}R_{i\ell jm}x^\ell x^m\partial^i\partial^j A_k + \frac{2}{3}\delta^{ij}\left(R^\alpha{}_{kim} + R^\alpha{}_{ikm}\right)x^m\partial_j A_\alpha - \frac{2}{3}R^\sigma{}_k A_\sigma - \frac{1}{3}R^p{}_{00k}A_p$$
$$- \frac{1}{3}R_{i00m}x^m\partial^i A_k - \frac{2}{3}R_{pm}x^m\partial^p A_k = j_k \tag{607}$$

For a one electron atom with a nucleus of charge Ze, one can take

$$j^0 = -Ze\delta(\mathbf{r}), \quad j^k = 0 \tag{608}$$

In what follows, we determine the vector perturbatively in the curvature. At zeroth order, in flat space–time, the equation for the time component reduces to

$$\delta^{ij}\partial_i\partial_j A_0 = Ze\,\delta(\mathbf{r}) \tag{609}$$

whose solution is, of course,

$$A_0^{(0)} = -\frac{Ze}{4\pi}\frac{1}{r} \tag{610}$$

Then, we write $A_0 = A_0^{(0)} + \delta A_0$, where δA_0 is first order in the curvature. By inserting this expression into the full equation (606) and using the zeroth order solution (610), one arrives at [106]

$$\delta^{ij}\partial_i\partial_j\delta A_0 + \frac{1}{3}\frac{Ze}{4\pi r^3}\left(3R^0{}_{\ell 0m} - R_{\ell m}\right)x^\ell x^m = 0 \tag{611}$$

The solution of the above equation reads [106]

$$\delta A_0 = \frac{1}{12}\frac{Ze}{4\pi}(R + 4R_{00})r + \frac{1}{12}\frac{Ze}{4\pi}\left(3R^0{}_{\ell 0m} - R_{\ell m}\right)\frac{x^\ell x^m}{r} \tag{612}$$

In the same manner, one can determine the space component. Obviously, at zeroth order, one has $A_k^{(0)} = 0$. The first order perturbation δA_k obeys the equation [106]

$$\delta^{ij}\partial_i\partial_j\delta A_k + \frac{2}{3}\frac{Ze}{4\pi}R^0{}_k\frac{1}{r} + \frac{2}{3}\frac{Ze}{4\pi}R^0{}_{ikm}\frac{x^i x^m}{r^3} = 0 \tag{613}$$

whose solution can be written as [106]

$$\delta A_k = \frac{1}{2}\frac{Ze}{4\pi}R_{0k}r + \frac{1}{6}\frac{Ze}{4\pi}R^0{}_{lkm}\frac{x^\ell x^m}{r} \tag{614}$$

This completes our calculation of the electromagnetic field.

Having determined the vierbeins fields (i.e. the quantities describing the gravitational force) and the gauge field (i.e. the quantity describing the electromagnetic force), we are now in a position where we can calculate the Hamiltonian (592). In order to make contact with the usual treatment of the Hydrogen atom, it is more convenient to introduce the standard Dirac matrices α_i and β related to the flat space–time Dirac matrices by

$$\gamma_0 = i\beta, \quad \gamma_i = -i\beta\alpha_i \tag{615}$$

As is well known, they satisfy

$$\alpha_i\alpha_j + \alpha_j\alpha_i = 2\delta_{ij} \tag{616}$$

$$\alpha_i\beta + \beta\alpha_i = 0, \quad \beta^2 = 1 \tag{617}$$

Then, we use the expressions of the spinorial connections (602), of the gauge field (610), (614) and of the Dirac matrices in Eq. (592) of the Hamiltonian. This leads to the following expression:

$$\begin{aligned}
H = {} &-i\alpha^i\partial_i + m\beta - \frac{\zeta}{r} - \frac{i}{2}R_{0\ell 0m}x^\ell x^m\alpha_i\partial^i - \frac{i}{6}R_{i\ell jm}x^\ell x^m\alpha^j\partial^i - \frac{i}{6}R_{0\ell jm}x^\ell x^m\alpha^j\alpha^i\partial_i - \frac{i}{2}R_{0\ell im}x^\ell x^m\partial^i \\
&+ \frac{i}{8}\alpha^i\alpha^j\alpha^k R_{jkim}x^m + \frac{i}{4}\alpha^i\alpha^j R_{0jmi}x^m + \frac{i}{4}\alpha^i\alpha^j R_{0mij}x^m - \frac{i}{2}\alpha^j R_{0j0m}x^m + \frac{1}{6}\zeta R_{0\ell mi}x^\ell x^m\frac{\alpha^i}{r} + \frac{1}{2}\zeta R_{0i}\alpha^i r \\
&+ \frac{1}{12}\zeta(R + 4R_{00})r - \frac{1}{12}\zeta(R_{\ell m} + 3R_{0\ell 0m})\frac{x^\ell x^m}{r} - \frac{m}{6}R_{i\ell 0m}x^\ell x^m\beta\alpha^i + \frac{m}{2}R_{0\ell 0m}x^\ell x^m\beta
\end{aligned} \tag{618}$$

In this expression we have defined $\zeta \equiv Ze^2/(4\pi)$ and we have taken $q = -e$. The three first terms are the usual Dirac Hamiltonian. The other terms represent the correction due to space–time curvature. Since the above expression is quite complicated, it is interesting to estimate the order of magnitude of each term. The curvature is given by D^{-2} where D is a typical distance. Each x^i is typically of the order of the Bohr radius, i.e. $x^i \simeq a_0 \simeq m^{-1}e^{-2} \simeq m^{-1}\zeta^{-1}$. As a consequence, ∂_i is of order $m\zeta$. This also means that the typical velocity is given by $m\zeta/m \simeq \zeta$. Of course the matrices α^i and β are dimensionless but, in the non-relativistic limit $\alpha^i\partial_i$ must reduce to $\simeq mv^2$ which implies that $\alpha^i \simeq v \simeq \zeta$. In the following we consider that $\zeta = Z\alpha$ where α is the fine structure constant is a small number. Then, the dominant term in the above expression is the last one, which means that the Hamiltonian can be approximated as

$$H = H_0 + H_{\text{int}} \simeq \boldsymbol{\alpha} \cdot \boldsymbol{p} + m\beta - \frac{\zeta}{r} + \frac{1}{2}mR_{0\ell 0m}x^\ell x^m\beta \tag{619}$$

where we have used $p_i = -i\partial_i$. If we use the following representation for the Dirac matrices

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix} \tag{620}$$

and write the Dirac spinor as

$$\Psi = e^{iEx^0}\begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{621}$$

then the Dirac equation reduces to two equations that can be expressed as

$$\boldsymbol{\alpha} \cdot \boldsymbol{p}\chi + \left(m - \frac{\zeta}{r} + \frac{1}{2}mR_{0\ell 0m}x^\ell x^m\right)\phi = E\phi \tag{622}$$

$$\boldsymbol{\alpha} \cdot \boldsymbol{p}\phi - \left(m + \frac{\zeta}{r} + \frac{1}{2}mR_{0\ell 0m}x^\ell x^m\right)\chi = E\chi \tag{623}$$

From the second relation we can express χ in terms of $\boldsymbol{\alpha} \cdot \boldsymbol{p}\phi$ and then we can insert this expression into the first equation. This leads to

$$\left(E + m + \frac{\zeta}{r} + \frac{1}{2}mR_{0\ell 0m}x^\ell x^m\right)^{-1}(\boldsymbol{\alpha} \cdot \boldsymbol{p})^2\phi + \left(m - \frac{\zeta}{r} + \frac{1}{2}mR_{0\ell 0m}x^\ell x^m\right)\phi = E\phi \tag{624}$$

or, using $\boldsymbol{\alpha} \cdot \boldsymbol{p}\phi = p^2\phi$ and expanding the denominator of the first term in the above equation

$$\left[1 - \frac{1}{E+m}\frac{\zeta}{r} - \frac{m}{2(E+m)}R_{0\ell 0m}x^\ell x^m\right]\frac{p^2}{E+m}\phi - \left(\frac{\zeta}{r} - \frac{1}{2}mR_{0\ell 0m}x^\ell x^m\right)\phi = (E - m)\phi \tag{625}$$

But, if we now perform an additional expansion assuming non-relativistic velocities for the electron (which, in practice, is a good approximation), one arrives at the following expression for the term $p^2/(E+m)$

$$\frac{p^2}{E+m} = \frac{p}{m+m+p^2/(2m)+\cdots} \simeq \frac{p^2}{2m} + \cdots \tag{626}$$

As a consequence, we conclude that the non-relativistic limit is given by

$$H_{\text{NR}}\phi = (E - m)\phi \tag{627}$$

where the non-relativistic Hamiltonian can be written as [106]

$$H_{\text{NR}} = H_0 + H_{\text{int}} = \frac{p^2}{2m} - \frac{\zeta}{r} + \frac{1}{2}mR_{0\ell0m}x^{\ell}x^{m}\beta \tag{628}$$

This is our final result. Endowed with this equation, one can calculate, by means of perturbations theory, the modification of the ground state due to the presence of a cosmological constant. This leads to (the arguments of E indicates the value of the quantum numbers n and ℓ)

$$\Delta E(1,0) = \frac{1}{2}mR_{0\ell0m}\langle 1s|x^{\ell}x^{m}|1s\rangle \tag{629}$$

$$= \frac{1}{2}mR_{0\ell0m}\frac{1}{3}\delta^{\ell m}\langle 1s|r^2|1s\rangle \tag{630}$$

$$= \frac{1}{6}mR_{00}\langle 1s|r^2|1s\rangle \tag{631}$$

The ground state wave function is given by (at this order of perturbations theory, there is no need to take into account the modification on the wave-function; one can use the standard one)

$$|1s\rangle = \frac{1}{\sqrt{\pi a_0^3}}e^{-r/a_0} \tag{632}$$

where we recall that a_0 is the Bohr radius. This implies that

$$\langle 1s|r^2|1s\rangle = \frac{4\pi}{\pi a_0^3}\int_0^{\infty} r^4 e^{-r/a_0}\,\mathrm{d}r = \frac{4\pi}{\pi a_0^3}\frac{3a_0^5}{4} = 3a_0^2 \tag{633}$$

The vacuum Einstein equations are, see Eq. (542), $R_{\mu\nu} = \Lambda_{\text{eff}}g_{\mu\nu}$ and, therefore, at first order in the cosmological constant one $R_{00} = -\Lambda_{\text{eff}}$ (here, of course, we only compute the corrections due to the cosmological constant and not those originating from the Earth gravitational field). This means that

$$\Delta E(1,0) = -\frac{1}{2}mc^2\Lambda_{\text{eff}}a_0^2 \tag{634}$$

where we have re-established the speed of light for convenience. In fact, this result could have been anticipated from the Newtonian approach. Indeed, we have seen that the Newtonian potential, in presence of a cosmological constant, is given by $-\Lambda_{\text{eff}}mc^2r^2/6$, see Eqs. (552) and (553). Using the perturbation theory with this potential exactly leads Eq. (634). Let us now evaluate the correction. Using that the reduced mass of the atom is essentially the electron mass, $m \simeq 9.1 \times 10^{-31}$ kg and that the Bohr radius is $a_0 = 0.52 \times 10^{-10}$ m, one obtains

$$\Delta E(1,0) = -6 \times 10^{-69}h^2\left(\frac{\Lambda}{\Lambda_{\text{obs}}}\right) \text{eV} \tag{635}$$

a quantity completely unobservable, probably forever. For comparison the Lamb shift is of the order 10^{-6} eV. At this stage this does not come as a surprise. As discussed in the case of planet orbits, the size of the system is so small that the influence of the cosmological constant is negligible. However, requiring that the shift due to the cosmological constant be smaller than the Lamb shift leads to

$$\rho_{\text{vac}} \lesssim 1.6 \times 10^{62}\rho_{\text{cri}} \simeq 1.3 \times 10^{16} \text{ GeV}^4 \tag{636}$$

This result is less good than the one obtained from planet motion by about 50 orders of magnitude. Moreover, the theoretical expectation (508) satisfies this inequality. Let us notice in passing that the result based on the wrong regularization scheme, $\rho_{\text{vac}} \simeq M_{\text{Pl}}^4$ does not satisfy the previous bound.

We now have a clearer view of what the cosmological constant problem is. On one hand, one can calculate the vacuum energy density by means of, apparently, well-known techniques of regularization. These techniques usually lead very robust results, fully consistent with the observations. For instance, the radiative corrections to the calculation of cross-sections have proven in very good agreement with various high energy physics experiments. One the other hand, various observational results (and not only in cosmology) indicate that the previous calculation is not correct. The theoretical expectation value for the vacuum energy density is so large that it should have already be seen by many experiences. Moreover, in cosmology, under a certain number of hypothesis, one even measures ρ_{vac} which is found to be many orders of magnitude smaller than the number derived from the theory. Since, as mentioned before, the theoretical framework seems to be robust (let us recall again that it involves calculations that can be found in the first pages of any textbook on quantum field theory) and since the observations seem to be convincing, we are facing a genuine mystery. However, before accepting this conclusion, it is worth checking that no loophole is present in the previous reasoning. This will be the goal of the next sections. A first question that can be asked is whether the zero-point fluctuations that are at the origin of the problem are a real physical phenomenon or just an artifact of the formalism of quantum field theory. We turn to this issue in the next section.

12. Do vacuum fluctuations really exist?

We have seen in the previous sections that the zero-point fluctuations manifest themselves as the vacuum energy. Since this leads to conclusions that seem to be in contradiction with observations, it is legitimate to ask whether these fluctuations really exist in Nature. In particular, we want to investigate whether the zero-point fluctuations could manifest themselves in other physical phenomena. The answer to this question is usually positive and one experiment which is considered as a proof that the vacuum fluctuations are real is the measurement of the Lamb shift [18,112]. In the next subsection, we discuss this phenomenon.

12.1. The Lamb shift

Let us consider the "motion" of an electron in an atom. Its "location" is described by the vector $\mathbf{r}$. The vector $\mathbf{r}$ is going to fluctuate because of the interaction between the electron and the zero-point fluctuations of the electromagnetic field. This interaction will slightly modify the position of the atomic levels leading to the Lamb shift. Let us study in details how this effect can be derived. Let $V(\mathbf{r})$ be the Coulomb potential which determines the properties of the atom. At the point $\mathbf{r} + \delta\mathbf{r}$, the potential can be Taylor expanded according to

$$V(\mathbf{r} + \delta\mathbf{r}) = V(\mathbf{r}) + \frac{\partial V}{\partial x_i}\delta x_i + \frac{1}{2}\frac{\partial^2 V}{\partial x_i \partial x_j}\delta x_i \delta x_j + \cdots \tag{637}$$

Then, if we time average the previous expression, one obtains

$$\langle V(\mathbf{r} + \delta\mathbf{r})\rangle \simeq V(\mathbf{r}) + \frac{\partial V}{\partial x_i}\langle \delta x_i \rangle + \frac{1}{2}\frac{\partial^2 V}{\partial x_i \partial x_j}\langle \delta x_i \delta x_j \rangle + \cdots \tag{638}$$

and, because the problem is spherically symmetric, we have $\langle \delta x_i \rangle = 0$ and $\langle \delta x_i \delta x_j \rangle = \langle \delta r^2 \rangle \delta_{ij}/3$. The first term in Eq. (638) is the unperturbed Coulomb potential and the last one (since the second one vanishes) can be viewed as a small perturbation of the Hamiltonian, namely

$$\Delta H = \frac{1}{6}\frac{\partial^2 V}{\partial x_i \partial x_j}\langle \delta r^2 \rangle \delta_{ij} \tag{639}$$

In an atom with atomic number Z, the Coulomb potential is given by

$$V = -\frac{Z\alpha}{r} \tag{640}$$

where $\alpha \simeq 1/137$ is the fine structure constant. This means that $\delta_{ij}\partial^2 V/(\partial x_i \partial x_j) = 4\pi Z\alpha\delta(\mathbf{r})$.

Let now us assume that the atom is placed in the quantum state $|\ell, m\rangle$. As we have already discussed in Section 11.3, at first order, the time-independent perturbations theory tells us that the correction to the energy of the level (ℓ, m) is given by [57]

$$\Delta E(n, \ell) = \langle \ell, m|\Delta H|\ell, m\rangle \tag{641}$$

Using Eq. (639), one obtains the following expression for the energy displacement

$$\Delta E(n, \ell) = \frac{2\pi}{3}Z\alpha\left|\Psi_{n\ell}(0)\right|^2\langle \delta r^2 \rangle \tag{642}$$

where $\Psi_{n\ell}$ is the wave-function in the corresponding quantum state. Since the wave-function at the origin can be expressed as [57]

$$\left|\Psi_{n\ell}(0)\right| = \frac{1}{\pi^{1/2}}\left(\frac{mZ\alpha}{n}\right)^{3/2}\delta_{\ell 0} \tag{643}$$

where m is the electron mass ($m \simeq 0.511$ MeV) one finally arrives at [18]

$$\Delta E(n, \ell) = \frac{(2mZ\alpha)^3}{12}\frac{Z\alpha}{n^3}\langle \delta r^2 \rangle\delta_{\ell 0} \tag{644}$$

Therefore, we have reduced the problem to the calculation of the quantity $\langle \delta r^2 \rangle$. In order to evaluate this quantity, we must now evaluate the displacement of the electron. The electron "moves" under the influence of the electric field within the atom and, as a consequence, δr_i obeys the following equation

$$m\frac{d^2}{dt^2}\delta r_i = eE_i \tag{645}$$

where $E_\mu \equiv u^\nu F_{\mu\nu} = -\dot{A}_i$ is the electric field (or, rather, its vacuum fluctuating component). Using the usual expansion of the vector potential in terms of creation and annihilation operators, see Eq. (387), one can write

$$E_i(t, \boldsymbol{x}) = \frac{-1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega(k)}} \sum_{\alpha=1}^{2} \epsilon_i^\alpha(\boldsymbol{k}) \left[-i\omega a_{\boldsymbol{k}}^\alpha \mathrm{e}^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}} + i\omega \left(a_{\boldsymbol{k}}^\alpha\right)^\dagger \mathrm{e}^{i\omega t - i\boldsymbol{k}\cdot\boldsymbol{x}} \right] \tag{646}$$

In the same manner, one has to Fourier expand δr_i which is considered as an operator in this context. One obtains

$$\delta r_i(t, \boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}\boldsymbol{k} \sum_{\alpha=1}^{2} \epsilon_i^\alpha(\boldsymbol{k}) \left[\delta r(\boldsymbol{k}, t) a_{\boldsymbol{k}}^\alpha \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{x}} + \delta r^*(\boldsymbol{k}, t) \left(a_{\boldsymbol{k}}^\alpha\right)^\dagger \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \right] \tag{647}$$

The Fourier amplitude of the displacement operator can be calculated from the equation of motion (645). This leads to the following expression

$$\delta r(\boldsymbol{k}, t) = \frac{e}{m} \frac{1}{\sqrt{2\omega^3}} \mathrm{e}^{-i\omega t} \tag{648}$$

From this expression, one can now evaluate the quantity $\langle \delta r^2 \rangle$. One finds

$$\langle \delta r^2 \rangle = \langle 0 | \delta^{ij} \delta r_i \delta r_j | 0 \rangle = \frac{1}{(2\pi)^3} \int \mathrm{d}\boldsymbol{k} \sum_{\alpha=1}^{2} \delta^{ij} \epsilon_i^\alpha \epsilon_j^\alpha \delta r(\boldsymbol{k}, t) \delta r^*(\boldsymbol{k}, t) \tag{649}$$

$$= \frac{2\alpha}{m^2 \pi} \int_0^\infty \frac{\mathrm{d}\omega}{\omega} \tag{650}$$

where we have used $\alpha = e^2/(4\pi)$. Therefore, inserting the above formula into Eq. (644), one obtains the following equation for the energy shift [18]

$$\Delta E(n, \ell) = \frac{4mZ^4\alpha^5}{3\pi n^3} \int_0^\infty \frac{\mathrm{d}\omega}{\omega} \delta_{\ell 0} \tag{651}$$

This equation is the main result of this section. It gives the energy levels displacement due to the presence of the electromagnetic zero-point fluctuations.

As usual this expression is divergent and needs to be regularized. In a simplified treatment, one simply assumes, see Ref. [18], that the wavelengths in the above sum must be larger than the Compton wavelength of the electron, which implies $\omega < \omega_{\max} \simeq m$ and that $\omega > \omega_{\min} \simeq 1/a_0 \simeq 1/[1/(\alpha m)] \simeq m\alpha$, where a_0 is the Bohr radius. In this case, the previous results reduces to

$$\Delta E(n, \ell) \simeq \frac{4mZ^4\alpha^5}{3\pi n^3} \ln\left(\frac{1}{\alpha}\right) \delta_{\ell 0} \tag{652}$$

Let us now evaluate this quantity for the Hydrogen atom. Clearly, there is no effect for $n = 1$. But for $n = 2$ and $\ell = 0$, one has

$$\Delta E(2, 0) \simeq \frac{4m\alpha^5}{3\pi 8} \ln\left(\frac{1}{\alpha}\right) \delta_{\ell 0} \simeq 668 \text{ MHz} \tag{653}$$

where we have used $m \simeq 0.511 \text{ MeV} = h\nu$ with $h = 4.135 \times 10^{-15} \text{ eV} \times \text{s}$.

This shift has been experimentally observed and is usually taken as a proof that the zero-point fluctuations are real since they lead to an observed physical phenomenon. Based on this result, it now seems difficult to argue that the zero-point fluctuations are just an artifact of the formalism of quantum field theory. Moreover, there exists another experiment, the Casimir effect, where the effects of the vacuum fluctuations can be observed. In the next section, we investigate this case.

12.2. The Casimir effect

In this section, we study the Casimir effect [113–119]. Since its discovery, this effect has been studied in great details and it is clear that, here, one cannot review all the literature. We will give a description of this phenomenon in a simplified context only and will refer to Refs. [113–115] for readers interested in learning more about this interesting subject.
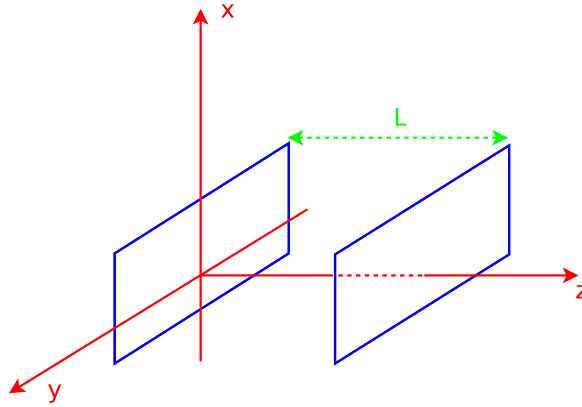
Fig. 8. Sketch of a Casimir cavity of width L.

12.2.1. The Casimir force

Let us consider an experimental situation where we have two conducting plates in the (x, y) plane separated by a distance L long the z direction, see Fig. 8. For simplicity, instead of considering the electromagnetic field in the cavity (which is the case in the real world, especially when measurements are performed), we will just treat the case of a real massive scalar field $\Phi(t, \boldsymbol{x})$. This field obeys the Klein–Gordon equation already studied and solved before, see Eq. (59), except that now the boundary conditions are modified by the presence of the two plates. Let us write $\Phi(t, \boldsymbol{x}) = X(x)Y(y)Z(z)T(t)$. Then, inserting this anzatz into the Klein–Gordon equation, one obtains

$$-\frac{\ddot{T}}{T} + \frac{1}{X}\frac{d^2 X}{dx^2} + \frac{1}{Y}\frac{d^2 Y}{dy^2} + \frac{1}{Z}\frac{d^2 Z}{dz^2} - m^2 = 0 \tag{654}$$

where m is the mass of the scalar particle. Clearly, one can separate the variables and, as a consequence, the above equation can be split into four differential equations, namely

$$-\frac{\ddot{T}}{T} - m^2 = C^2 \tag{655}$$

$$\frac{1}{X}\frac{d^2 X}{dx^2} = k_x^2 \tag{656}$$

$$\frac{1}{Y}\frac{d^2 Y}{dy^2} = k_y^2 \tag{657}$$

$$\frac{1}{Z}\frac{d^2 Z}{dz^2} = k_z^2 \tag{658}$$

where the constants C, k_x, k_y and k_z are related by

$$C - k_x^2 - k_y^2 - k_z^2 = 0 \tag{659}$$

The solution in the x and y direction can be expressed in terms of standard plane waves, $X(x) = A_x e^{ik_x x} + B_x e^{-ik_x x}$ and $Y(y) = A_y e^{ik_y y} + B_y e^{-ik_y y}$. The quantities $A_{x,y}$ and $B_{x,y}$ are just integration constants. Of particular interest is of course the solution along the z-axis. It can also be written as

$$Z(z) = A_z e^{ik_z z} + B_z e^{-ik_z z} \tag{660}$$

where A_z and B_z are two integration constants. But the difference with the case of a free field is that the boundary conditions "feel" the presence of the two plates. These boundary conditions are given by $Z(0) = 0$ and $Z(L) = 0$. The first one implies that $A_z = -B_z$ while the second one gives

$$2i \sin(k_z L) = 0 \tag{661}$$

or, equivalently,

$$k_z = n\frac{\pi}{L} \tag{662}$$

where n is an integer. Finally, the equation for the function $T(t)$ can also be easily solved and the solution reads

$$T(t) = A_t e^{i\omega t} + B_t e^{-i\omega t} \tag{663}$$

where $\omega^2 = C^2 + m^2$. Therefore, upon using Eq. (659), the frequency ω takes the form

$$\omega = \sqrt{k_x^2 + k_y^2 + n^2 \frac{\pi^2}{L^2} + m^2} \tag{664}$$

In this expression, k_x and k_y take continuous values.

Based on the previous considerations, one can now write the field operator as an expansion in terms of creation and annihilation operators. But this expansion encodes the fact that, along the z-axis, the wave-number is discrete. Concretely, we have

$$\Phi(t, \boldsymbol{x}) = \int \frac{dk_x}{(2\pi)^{1/2}} \int \frac{dk_y}{(2\pi)^{1/2}} \sum_{n=1}^{\infty} \sqrt{\frac{2}{L}} \frac{1}{\sqrt{2\omega}} \sin\left(n\frac{\pi}{L}z\right) \left(c_{k_x,k_y,n} e^{-i\omega t + ik_x x + ik_y y} + c_{k_x,k_y,n}^\dagger e^{i\omega t - ik_x x - ik_y y}\right) \tag{665}$$

The operators $c_{k_x,k_y,n}$ and $c_{k_x,k_y,n}^\dagger$ obey the usual commutation relations for bosons. It is interesting to check that the above expansion is indeed consistent with the commutation relations between the field and its conjugate momentum. For this purpose, let us compute the following quantity:

$$\int dx \int dy \int_0^L dz\, \Phi e^{-ip_x x - ip_y y} \sin\left(m\frac{\pi}{L}z\right)$$

$$= \int \frac{dk_x}{(2\pi)^{1/2}} \int \frac{dk_y}{(2\pi)^{1/2}} \sum_{n=1}^{\infty} \int dx \int dy \sqrt{\frac{2}{L}} \frac{1}{\sqrt{2\omega}} \left(c_{k_x,k_y,n} e^{-i\omega t + ik_x x + ik_y y}\right.$$

$$\left. + c_{k_x,k_y,n}^\dagger e^{i\omega t - ik_x x - ik_y y}\right) e^{-ip_x x - p_y y} \int_0^L dz \sin\left(m\frac{\pi}{L}z\right) \sin\left(n\frac{\pi}{L}z\right) \tag{666}$$

Using the fact that $\int_0^L dz \sin(n\pi z/L)\sin(m\pi z/L) = (L/2)\delta_{mn}$ (as can be checked directly) and $\int dx\, e^{i(k_x - p_x)x} = (2\pi)\delta(k_x - p_x)$ one arrives at

$$\int dx \int dy \int_0^L dz\, \Phi(t, \boldsymbol{x}) e^{-ip_x x - ip_y y} \sin\left(m\frac{\pi}{L}z\right) = \frac{2\pi}{2}\sqrt{\frac{L}{\omega}}\left(c_{p_x,p_y,m} e^{-i\omega t} + c_{-p_x,-p_y,m}^\dagger e^{i\omega t}\right) \tag{667}$$

In the same manner, one can also evaluate the following integral involving the conjugate momentum $\Pi(t, \boldsymbol{x}) = \dot{\Phi}(t, \boldsymbol{x})$ (a dot means a derivative with respect to time):

$$\int dx \int dy \int_0^L dz\, \Pi(t, \boldsymbol{x}) e^{-ip_x x - ip_y y} \sin\left(m\frac{\pi}{L}z\right) = \frac{2\pi}{2} i\omega \sqrt{\frac{L}{\omega}}\left(-c_{p_x,p_y,m} e^{-i\omega t} + c_{-p_x,-p_y,m}^\dagger e^{i\omega t}\right) \tag{668}$$

From the two above expression, one can deduce a formula expressing the creation and the annihilation operators in terms of the field operator and its conjugate momentum. One obtains

$$c_{p_x,p_y,m} = \sqrt{\frac{\omega}{L}}\frac{e^{i\omega t}}{2\pi}\left[\int dx\, dy \int_0^L dz\, \Phi\, e^{-ip_x x - ip_y y} \sin\left(m\frac{\pi}{L}z\right) + \frac{i}{\omega}\int dx\, dy \int_0^L dz\, \Pi\, e^{-ip_x x - ip_y y} \sin\left(m\frac{\pi}{L}z\right)\right]$$

$$c_{p_x,p_y,m}^\dagger = \sqrt{\frac{\omega}{L}}\frac{e^{-i\omega t}}{2\pi}\left[\int dx\, dy \int_0^L dz\, \Phi\, e^{ip_x x + ip_y y} \sin\left(m\frac{\pi}{L}z\right) - \frac{i}{\omega}\int dx\, dy \int_0^L dz\, \Pi\, e^{ip_x x + ip_y y} \sin\left(m\frac{\pi}{L}z\right)\right]$$

We are now in a position where we can compute the commutator of a creation and an annihilation operators. It can be written as

$$\left[c_{p_x,p_y,n}, c_{q_x,q_y,m}^\dagger\right] = \frac{e^{i(\omega_p - \omega_q)t}}{(2\pi)^2}\frac{\sqrt{\omega_p \omega_q}}{L}\int dx\, dy \int_0^L dz \int d\bar{x}\, d\bar{y} \int_0^L d\bar{z}\left[\Phi(t, \boldsymbol{x}) e^{-ip_x x - ip_y y} \sin\left(n\frac{\pi}{L}z\right)\right.$$

$$\left. + \frac{i}{\omega_p}\Pi(t, \boldsymbol{x}) e^{-ip_x x - ip_y y} \sin\left(n\frac{\pi}{L}z\right), \Phi(t, \bar{\boldsymbol{x}}) e^{iq_x \bar{x} + iq_y \bar{y}} \sin\left(m\frac{\pi}{L}\bar{z}\right)\right.$$

$$- \frac{i}{\omega_q} \Pi(t, \bar{\boldsymbol{x}}) \, e^{iq_x \bar{x} + iq_y \bar{y}} \sin\left(m \frac{\pi}{L} \bar{z} \right) \Bigg] \tag{669}$$

Then, we make use of the canonical relation $[\Phi(t, \boldsymbol{x}), \Pi(t, \boldsymbol{y})] = i\delta^3(\boldsymbol{x} - \boldsymbol{y})$ in the previous formula. This leads to the following complicated expression:

$$
[c_{p_x, p_y, n}, c^\dagger_{q_x, q_y, m}] = \frac{e^{i(\omega_p - \omega_q)t}}{(2\pi)^2} \frac{\sqrt{\omega_p \omega_q}}{L} \int dx\, dy \int_0^L dz \int d\bar{x}\, d\bar{y} \int_0^L d\bar{z} \Bigg\{ -\frac{i}{\omega_q} e^{-ip_x x - ip_y y + iq_x \bar{x} + iq_y \bar{y}} \sin\left(n \frac{\pi}{L} z \right)
$$
$$
\times \sin\left(m \frac{\pi}{L} \bar{z} \right) [\Phi(t, \boldsymbol{x}), \Pi(t, \bar{\boldsymbol{x}})]
$$
$$
+ \frac{i}{\omega_p} e^{-ip_x x - ip_y y + iq_x \bar{x} + iq_y \bar{y}} \sin\left(n \frac{\pi}{L} z \right) \sin\left(m \frac{\pi}{L} \bar{z} \right) [\Pi(t, \boldsymbol{x}), \Phi(t, \bar{\boldsymbol{x}})] \Bigg\} \tag{670}
$$
$$
= \frac{e^{i(\omega_p - \omega_q)t}}{(2\pi)^2} \frac{\sqrt{\omega_p \omega_q}}{L} \left(\frac{1}{\omega_q} + \frac{1}{\omega_p} \right) \int dx\, dy \int_0^L dz\, e^{-i(p_x - q_x)x - i(p_y - q_y)y} \sin\left(n \frac{\pi}{L} z \right) \sin\left(m \frac{\pi}{L} z \right) \tag{671}
$$
$$
= \frac{\omega}{L} \frac{2}{\omega} \delta(p_x - q_x) \delta(p_y - q_y) \frac{L}{2} \delta_{mn} = \delta(p_x - q_x) \delta(p_y - q_y) \delta_{mn} \tag{671}
$$

We have thus reached our goal, namely we have shown that the canonical commutation relation between the field and its conjugate momentum implies the usual commutation relation between the creation and the annihilation operator.

We are now in a position where we can evaluate the energy density of the field in the Casimir cavity. It is of course given by the same expression as in the free case, see Eq. (66). This means that one should first calculate the three terms that participate in the expression of T_{00}, see Eqs. (63), (64) and (65). Let us start with $\langle 0|\dot{\Phi}^2|0\rangle$. Using the Fourier expansion of the field and the commutation relations established above, one obtains

$$\langle 0|\dot{\Phi}^2|0\rangle = \int \frac{dk_x}{2\pi} \int \frac{dk_y}{2\pi} \sum_{n=1}^\infty \frac{2}{L} \frac{\omega^2}{2\omega} \sin^2\left(n \frac{\pi}{L} z \right) \tag{672}$$

and, if we integrate over the cavity volume

$$\int_{\text{Cavity}} d\boldsymbol{x} \, \langle 0|\dot{\Phi}^2|0\rangle = \frac{D^2}{2} \int \frac{dk_x}{2\pi} \int \frac{dk_y}{2\pi} \sum_{n=1}^\infty \omega \tag{673}$$

where D^2 is the area of the plates (of course, the two parallel plates have the same surface). In the same manner, one finds

$$\int_{\text{Cavity}} d\boldsymbol{x} \langle 0|\delta^{ij} \partial_i \Phi \partial_j \Phi|0\rangle = \frac{D^2}{2} \int \frac{dk_x}{2\pi} \int \frac{dk_y}{2\pi} \sum_{n=1}^\infty \frac{1}{2\omega} \left(k_x^2 + k_y^2 + n^2 \frac{\pi^2}{L^2} \right) \tag{674}$$

Finally, the last relevant term reads

$$\int_{\text{Cavity}} d\boldsymbol{x} \, \langle 0|\Phi^2|0\rangle = \frac{D^2}{2} \int \frac{dk_x}{2\pi} \int \frac{dk_y}{2\pi} \sum_{n=1}^\infty \frac{1}{2\omega} \tag{675}$$

Putting everything together and using the definition (66), the expression of the energy of the field within the empty cavity takes the form

$$\int_{\text{Cavity}} d\boldsymbol{x} \, \langle 0|T_{00}|0\rangle \tag{676}$$

$$= \left(\frac{D}{2\pi} \right)^2 \int dk_x \int dk_y \sum_{n=1}^\infty \frac{\omega}{2} \tag{677}$$

$$= \left(\frac{D}{2\pi} \right)^2 \int dk_x \int dk_y \sum_{n=1}^\infty \frac{1}{2} \sqrt{k_x^2 + k_y^2 + n^2 \frac{\pi^2}{L^2} + m^2} \tag{678}$$

Of course, this expression has exactly the expected form. Roughly speaking it can be written as "$\sum \frac{1}{2}\omega$", see Eq. (664), the only difference being that, along the z-axis, the sum takes into account the fact that the momentum is discrete. Another

common point with our previous considerations is that the above number is actually divergent and must be regularized. In the following, we assume $m = 0$ for simplicity and use dimensional regularization. Therefore, the expression of the energy can now be written as

$$E = \left(\frac{D}{2\pi}\right)^d \int d^d \boldsymbol{k}_\perp \sum_{n=1}^{\infty} \frac{1}{2} \sqrt{k_\perp^2 + n^2 \frac{\pi^2}{L^2}} \tag{679}$$

$$= \left(\frac{D}{2\pi}\right)^d \int dk_\perp k_\perp^{d-1} d^{d-1}\Omega \sum_{n=1}^{\infty} \frac{1}{2} \sqrt{k_\perp^2 + n^2 \frac{\pi^2}{L^2}} = \left(\frac{D}{2\pi}\right)^d \frac{2\pi^{d/2}}{\Gamma(d/2)} \int dk_\perp k_\perp^{d-1} \sum_{n=1}^{\infty} \frac{1}{2} \sqrt{k_\perp^2 + n^2 \frac{\pi^2}{L^2}} \tag{680}$$

where d is the "dimension" of the plates (i.e. $d = 2$ in the real world) and $k_\perp$ denotes the wave-vector living in the sub-manifold defined by the plates. Then using the change of variables $y = Lk_\perp/(n\pi)$ and the following definition [48,49] of the Euler's integral of first kind:

$$B(x, y) = 2 \int_0^{\infty} \frac{t^{2x-1}}{(1+t^2)^{x+y}} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{681}$$

the energy can be re-expressed as

$$E = \left(\frac{D}{2\pi}\right)^d \frac{\pi^{d/2}}{2} \left(\frac{\pi}{L}\right)^{d+1} \frac{\Gamma(-d/2 - 1/2)}{\Gamma(-1/2)} \sum_{n=1}^{\infty} n^{d+1} \tag{682}$$

The above formula can be expressed in term of the Riemann zeta function defined by the following expression:

$$\zeta(s) \equiv \sum_{n=1}^{\infty} n^{-s} \tag{683}$$

Moreover, upon using the two following equations:

$$\zeta(-1 - d) = 2^{1-d} \pi^{-2-d} \sin\left[\frac{\pi}{2}(-1 - d)\right] \Gamma(2 + d) \zeta(2 + d) \tag{684}$$

$$\Gamma\left(-\frac{d}{2} - \frac{1}{2}\right) \sin\left[\frac{\pi}{2}(-1 - d)\right] = \frac{\pi}{\Gamma(d/2 + 3/2)} \tag{685}$$

one can re-write the energy as

$$\frac{E}{D^d} = -\frac{1}{L^{d+1}} \frac{\Gamma(1 + d/2)}{2^{d+2} \pi^{d/2+1}} \zeta(2 + d) \tag{686}$$

The Casimir force arises because of a shift in the vacuum energy when the plates are present compared to the situation where they are absent. So the relevant quantity is in fact $\Delta E/D^d \equiv (E - E_0)/D^d$ where E_0/D^d is the above quantity in the limit $L \to +\infty$. As a consequence, for $d = 2$, one arrives at

$$\frac{\Delta E}{D^2} = -\frac{\pi^2}{1440L^3} \tag{687}$$

where we have used $\zeta(4) = \pi^4/90$. Therefore, the corresponding force is given by [113–116]

$$F = -\frac{\partial}{\partial L}\left(\frac{\Delta E}{D^2}\right) = -\frac{\pi^2}{480L^4} \tag{688}$$

which is one half the result obtained in the case of the electromagnetic field since we have only one state of polarization instead of two. We notice that the force is attractive (this could change in the case of more complicated geometries).

It is also interesting to check that the final result is independent from the regularization scheme used to obtain it. In order to test this idea in the simplest situation (i.e. where the calculations are easy), let us consider again the energy of a Casimir cavity but, this time, with $d = 0$. In this case, our zeta-function regularized result reads

$$\frac{E}{D^0} = -\frac{1}{L} \frac{\Gamma(1)\zeta(2)}{4\pi} = -\frac{\pi}{24L} \tag{689}$$

since $\Gamma(1) = 1$ and $\zeta(2) = \pi^2/6$. One the other hand, we can return to the expression of E in Eq. (682). In the case $d = 0$, the energy density $\rho(L) = E/L$ reads

$$\rho(L) = \frac{\pi}{2L^2} \sum_{n=1}^{\infty} n \tag{690}$$

In order to regularize this divergent quantity, we now proceed differently and introduce an exponential cut-off such that (previously, we argued that introducing a cut-off was not a good method because this breaks Lorentz invariance. Here, of course, this objection is no longer valid as the presence of the two plates along the z-axis obviously breaks translation invariance; as a consequence working with a cut-off seems legitimate in the present context)

$$\rho(L, \alpha) = \frac{\pi}{2L^2} \sum_{n=1}^{\infty} n e^{-n\alpha/L} \tag{691}$$

The sum is easy to perform and one obtains

$$\rho(L, \alpha) = \frac{\pi}{8L^2} \sinh^{-2}\left(\frac{\alpha}{2L}\right) \tag{692}$$

$$\simeq \frac{\pi}{2\alpha^2} - \frac{\pi}{24L^2} + \cdots \tag{693}$$

Now, the regularized Casimir energy is the shift energy in the vacuum after having sent the cut-off to zero; in other words

$$E = L \lim_{\alpha \to 0}\left[\rho(L, \alpha) - \lim_{L \to \infty} \rho(L, \alpha)\right] \tag{694}$$

$$= -\frac{\pi}{24L} \tag{695}$$

that is to say exactly the same expression obtained with another regularization scheme. It is reassuring to notice that the final result is independent of the regularization method used.

12.2.2. The Casimir stress–energy tensor

In this section, we aim at calculating the stress–energy tensor of a scalar field in a Casimir cavity [113–116,119]. Our goal is of course to compare this tensor with the vacuum stress–energy tensor $T_{\mu\nu} = -\rho_{\mathrm{vac}} g_{\mu\nu}$. Here, we use a method based on the calculations of the Green function. This is yet another method of calculation of the Casimir effect, different from the two ones exposed before and, therefore, it will be interesting to compare this approach with the treatments used previously.

Let us start with the calculation of the Green function. It obeys the following equation

$$-\eta^{\mu\nu} \partial_\mu \partial_\nu G(x, \bar{x}) = \delta^{(4)}(x^\mu - \bar{x}^\mu) \tag{696}$$

It is interesting to compare the above formula with Eq. (474). It is clear that this is the same equation but with $g_{\mu\nu} = \eta_{\mu\nu}$ (flat space–time) and $m = 0$ (mass-less field). In order to solve this equation, it is convenient to Fourier transform $G(x, \bar{x})$ in frequency and transverse momentum. Explicitly, one writes

$$G(x, \bar{x}) = \int \frac{\mathrm{d}\omega}{2\pi} \frac{\mathrm{d}\boldsymbol{k}_\perp}{(2\pi)^2} e^{-i\omega(t-\bar{t})+ik_x(x-\bar{x})+ik_y(y-\bar{y})}$$
$$\times g(z, \bar{z}; \omega, k_x, k_y) \tag{697}$$

Inserting this expression into the Green equation (696), one obtains

$$-\left(\frac{\mathrm{d}^2}{\mathrm{d}z^2} + \lambda^2\right) g(z, \bar{z}) = \delta(z - \bar{z}) \tag{698}$$

where we have defined

$$\lambda^2 \equiv \omega^2 - k_\perp^2 = \omega^2 - k_x^2 - k_y^2 \tag{699}$$

Our goal is now to determine explicitly the Green function. Let us start with the solution within the plates, i.e. $0 < z, \bar{z} < L$. When $z \neq \bar{z}$, the solution can be written

$$g(z, \bar{z}) = A(\bar{z}) \cos(\lambda z) + B(\bar{z}) \sin(\lambda z) \tag{700}$$

where $A(\bar{z})$ and $B(\bar{z})$ are two unknown function of $\bar{z}$. We must now take into account the boundary conditions. If $z < \bar{z}$, the condition $g(z = 0, \bar{z}) = 0$ implies that $A = 0$ and therefore $g = B \sin(\lambda z)$ (for $z < \bar{z}$). On the other hand, if $z > \bar{z}$, the condition $g(z = L, \bar{z}) = 0$ means that

$$g = A \cos(\lambda z) - A \frac{\cos(\lambda L)}{\sin(\lambda L)} \sin(\lambda z), \quad z > \bar{z} \tag{701}$$

We have two unknown quantities and, therefore, we need two equations. One is provided by writing the continuity of the Green function when $z = \bar{z}$. The other is obtained by integrating Eq. (698) from $\bar{z} - \epsilon$ to $\bar{z} + \epsilon$. This gives

$$-\int_{\bar{z}-\epsilon}^{\bar{z}+\epsilon} \frac{\mathrm{d}^2 g}{\mathrm{d}z^2}\,\mathrm{d}z - \lambda^2 \int_{\bar{z}-\epsilon}^{\bar{z}+\epsilon} g\,\mathrm{d}z = \int_{\bar{z}-\epsilon}^{\bar{z}+\epsilon} \delta(z-\bar{z})\,\mathrm{d}z \tag{702}$$

which leads to

$$-\left[\frac{\mathrm{d}g}{\mathrm{d}z}(\bar{z}+\epsilon) - \frac{\mathrm{d}g}{\mathrm{d}z}(\bar{z}-\epsilon)\right] = 1 \tag{703}$$

Therefore, taking the limit $\epsilon \to 0$, the solutions of these two equations (i.e. the one expressing the continuity of the Green function and the one we have just derived) read

$$A(\bar{z}) = \frac{1}{\lambda}\sin(\lambda\bar{z}) \tag{704}$$

$$B(\bar{z}) = \frac{1}{\lambda}\cos(\lambda\bar{z}) - \frac{\cos(\lambda L)}{\sin(\lambda L)}A(\bar{z}) \tag{705}$$

As a consequence, the Green function between the plates can be expressed as [116]

$$g(z,\bar{z}) = \begin{cases} \dfrac{\sin(\lambda z)\sin(\lambda L - \lambda\bar{z})}{\lambda \sin(\lambda L)}, & z < \bar{z} \\[2mm] \dfrac{\sin(\lambda\bar{z})\sin(\lambda L - \lambda z)}{\lambda \sin(\lambda L)}, & z > \bar{z} \end{cases} \tag{706}$$

The same method can be used to determine the Green function outside the cavity. Let us apply it for z and $\bar{z}$ larger than L, i.e. on the right-hand side of the cavity. In this case, it is more convenient to re-write the general solution (700) as

$$g(z,\bar{z}) = C(\bar{z})e^{i\lambda z} + D(\bar{z})e^{-i\lambda z} \tag{707}$$

When $z > \bar{z}$, we require only one branch to be present which amounts to take $D = 0$. For $z < \bar{z}$, the boundary condition is still $g(L,\bar{z}) = 0$. Then, straightforward manipulations lead to [116]

$$g(z,\bar{z}) = \begin{cases} \frac{1}{\lambda}e^{i\lambda(\bar{z}-L)}\sin(\lambda z - \lambda L), & z < \bar{z} \\[2mm] \frac{1}{\lambda}e^{i\lambda(z-L)}\sin(\lambda\bar{z} - \lambda L), & z > \bar{z} \end{cases} \tag{708}$$

which completes our determination of the Green function. The calculation of the Green function outside the cavity, but on the left-hand side, proceeds exactly in the same way. The corresponding expression is obtained from the one above with $L = 0$, reflecting the fact that the boundary condition is now $g(0,\bar{z}) = 0$.

Endowed with the Green function of the problem, one can now turn to our main goal, namely the calculation of the stress–energy tensor. We have already seen that, see also Eq. (100),

$$G(\boldsymbol{r},\bar{\boldsymbol{r}}) = i\langle 0|T[\Phi(\boldsymbol{r})\Phi(\bar{\boldsymbol{r}})]|0\rangle \tag{709}$$

But we know that the T_{00} component of the stress–energy tensor can be expressed as, see Eq. (7),

$$T_{00} = \frac{1}{2}\left[(\partial_0\Phi)^2 + (\partial_x\Phi)^2 + (\partial_y\Phi)^2 + (\partial_z\Phi)^2\right] \tag{710}$$

Therefore, if we apply the following operator to the Green function in Eq. (709)

$$\frac{1}{2}[\partial_0\partial_{\bar{0}} + \partial_x\partial_{\bar{x}} + \partial_y\partial_{\bar{y}} + \partial_z\partial_{\bar{z}}] \tag{711}$$

where $\partial_{\bar{\mu}}$ means a derivative with respect to $\bar{x}^\mu$, to $G(\boldsymbol{r},\bar{\boldsymbol{r}})$ and, then, take the space–time points to be the same $\boldsymbol{r} = \bar{\boldsymbol{r}}$, one should obtain an expression for $\langle T_{00}\rangle$. Using Eqs. (697) and (699), the result reads

$$\langle T_{00}\rangle = \frac{1}{2i}\int \frac{\mathrm{d}\omega}{2\pi}\frac{\mathrm{d}k_x\,\mathrm{d}k_y}{(2\pi)^2}(\omega^2 + k_x^2 + k_y^2 + \partial_z\partial_{\bar{z}})g(z,\bar{z})\Big|_{z=\bar{z}} \tag{712}$$

Then, we use the expression of the Green function inside the plates (706) and one arrives at

$$\langle T_{00}\rangle = \frac{1}{2i}\int \frac{\mathrm{d}\omega}{2\pi}\frac{\mathrm{d}k_x\,\mathrm{d}k_y}{(2\pi)^2}\frac{1}{\lambda\sin(\lambda L)}\left[(\omega^2 + k_x^2 + k_y^2)\sin(\lambda z)\sin(\lambda L - \lambda L) - \lambda^2\cos(\lambda z)\cos(\lambda L - \lambda z)\right] \tag{713}$$

$$= -\frac{1}{2i}\int \frac{\mathrm{d}\omega}{2\pi}\frac{\mathrm{d}k_x\,\mathrm{d}k_y}{(2\pi)^2}\frac{1}{\lambda\sin(\lambda L)}\left[\omega^2\cos(\lambda L) - k^2\cos(2\lambda z - \lambda L)\right] \tag{714}$$

In order to evaluate this integral, we perform the two Wick rotations $\omega \to i\zeta$ and $\lambda \to i\kappa$. This leads to

$$\langle T_{00} \rangle = -\frac{1}{2} \int_0^\infty \frac{d\zeta}{2\pi} \int \frac{dk_x\, dk_y}{(2\pi)^2} \frac{1}{\kappa \sinh(\kappa L)} \big[\zeta^2 \cosh(\kappa L) + k^2 \cosh(2\kappa z - \kappa L) \big] \tag{715}$$

Then, we introduce polar coordinates in the plane (ζ, k), i.e. $\zeta = \kappa \cos\theta$ and $k = \kappa \sin\theta$. This gives

$$\langle T_{00} \rangle = -\frac{1}{4\pi^2} \int_0^\infty d\kappa\, \kappa \int_0^{\pi/2} d\theta\, \kappa^2 \frac{\sin\theta}{\sinh(\kappa L)} \big[\cos^2\theta \cosh(\kappa L) + \sin^2\theta \cosh(2\kappa z - \kappa L) \big] \tag{716}$$

$$= -\frac{1}{12\pi^2} \int_0^\infty d\kappa\, \kappa^3 \left[\frac{\cosh(\kappa L)}{\sinh(\kappa L)} + 2\frac{\cosh(2\kappa z - \kappa L)}{\sinh(\kappa L)} \right] \tag{717}$$

Finally, writing that

$$\frac{\cosh(\kappa L)}{\sinh(\kappa L)} = 1 + \frac{2}{e^{2\kappa L} - 1} \tag{718}$$

the expression of $\langle T_{00} \rangle$ can be re-expressed as

$$\langle T_{00} \rangle = -\frac{1}{6\pi^2} \int_0^\infty d\kappa\, \kappa^3 \left[\frac{1}{e^{2\kappa L} - 1} + \frac{1}{2} + \frac{e^{2\kappa z} + e^{2\kappa(L-z)}}{e^{2\kappa L} - 1} \right] \tag{719}$$

The last step of the calculation consists in evaluating the last term [we call it $g(z)$ in what follows; not to be confused with the Green function] of the above equation. One has

$$g(z) = -\frac{1}{6\pi^2} \frac{1}{16L^4} \int_0^\infty dy\, y^3 \frac{e^{y/L} + e^{y(1-z/L)}}{e^y - 1} \tag{720}$$

$$= -\frac{1}{6\pi^2} \frac{1}{16L^4} \sum_{n=1}^\infty \int_0^\infty dy\, y^3 \big[e^{y(z/L-n)} + e^{y(1-n-z/L)} \big] \tag{721}$$

where we have expanded the denominator in a geometric series. The two above integrals can be expressed in terms of the Hurwitz zeta function defined by [48,49]

$$\zeta(s, a) \equiv \sum_{m=0}^\infty \frac{1}{(m+a)^s} \tag{722}$$

and this leads to

$$g(z) = -\frac{1}{16\pi^2 L^4} \left[\zeta\left(4, \frac{z}{L}\right) + \zeta\left(4, 1 - \frac{z}{L}\right) \right] \tag{723}$$

If we use the fact that $\int_0^{+\infty} \kappa^3/(e^{2\kappa L} - 1)\, d\kappa = \pi^4/(240L^4)$, one arrives at our final expression for the time–time component of the stress–energy tensor, namely

$$\langle T_{00} \rangle = -\frac{\pi^2}{1440 L^4} - \frac{1}{6\pi^2} \int_0^\infty d\kappa \frac{\kappa^3}{2} + g(z) \tag{724}$$

Of course we notice that the first term in the above equation is similar to Eq. (687).

Let us now evaluate $\langle T_{00} \rangle$ outside the cavity. The calculation proceeds along the same lines but we now need to use the Green function given by Eq. (708). Straightforward manipulations lead to

$$\langle T_{00} \rangle = -\frac{1}{6\pi^2} \int_0^\infty d\kappa \frac{\kappa^3}{2} - \frac{1}{16\pi^2} \frac{1}{(z-L)^4} \tag{725}$$

We obtain the same structure as inside the cavity, namely a divergent term and a surface divergence term, i.e. a term which is divergent only on the plate. As explained in Ref. [116], this type of terms can be removed by restoring conformal invariance.

Let us now calculate the quantity $\langle T_{zz} \rangle$ inside the cavity. Clearly, following the previous considerations, it can be written as

$$\langle T_{zz} \rangle = \frac{1}{2i} \int \frac{d\omega}{2\pi} \frac{dk_x \, dk_y}{(2\pi)^2} (\partial_t \partial_{\bar{t}} - \partial_x \partial_{\bar{x}} - \partial_y \partial_{\bar{y}} + \partial_z \partial_{\bar{z}}) g(z, \bar{z}) \Big|_{z=\bar{z}}. \tag{726}$$

Following exactly the same procedure as before, we arrive at

$$\langle T_{zz} \rangle = -3 \frac{\pi^2}{1440 L^4} - \frac{3}{6\pi^2} \int\limits_0^\infty d\kappa \, \frac{\kappa^3}{2} \tag{727}$$

Let us now calculate $\langle T_{zz} \rangle$ outside the plates. As we did for the time component, we must now use the Green function (708). This gives

$$\langle T_{zz} \rangle = \frac{1}{2i} \int \frac{d\omega}{2\pi} \frac{k_x \, dk_y}{(2\pi)^2} \frac{e^{i\lambda(z-L)}}{\lambda} \lambda^2 \big[\sin(\lambda z - \lambda L) + i \cos(\lambda z - \lambda L) \big] \tag{728}$$

$$= -\frac{3}{6\pi^2} \int\limits_0^\infty \frac{\kappa^3}{2} \, d\kappa \tag{729}$$

Of course, this quantity is divergent.

Finally, in order to complete the determination of the stress–energy tensor, one must calculate $\langle T_{xx} \rangle = \langle T_{yy} \rangle$. It can be expressed as

$$\langle T_{xx} \rangle = \frac{1}{2i} \int \frac{d\omega}{2\pi} \frac{dk_x \, dk_y}{(2\pi)^2} (\partial_t \partial_{\bar{t}} + \partial_x \partial_{\bar{x}} - \partial_y \partial_{\bar{y}} - \partial_z \partial_{\bar{z}}) g(z, \bar{z}) \Big|_{z=\bar{z}} \tag{730}$$

$$= \frac{1}{2i} \int \frac{d\omega}{2\pi} \frac{dk_x \, dk_y}{(2\pi)^2} \frac{1}{\lambda \sin(\lambda L)} \big[(\omega^2 + k_x^2 - k_y^2) \sin(\lambda z) \sin(\lambda L - \lambda z) + \lambda^2 \cos(\lambda z) \cos(\lambda L - \lambda z) \big] \tag{731}$$

A first step consists in dealing with the terms proportional to $k_x^2 - k_y^2$. In fact these terms are of the form $\int d\mathbf{k}_\perp (k_x^2 - k_y^2) \mathcal{F}(k_\perp)$, where $\mathcal{F}(k_\perp)$ represents the other terms in the integral which are function of $k_x^2 + k_y^2$ only. Therefore, one can evaluate them by going to polar coordinates in the plane (k_x, k_y). This leads to an expression of the form $\int dk_\perp d\psi \, k_\perp^2 (\cos^2 \psi - \sin^2 \psi) \mathcal{F}(k_\perp) = 0$, which shows that the corresponding contribution vanishes. Then, following the same steps as before, one reduces the above expression to

$$\langle T_{xx} \rangle = \frac{1}{12\pi^2} \int\limits_0^\infty d\kappa \, \frac{\kappa^3}{\sinh(\kappa L)} \big[\cosh(2\kappa z - \kappa L) + 2 \cosh(\kappa z) \cosh(\kappa L - \kappa z) \big] = -\langle T_{00} \rangle \tag{732}$$

where we have used that $2 \cosh(\kappa z) \cosh(\kappa L - \kappa z) = \cosh(\kappa L) + \cosh(2\kappa z - \kappa L)$. Of course, as already mentioned, this also shows that $\langle T_{yy} \rangle = -\langle T_{00} \rangle$. All the other components being zero, this completes our calculation of the stress–energy in the cavity. Only the calculations of $\langle T_{xx} \rangle = \langle T_{yy} \rangle$ outside the plates remains to be done and we now turn to this question. Considering again Eq. (730), we have

$$\langle T_{xx} \rangle = \frac{1}{2i} \int \frac{d\omega}{2\pi} \frac{dk_x \, dk_y}{(2\pi)^2} \frac{e^{i\lambda(z-L)}}{\lambda} \big[\omega^2 \sin(\lambda z - \lambda L) - i\lambda^2 \cos(\lambda z - \lambda L) \big] = -\langle T_{00} \rangle \tag{733}$$

We are now in a position where we can write the full stress–energy tensor. Inside the cavity, it takes the form

$$\langle T_{\mu\nu} \rangle_{\text{inside}} = (\rho_{\text{vac}} + \rho_{\text{Casimir}}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} + g(z) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{734}$$

where we have defined the two quantities ρ_{vac} and ρ_{Casimir} by

$$\rho_{\text{vac}} = -\frac{1}{6\pi^2} \int\limits_0^\infty d\kappa \, \frac{\kappa^3}{2}, \qquad \rho_{\text{Casimir}} = -\frac{\pi^2}{1440 L^4}. \tag{735}$$

Outside the cavity (more precisely on the right-hand side), the results derived above imply that

$$\langle T_{\mu\nu}\rangle_{\text{outside}} = \rho_{\text{vac}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} - \frac{1}{16\pi^2(z-L)^4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{736}$$

As discussed in Ref. [116], the terms proportional to $g(z)$ and $(z-L)^{-4}$ are surface divergent terms. They can be removed in the case where conformal invariance is re-established and, for that reason, are not present in electromagnetism, i.e. when one considers an electric field inside the cavity rather than a scalar field. In this case, we see that the difference between the stress–energy tensors inside and outside the cavity reads

$$\Delta\langle T_{\mu\nu}\rangle = \rho_{\text{Casimir}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \tag{737}$$

which is a finite result. The infinite part has been removed by subtraction. Finally, the stress–energy tensor can be written in a covariant form, namely [119]

$$\Delta\langle T_{\mu\nu}\rangle = \rho_{\text{Casimir}}\left(4\hat{z}^{\mu}\hat{z}^{\nu} - g^{\mu\nu}\right) \tag{738}$$

where $\hat{z}^{\mu}$ in the unit vector in the z direction.

At this point, several comments are in order. Firstly, we notice in Eqs. (737) and (738) that the energy density (i.e. the time–time component) is similar to the expressions already derived in Section 12.2.1 by other methods. This confirms that this calculation is independent of the regularization scheme used to obtain the result. Secondly, it is clear that the stress–energy tensor (738) is not similar to the cosmological constant stress–energy tensor $T_{\mu\nu} = -\rho_{\text{vac}}g_{\mu\nu}$. This illustrates the difference between the Casimir case and the cosmological constant case. In the latter case, one measures the absolute value of the vacuum energy while, in the former case, one is only sensitive to the difference of vacuum energy from one side of the plates to the other.

Thirdly, the Casimir Force has been observed in the laboratory and this is usually taken as another evidence in favor of the existence of the zero-point fluctuations. After all, since we observe the Casimir force and since this force is due to a change in the structure of the vacuum that is perturbed by the presence of the plates, the fact is that the vacuum must exist as a real physical phenomenon and is not an artifact of the quantum field theory formalism. In addition, this also validates the methods of regularization used to calculate ρ_{vac}. In the context of the Casimir effect, these methods seem to work very well and to lead to predictions that are in agreement with the experiments. Therefore, why would the very same techniques, used in a similar context, fail to regularize the cosmological constant?

Fourthly, the claim that the calculation and the observation of the Casimir force are evidences in favor of the reality of vacuum fluctuations has been challenged in Refs. [120,121]. It is indeed possible to derive the expression of the Casimir force without referring at all to zero point fluctuations, by means of the "source theory" approach to quantum field theory. However, to our knowledge, there is no attempt to do the same for the Lamb shift (but this would certainly be an interesting exercise). Therefore, it seems that zero-points fluctuations are nevertheless necessary to quantum field theory even if the above remark should lead us to tone down the claim that they have been seen in the laboratory. Let us also notice that, very recently, another situation where the vacuum fluctuations can lead to an observable effect has been studied in Ref. [122]. This article investigates whether the quantum zero point fluctuations could cause the rotation of a small nano device. This "rotational vacuum effect" has in fact a similar origin than the Casimir effect.

Finally, it should be clear that, if the Lamb shift and the Casimir effect seem to indicate that the vacuum fluctuations are real, they do not say anything with regards to their gravitational properties. In other words, they do not say how the vacuum fluctuations weigh (it is worth noticing here that the gravitational properties of a Casimir cavity have been studied in Refs. [123–129]). It is clear that there exists the possibility that the zero-point fluctuations do exist but have non-standard gravitational properties, i.e. that they abnormally weigh. It is therefore interesting to investigate if there is an experiment which measures whether this is true or not. In the next section, we turn to this question.

13. Do the quantum fluctuations gravitate?

In this section, we study the gravitational property of the Lamb shift [12,130]. We will follow the treatment of Ref. [130]. In Section 12.1, we showed that the interaction of the electron with the "vacuum electric field" inside the atom leads to a shift of the atomic levels. In some sense, this means that the zero-point fluctuations affect the mass (energy) of the atom. Therefore, by studying how the atom falls down, one can put constraints on the gravitational properties of the vacuum fluctuations. This is the idea that we pursue in this section. However, since the corresponding effect is small for an atom and is larger for a nucleus (of course, the Lamb shift also exists in this case), it is better to study the latter case. For this reason, we start with recalling some basics facts about nuclear models.

The main idea is that the interaction between a given nucleon and the other $A - 1$ nucleons can be mimicked by a Woods–Saxon potential given by

$$V_{\text{WS}}(r) = -V_0\left[1 - \exp\left(\frac{r-R}{a}\right)\right] \tag{739}$$

where a describes how abrupt the potential is while R is the radius of the nucleus. As is well known one can take $R = r_0 A^{1/3}$ where $r_0 \simeq 1.3 \times 10^{-15}$ m $= 1.3$ fm. The quantity V_0 represents the depth of the potential and can be taken to be $V_0 \simeq 45$ MeV. In fact, in order to have a reasonable model, one must add to the Woods–Saxon potential a spin orbit interaction term which can be expressed as

$$V_{\text{L,s}} = -f(r)\boldsymbol{L} \cdot \boldsymbol{s} \tag{740}$$

where $\boldsymbol{L}$ is the angular momentum of the nucleon and $\boldsymbol{s}$ its spin. With a good approximation, one can write $f(r) \simeq 24 A^{-2/3}$ MeV. Then, the Hamiltonian of the nucleus can be written as

$$H = \sum_{j=1}^{A}\left[\frac{p_j^2}{2m_j} + V_{\text{WS}}(r_j) - f(r_j)\boldsymbol{L}_j \cdot \boldsymbol{s}_j\right] \tag{741}$$

The total wave-function can be taken as the product of the individual wave functions (in fact, more precisely, one should consider a Slater determinant in order to take into account the fact that the total wave function must be anti-symmetrized), that is to say

$$\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_A) = \prod_{j=1}^{A} \Psi_j(\boldsymbol{r}_j) \tag{742}$$

Each individual wave-function can be obtained from a Schrödinger equation given by (for the moment, one ignores the spin orbit term since it can be treated as a perturbation)

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + V_{\text{WS}}(r_j)\right]\Psi_j(\boldsymbol{r}_j) = E_j \Psi_j(\boldsymbol{r}_j) \tag{743}$$

where μ is the reduced mass,

$$\frac{1}{\mu} \equiv \frac{1}{m_{\text{p}}} + \frac{1}{(A-1)m_{\text{p}}} \tag{744}$$

or $\mu \simeq m_{\text{p}}(A-1)/A \simeq m_{\text{p}} \simeq 1.6 \times 10^{-27}$ kg, m_{p} being the proton mass. Unfortunately, the Schrödinger equation cannot be solved analytically for the Woods–Saxon potential and, therefore, we are forced to make approximations. The simplest one is just to describe the Woods–Saxon potential with a parabola, i.e.

$$V_{\text{WS}}(r) \simeq -V_0\left(1 - \frac{r^2}{R^2}\right) = -V_0 + \frac{1}{2}\mu\omega_0^2 r^2 \tag{745}$$

with $\omega_0 \equiv \sqrt{2V_0/(\mu R^2)}$. In this case, the question becomes solvable. We can use the spherical symmetry of the problem and write the wave function as

$$\Psi_j(\boldsymbol{r}_j) = \frac{u_{n,\ell}(r_j)}{r_j} Y_{\ell m}(\theta_j, \varphi_j) \tag{746}$$

where $Y_{\ell m}$ is a spherical harmonics [48,49]. The radial function $u_{n\ell}$ is controlled by the equation

$$\left[-\frac{\hbar^2}{2\mu}\frac{\text{d}^2}{\text{d}r_j^2} + \frac{\ell(\ell+1)}{2\mu r_j^2} + \frac{1}{2}\mu\omega_0^2 r^2\right]u_{n,\ell} = (E_j + V_0)u_{n,\ell} \tag{747}$$

which can be solved in terms of Laguerre polynomials L_n^{α} [48,49]. Explicitly, one has

$$\Psi(\boldsymbol{r}_j) = C r_j^{\ell} e^{-\mu\omega_0 r_j^2/(2\hbar)} L_{n-1}^{\ell+1/2}\left(\frac{\mu\omega_0}{\hbar}r^2\right) Y_{\ell m}(\theta_j, \varphi_j) \tag{748}$$

where C is a normalization constant to be determined. In the above expression, one has $n \geqslant 1$ and $\ell \geqslant 0$. The energy levels are given by the expression

$$E(n, \ell) = -V_0 + \hbar\omega_0\left[2(n-1) + \ell + \frac{3}{2}\right] \tag{749}$$

The factor $3/2 = 3 \times 1/2$ originates from the fact that we have a three-dimensional harmonic oscillator. The quantum numbers have been chosen so that the ground state is $n = 1$, $\ell = 0$ which allows the spectroscopic notation "1s", see Fig. 9.
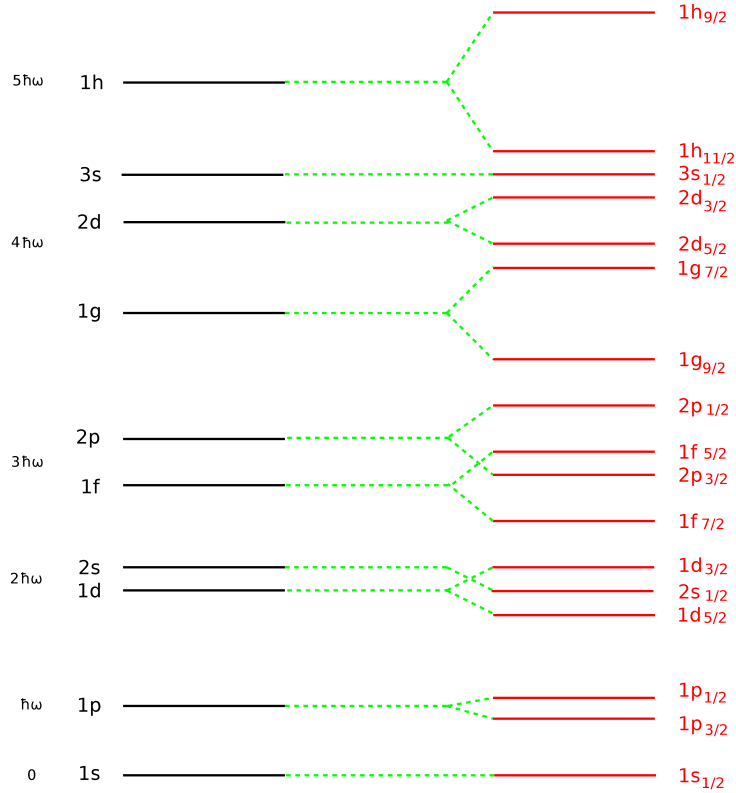
Fig. 9. Energy levels of the nucleus. On the left-hand side are represented the levels in the approximation where the Woods–Saxon potential is approximated by a parabola. In this case, each level (which includes a series of sub-levels) is separated by the energy $\hbar\omega$. On the right-hand side are represented the levels where the spin orbit interaction (740) is taken into account. This breaks the degeneracy of the sub-levels and reproduces the magic numbers.

The only thing which remains to be done is to determine the constant C. Obviously, this is done by normalizing the wave-function, i.e.

$$|C|^2 \int_0^\infty dr_j\, r_j^2 \frac{|u_{n,\ell}|^2}{r_j^2} \int d\Omega_j\, Y_{\ell m} = 1 \tag{750}$$

Using the solution obtained above and performing the change of variable $\rho = (\mu\omega_0/\hbar)r^2$, the previous condition reduces to

$$\frac{|C|^2}{2}\left(\frac{\mu\omega_0}{\hbar}\right)^{-\ell-3/2} \int_0^\infty \rho^{\ell+1/2} e^{-\rho} L_{n-1}^{\ell+1/2}(\rho) L_{n-1}^{\ell+1/2}(\rho) = 1 \tag{751}$$

Then, using Eq. (7.414.3) of Ref. [49], one obtains

$$|C|^2 = 2(n-1)!\left(\frac{\mu\omega_0}{\hbar}\right)^{\ell+3/2} \Gamma^{-1}\left(n+\ell+\frac{1}{2}\right) \tag{752}$$

We are now in a position where we can compute the value of the wave-function at $r_j = 0$ for $\ell = 0$. This gives

$$\big|\Psi(0)\big|^2 = |C|^2 \big[L_{n-1}^{1/2}(0)\big]^2 \frac{1}{4\pi} \tag{753}$$

Using the fact that $L_{n-1}^{1/2}(0) = \binom{n-1/2}{n-1} = \Gamma(n+1/2)/[\sqrt{\pi}(n-1)!/2]$, one obtains

$$\big|\Psi_{n,\ell=0}(0)\big|^2 = \frac{2}{\pi^2}\frac{\Gamma(n+1/2)}{(n-1)!}\left(\frac{\mu\omega_0}{\hbar}\right)^{3/2} \tag{754}$$

This completes the first part of the calculation.

We have established on very general grounds that the shift in energy due to the interaction of a charged particle (in Section 12.1, the electron in the atom, here the proton in the nucleus) is given by Eq. (642). The means value of the square of the displacement has been calculated in Eq. (650). Applying these formulas to the case of the protons in the nucleus, one obtains [130]

$$\Delta E(n, 0) \simeq \frac{4\alpha^2}{3m_p^2} \sum_i Z^{(i)} \big| \Psi_{n,\ell=0}^{(i)}(0) \big|^2 \int_0^\infty \frac{d\omega}{\omega} \tag{755}$$

As usual, this infinite integral must be regularized and usually leads to logarithmic corrections, see for instance Eq. (652). We will ignore this factor in what follows since this does not affect too much the final result. As explained in Ref. [130], the sum over "(i)" runs only over s-wave since this is the only way to get $\Psi_{n,\ell}(0) \neq 0$. Moreover, still following Ref. [130], it is reasonable to approximate the field experienced by a proton by a field produced by the other protons in inner shells. This is why $Z^{(i)}$ represents the total number of protons present in shells inner than the proton "(i)" under consideration.

Our goal is now to test how the vacuum fluctuations weigh. For this purpose, we write the gravitational mass of a given nucleus as [130]

$$m_g = m_i - \eta \Delta E \tag{756}$$

where m_i is the inertial mass. The parameter η is a priori a free parameter which controls to which extent the weak equivalence principle would be violated. If $\eta = 0$, there is no violation of the universality of free fall. A convenient measure of potential violations of the weak equivalence principle is given by the Eötvös ratio

$$\eta(1, 2) \equiv \frac{m_g(1)}{m_i(1)} - \frac{m_g(2)}{m_i(2)} \tag{757}$$

for two bodies 1 and 2. Using the previous calculations, this ratio can be expressed as

$$\eta(1, 2) = \eta \frac{4\alpha^2}{3m_p^3} \bigg[\frac{1}{A_1} \sum_i Z_1^{(i)} \big| \Psi_{n,\ell=0}^{(1,i)}(0) \big|^2 - \frac{1}{A_2} \sum_i Z_2^{(i)} \big| \Psi_{n,\ell=0}^{(2,i)}(0) \big|^2 \bigg], \tag{758}$$

where we have used $m_i \sim A m_p$, a relation valid since we work at first order in the parameter η. Then, we use the fact that

$$(\mu\omega_0)^{3/2} = \bigg(\frac{2m_p V_0}{r_0^2} \bigg)^{3/4} \frac{1}{A^{1/2}} \tag{759}$$

to express the Eötvös ratio as [130]

$$\eta(1, 2) = \eta \frac{4\alpha^2}{3} \bigg(\frac{2V_0}{m_p^3 r_0^2} \bigg)^{3/4} \frac{2}{\pi^2} \bigg[\frac{1}{A_1^{3/2}} \sum_i Z_1^{(i)} \frac{\Gamma(n+1/2)}{(n-1)!} - \frac{1}{A_2^{3/2}} \sum_i Z_2^{(i)} \frac{\Gamma(n+1/2)}{(n-1)!} \bigg] \tag{760}$$

Let us evaluate this number for aluminum ($Z = 13$, $A = 27$) and platinum ($Z = 78$, $A = 195$). In the case of aluminum, since we only have 13 protons, the only s-shell protons are on the $1s_{1/2}$ level (in fact, we have two protons in $1s_{1/2}$, four protons in $1p_{3/2}$, two protons in $1p_{1/2}$ and five protons in $1d_{5/2}$, see Fig. 9). As a consequence, there is no inner protons for those protons (since, obviously, there is no inner level than the ground state $1s_{1/2}$) and the corresponding contribution in the Eötvös ratio vanishes. In the case of platinum, all the levels are filled up until the level $1h_{11/2}$ which contains height protons, see Fig. 9. Therefore, we have two protons on $1s_{1/2}$, two protons on $2s_{1/2}$ and two on $3s_{1/2}$, see again Fig. 9. This means that the $3s_{1/2}$ protons see four inner protons and the $2s_{1/2}$ protons see two inner protons. As a consequence, one can write

$$\eta(\text{Al}, \text{Pt}) \simeq -\eta \frac{4\alpha^2}{3} \bigg(\frac{2V_0}{m_p^3 r_0^2} \bigg)^{3/4} \frac{2}{\pi^2} \frac{1}{A_{\text{Pt}}^{3/2}} \bigg[\frac{2}{(2-1)!} \Gamma\bigg(2+\frac{1}{2}\bigg) + \frac{4}{(3-1)!} \Gamma\bigg(3+\frac{1}{2}\bigg) \bigg] \tag{761}$$

Using the values of V_0 and r_0 chosen previously this leads to

$$\eta(\text{Al}, \text{Pt}) \simeq -5.6 \times 10^{-10} \eta \tag{762}$$

Given the fact that $|\eta(\text{Al}, \text{Pt})| \lesssim 10^{-12}$ [131,132], we reach the conclusion that

$$\eta \lesssim 1.7 \times 10^{-3} \tag{763}$$

Therefore, the vacuum fluctuations seem to have standard gravitational properties to a very good accuracy.

The conclusion of this section is that it seems difficult to argue that the zero-point fluctuations do not gravitate. Clearly, this makes the cosmological constant problem more acute. However, the results obtained here are by no means a final proof. For instance, it seems reasonable to assume that the stress–energy tensor of the Lamb shift is not of the cosmological
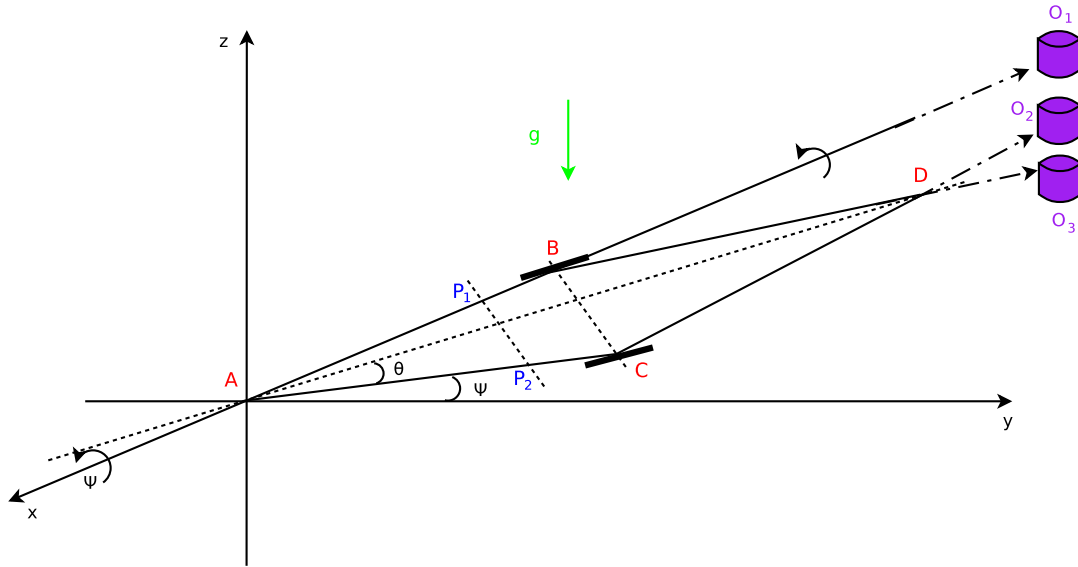
Fig. 10. Representation of the Colella, Overhauser and Werner (COW) interferometer. The interferometer is placed in a vertical gravitational field (represented by the green arrow) and is rotated around the axis AB.

constant type. Therefore, the fact that a nucleus falls down normally does not necessarily imply that the cosmological constant weighs according to the Einstein equations. Nevertheless, no sign of inconsistency appears in the above situation and one is obviously tempted to assume that zero-point fluctuations couple to gravity as any other type of matter. Again, this makes the cosmological constant problem more mysterious.

14. The weak equivalence principle in quantum mechanics

In the previous sections, we have argued that the zero-point fluctuations seem to be real and seem to gravitate normally. Therefore, these two reasons cannot be invoked to avoid the cosmological constant problem. However, there is another (but related) issue that it is interesting to investigate. When we consider the vacuum stress–energy tensor in the Einstein equations, we implicitly assume that the zero-point fluctuations obey the weak equivalence principle, i.e. that they "fall" in a gravitational field as any other type of matter. Clearly the zero point fluctuations are of quantum-mechanical origin. The weak equivalence principle is well established (and tested) in classical physics and general relativity is based on this principle. But what is the status of this principle in quantum mechanics? If it is not valid in quantum mechanics, maybe we do not have the right to couple the quantum zero point fluctuations in the standard way in the Einstein equations? Could it be a way to avoid the problem? In order to discuss this issue, we will consider two situations, the Colella, Overhauser and Werner experiment (COW) and the quantum Galileo experiment.

14.1. The Colella, Overhauser and Werner (COW) experiment

The COW experiment [133–135] consists in sending neutrons in an interferometer placed in a weak gravitational field, see Fig. 10. The presence of the gravitational field affects the wave-function of the neutrons and changes the relative phases between the two beams. When the interferometer is rotated, the gravitational phase shift is modified because the height difference between the two arms is changed. This effect can then be detected (and was detected) in the interference pattern. The COW experiment is conceptually very important. This was the first experiment to measure the effect of the gravitational field on the wave-function. In fact, it shows that the gravitation is just an ordinary force as far as the Schrödinger equation is concerned.

Let us now describe the experiment in more detail. The neutron interferometer is represented in Fig. 10. Let us assume that, along one of the two paths, a phase shift $\Delta\varphi$ is introduced. Our first goal is to evaluate the intensity seen by the detectors O_2 and O_3. After diffraction, the wave function of the neutron takes the form

$$\Psi(\boldsymbol{r}) = \Phi T(\theta)e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}} + \Phi D(\theta)e^{-i\boldsymbol{k}_G \cdot \boldsymbol{r}} \tag{764}$$

where $\boldsymbol{k}_0$ is the incident wave-number and $\boldsymbol{k}_G$ the wave number after Bragg diffraction (which, therefore, must satisfy the Bragg condition). The quantity Φ is the amplitude of the incident wave and $T = T(\theta)$ (for "transmitted") and $D = D(\theta)$ (for "diffracted") are two coefficients describing the amplitude of the two branches of the wave-function after the neutron has emerged from the crystal. Let us now consider the path $ABDO_3$. At point B, the wave function is given by (the origin of the coordinates is chosen to be at point A)

$$\Psi(AB) = \Phi T(\theta) e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}_B} \tag{765}$$

since the wave was just transmitted. Then, at point B, we have a diffraction and, therefore, the wave-function at point D reads

$$\Psi(ABD) = \Phi T(\theta) e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}_B} D(\theta) e^{i\boldsymbol{k}_G \cdot (\boldsymbol{r}_D - \boldsymbol{r}_B)} \tag{766}$$

Finally, at point D, we have another diffraction but this time with opposite angle. The leads to

$$\Psi(ABDO_3) = \Phi T(\theta) e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}_B} D(\theta) e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_D - \boldsymbol{r}_B)} D(-\theta) e^{i\boldsymbol{k}_G \cdot (\boldsymbol{r} - \boldsymbol{r}_D)} \tag{767}$$

We also assume that, along that path, there is a phase shift $\Delta\varphi$ that we do not specify for the moment. As a consequence, the total wave function can be written as

$$\Psi(ABDO_3) = \Phi T(\theta) e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}_B} D(\theta) e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_D - \boldsymbol{r}_B)} D(-\theta) e^{i\Delta\varphi} e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r} - \boldsymbol{r}_D)} \tag{768}$$

where we have introduced the factor $e^{i\Delta\varphi}$.

It is straightforward to follow the same procedure and to establish the expression for $\Psi(ACDO_3)$. This gives the expression of the wave function at detector O_3, namely $\Psi(O_3) = \Psi(ABDO_3) + \Psi(ACDO_3)$,

$$\Psi(O_3) = \Phi\left[T(\theta) e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}_B} D(\theta) e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_D - \boldsymbol{r}_B)} D(-\theta) e^{i\Delta\varphi} + D(\theta) e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}_C} D(-\theta) e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_D - \boldsymbol{r}_C)} T(\theta)\right] e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r} - \boldsymbol{r}_D)} \tag{769}$$

Using this expression, one can calculate the expected intensity. Straightforward manipulations leads to the following expression

$$I(O_3) = \left|\Psi(O_3)\right|^2 = \alpha(1 + \cos\Delta\varphi) \tag{770}$$

where α is a constant that can be determined from the above expression and which depends on the incident flux and the crystal structure. Then, one can repeat the same analysis and determine the wave function at detector O_2. It reads

$$\Psi(O_2) = \Phi\left[T(\theta) e^{i\boldsymbol{k}_0 \cdot \boldsymbol{r}_B} D(\theta) e^{i\boldsymbol{k}_G \cdot (\boldsymbol{r}_D - \boldsymbol{r}_B)} T(-\theta) e^{i\Delta\varphi} + D(\theta) e^{i\boldsymbol{k}_G \cdot \boldsymbol{r}_C} D(-\theta) e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_D - \boldsymbol{r}_C)} D(\theta)\right] e^{i\boldsymbol{k}_G \cdot (\boldsymbol{r} - \boldsymbol{r}_D)} \tag{771}$$

The intensity at detector O_2 can be computed along the same lines and one is led to

$$I(O_2) = \beta - \alpha \cos\Delta\varphi \tag{772}$$

where α and β are just constants that do not play an important role in what follows (of course, the constant α is the same as in Eq. (770)).

Having estimated the form of the signal expected for any phase shift $\Delta\varphi$, we now turn to its calculation in the case of the COW experiment. For this purpose, we return to the Schrödinger equation which can be written as

$$\nabla^2 \Psi + \omega^2 \Psi = 0 \tag{773}$$

with

$$\omega^2 = \frac{2m}{\hbar^2}(E - V) \tag{774}$$

Now, let us assume that the potential is given by

$$V = V_0 + \Delta V(\boldsymbol{r}) \tag{775}$$

where V_0 is a constant and ΔV a small space-dependent small perturbation, i.e. $\Delta V \ll V_0$. If the perturbation does not change abruptly, the Wentzel–Kramers–Brillouin (WKB) wave-function

$$\Psi \propto \exp\left[i \int \omega(\boldsymbol{r}) \, d\boldsymbol{r}\right] \tag{776}$$

is a very good approximation to the actual solution. Introducing the notation $k_0 = 2m(E - V_0)/\hbar^2$ and using that $p = mv = \hbar k_0$, an expansion of the phase in $\Delta V/V_0$ leads to the following expression

$$\Psi \propto \exp\left(i\boldsymbol{k}_0 \cdot \boldsymbol{r} - \frac{1}{\hbar} \int_{\mathcal{P}} \Delta V \, dt\right) \tag{777}$$

where the subscript "$\mathcal{P}$" indicates that one integrates along the non-perturbed path. From the above equation, we conclude that
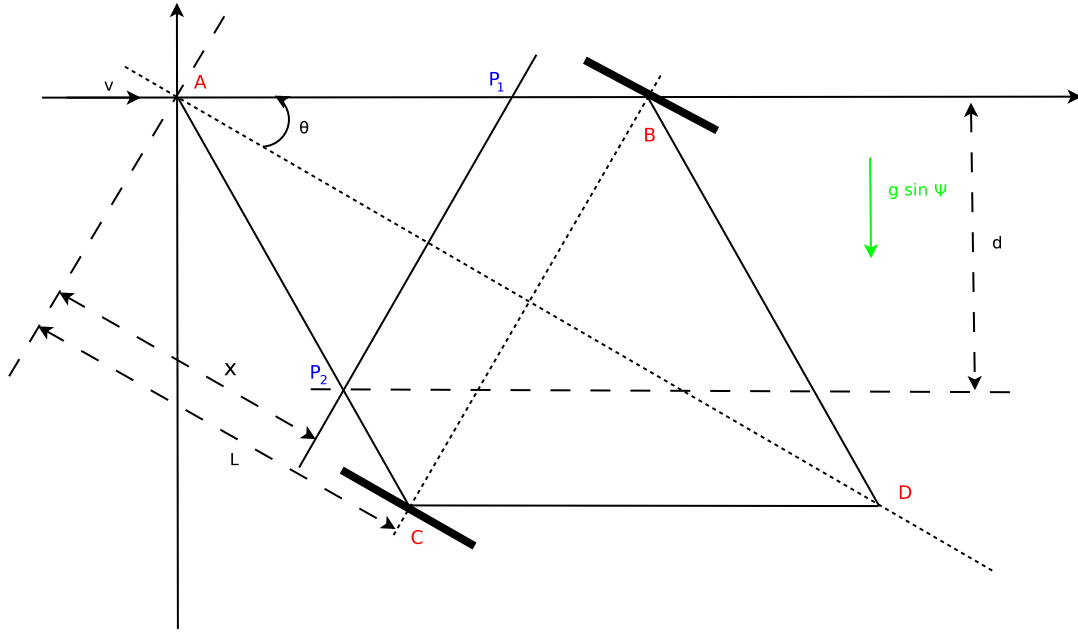
Fig. 11. Representation of the COW interferometer in the plane of the two beams.

$$\Delta\varphi = -\frac{1}{\hbar}\int_{\mathcal{P}}\Delta V\,dt \tag{778}$$

The same result can also be expressed differently. Using energy conservation, on can write

$$\frac{p^2}{2m} + V_0 + \Delta V = E \tag{779}$$

and, since $\boldsymbol{p} = \boldsymbol{p}_0 + \delta\boldsymbol{p}$, one obtains

$$\frac{p_0^2}{2m} + \frac{\boldsymbol{p}_0}{m}\cdot\delta\boldsymbol{p} + V_0 + \Delta V = E \tag{780}$$

or,

$$\boldsymbol{v}_0 \cdot \delta\boldsymbol{p} = -\Delta V \tag{781}$$

As a consequence, the phase shift can be re-written as the following expression

$$\Delta\varphi = \frac{1}{\hbar}\int_{\mathcal{P}}\delta\boldsymbol{p}\cdot d\boldsymbol{r} \tag{782}$$

where the integral should again be calculated along the path of the neutron.

Let us now determine the phase shift for the COW experiment. Let us consider two point P_1 and P_2 that correspond to one beam transmitted and one beam diffracted in the interferometer, see Fig. 11. The potential difference between those two points is given by

$$V(P_2) - V(P_1) = mgd\sin\psi \tag{783}$$

where d can be expressed as $d = AP_1\cos(\pi/2 - 2\theta)$. Given that the quantity x in Fig. 11 can be written as $x = AP_1\cos\theta$, one deduces that $d = 2x\sin\theta$. As a consequence,

$$V(P_2) - V(P_1) = 2m_{\mathrm{grav}}gx\sin\theta\sin\psi \tag{784}$$

where we have now carefully written that the mass which appeared in the above expression is the gravitational mass (since this is the "coupling constant" to the gravitational field). This implies that the phase shift can be written as

$$\Delta\varphi = -\frac{2}{\hbar}\int_0^L 2m_{\mathrm{g}}g_0 x\sin\theta\sin\psi\,\frac{dx}{\cos\theta} \tag{785}$$

where we have taken into account the fact that the line $P_1 P_2$ is "moving" with a speed $v \cos \theta$. The factor two in front of the whole expression comes from the fact that the phase shift accumulates and is, along the total path, twice its value in the fist part of the interferometer. The integration over the variable x is easily done and we obtain

$$\Delta \varphi = -2 \frac{m_g g_0}{\hbar v} L^2 \tan \theta \sin \psi \tag{786}$$

Finally, noticing that the total area of the interferometer is nothing but $\mathcal{A} \equiv 2L^2 \tan \theta$ and that the de Broglie relation is $v = h/(m_{in} \lambda)$, where m_{in} is the inertial mass and λ the wavelength of the incident beam, we arrive at [133–135]

$$\Delta \varphi = -\frac{2\pi}{h^2} m_{grav} m_{in} g_0 \mathcal{A} \lambda \sin \psi \tag{787}$$

This expression is quite remarkable as it involves the gravitational and the inertial mass together with the Planck constant. Moreover, we deal with an observable quantity which does not depend on the ratio of the gravitational mass to the inertial mass, as is usually the case, but on the product of these quantities. We also see that the intensity of the signal at detectors O_2 and/or O_3 depend on the orientation of the interferometer ψ.

However, in practice, one must deal with two other effects [133–135]. The first one is the Sagnac effect due to the rotation of Earth. Classically, the Hamiltonian of a neutron is given by

$$H = \frac{p^2}{2m_{in}} + m_{grav} \boldsymbol{g} \cdot \boldsymbol{r} - \boldsymbol{\omega} \cdot \boldsymbol{L} \tag{788}$$

where $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ is the angular momentum of the neutron and $\boldsymbol{\omega}$ the angular velocity of Earth. Upon using the Hamilton equations, this implies that

$$\boldsymbol{p} = m_{in} \dot{\boldsymbol{r}} + m_{in} \boldsymbol{\omega} \times \boldsymbol{r} \tag{789}$$

Therefore, one obtains a Sagnac phase shift which can be expressed as

$$\Delta \varphi_{Sagnac} = \frac{m_{in}}{\hbar} \oint (\boldsymbol{\omega} \times \boldsymbol{r}) \cdot d\boldsymbol{r} \tag{790}$$

where we have used Eq. (782). Using Stokes theorem, this can also be written as

$$\Delta \varphi_{Sagnac} = \frac{4\pi m_{in}}{h} \boldsymbol{\omega} \cdot \mathcal{A} \tag{791}$$

where $\mathcal{A}$ is the vector associated to the interferometer area. If Γ is the incident neutron beam west of due south and δ the co-latitude, the Sagnac phase shift takes the form

$$\Delta \varphi_{Sagnac} = \frac{4\pi m_{in}}{h} \omega \mathcal{A} (\cos \psi \cos \delta + \sin \psi \sin \Gamma \sin \delta) \tag{792}$$

The COW experiment is such that the incident beam is directed due south which implies that $\Gamma = 0$. In the case, one obtains

$$\Delta \varphi_{Sagnac} = \frac{4\pi m_i}{h} \omega \mathcal{A} \cos \delta \cos \psi \tag{793}$$

a dependence in ψ which is different from the gravitational phase shift, see Eq. (787).

Finally, there is a shift due to the fact that the interferometer bends and/or warps under its own weight. The COW experiment claim that this effect can be described by the following equation

$$\Delta \varphi_{bend} = -q_{bend} \sin \psi \tag{794}$$

This expression can be justified by noticing that the bending effect depends on the rotation angle. Clearly, this effect is difficult to estimate from first principles.

Therefore, the total phase shift expected in the COW experiment is the sum of the three contributions discussed before, namely the gravity, Sagnac and bending shifts. This leads to the following expression

$$\Delta \varphi = -q_{grav} \sin \psi + q_{Sagnac} \cos \psi - q_{bend} \sin \psi \tag{795}$$

where, in order to take into account Eqs. (787) and (793), we have defined $q_{grav} \equiv -2\pi m_{grav} m_{in} g_0 \mathcal{A} \lambda / h^2$ and $q_{Sagnac} = 4\pi m_{in} \omega \mathcal{A} \cos \delta / h$. The above expression can be re-written as

$$\Delta \varphi = q \sin(\psi - \psi_0) \tag{796}$$

where $q^2 = (q_{grav} + q_{bend})^2 + q_{Sagnac}^2$ and $\tan \psi_0 = q_{Sagnac}/(q_{grav} + q_{bend})$. In practice, the Sagnac effect is only 2.5% of the gravitational effect but we see it leads to a global shift of the oscillatory pattern.
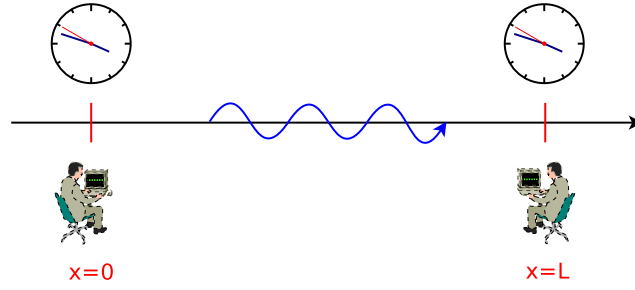
Fig. 12. Sketch of a quantum "Peres" clock. The clock is coupled to the particle traveling from the point $x = 0$ to the point $x = L$. When the particle crosses $x = 0$, the clock changes its state and starts running. When the particle reaches $x = L$, the clock stops and a comparison of the two clock states permits a measurement of the time of flight of the particle.

Let us now describe the result of the experiment. According to the previous considerations, the intensity observed at detector O_3 is

$$I(O_3) = \alpha \{1 + \cos[q\sin(\psi - \psi_0)]\} \tag{797}$$

By Fourier transforming this signal, one can extract the frequency q_{grav} of this oscillations. If one repeat this procedure for different wavelengths, one can fit the dependence of q_{grav} with λ which allows us to determine the product $m_{grav}m_{in}$. Let us recall that the neutron mass is usually obtained by mass spectroscopy on the deuteron which leads to the following value:

$$m_n = m_D - m_p + \frac{E_\gamma}{c^2} = 1.6747 \times 10^{-24} \text{ g} \tag{798}$$

where E_γ is deuteron binding energy. On the other hand, the COW experiment has found [133]

$$(m_{grav}m_{in})^{1/2} = (1.675 \pm 0.003) \times 10^{-24} \text{ g} \tag{799}$$

This shows that the weak equivalence principle seems to be satisfied even in a purely quantum mechanical situation. This also shows that gravity can be "coupled" to the Schrödinger equation in a standard way. The COW experiment is therefore very important since this was the first time that this was demonstrated. Moreover, if we consider that $m_n = m_{in}$ then the two previous equations show that the COW experiment proves that $m_{in} = m_{grav}$ even in the quantum domain.

14.2. The quantum Galileo experiment

It is well known that, classically, the motion of a body in a constant gravitational field only depends on the ratio m_{grav}/m_{in} (which is measured to be unity). However, if one writes the Schrödinger equation for the same situation, namely

$$i\hbar \frac{\partial \Psi(t,z)}{\partial t} = -\frac{\hbar^2}{2m_{in}} \frac{\partial^2 \Psi(t,z)}{\partial z^2} + m_{grav}gz\Psi(t,z) \tag{800}$$

it is apparent that m_{in} and m_{grav} no longer cancel out. Therefore, it is interesting to study how the universality of the free fall is recovered in quantum mechanics. We investigate this point in the following sections.

14.2.1. The Salecker–Wigner–Peres clock

In this section, we present a simple model of a quantum clock, the so-called "Salecker–Wigner–Peres clock" [136–140]. We introduce this system because we want to study how a quantum particle falls down in a constant gravitational field. One way to do that is to measure the time of flight of the particle and to compare it with its classical analogue. Hence the question of how to define the time of flight becomes important. As is well known, the measurement of time (or, to be more accurate, of time of arrival) is problematic in quantum mechanics [141–145]. In particular, there is no time operator because time is not a dynamical variable. One way out is to introduce a quantum system, the "clock", that is coupled to the particle and changes its state when the particle crosses the starting and the ending points along its "trajectory", see Fig. 12. Reading these two states then gives information about the time of flight. We now study how this general idea works in practice.

We assume that the clock has an odd number $N = 2j + 1$ of states. The Hamiltonian of the clock is given by

$$H_{clock} = \omega J = -i\hbar\omega \frac{\partial}{\partial \theta} \tag{801}$$

where the quantity θ can be viewed as describing the position of the clock's hand. The corresponding states of the clock are represented by the following wave-function

$$u_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta} \tag{802}$$

that are eigenfunction of H_{clock} since $H_{\text{clock}}u_n(\theta) = n\hbar\omega u_n(\theta)$. It is also easy to show that

$$e^{-iH_{\text{clock}}t/\hbar}u_n(\theta) = u_n(\theta - \omega t) \tag{803}$$

which gives the time evolution of the system.

It is also convenient to introduce another basis $v_s(\theta)$ defined by $(s = 0, \ldots, N-1)$

$$v_s(\theta) = \frac{1}{\sqrt{N}} \sum_{n=-j}^{n=j} e^{-\frac{2i\pi sn}{N}} u_n(\theta) \tag{804}$$

$$= \frac{1}{\sqrt{2\pi N}} \frac{\sin[N(\theta - 2\pi s/N)/2]}{\sin[(\theta - 2\pi s/N)/2]} \tag{805}$$

If N is large enough, these functions peak at $\theta = 2\pi s/N$ and this corresponds to the hand of the clock pointing to the s-th hour with an angle uncertainty of $\pm\pi/N$ (hence, as expected, if N is large, the accuracy is good). Moreover, using Eq. (803), we have

$$e^{-iH_{\text{clock}}t/\hbar}v_s(\theta) = v_s(\theta - \omega t) \tag{806}$$

This means that if we evaluate the above equation for a time τ such that $\omega\tau = 2\pi/N$, then one obtains

$$e^{-iH_{\text{clock}}\tau/\hbar}v_s(\theta) = v_{s+1 \text{ (mod } N)}(\theta) \tag{807}$$

which confirms the interpretation given before.

Let us now consider a free particle traveling along the x-axis. We want to measure the time of flight between $x = 0$ and $x = L$, see Fig. 12. In order for the clock to record this time, we must of course couple it to the particle. Following Ref. [137], we assume that the Hamiltonian of the total system "particle + clock" can be written as

$$H = \frac{p^2}{2m} + \mathcal{P}(x)H_{\text{clock}} \tag{808}$$

where $\mathcal{P}(x)$ is the projector operator which is one if $0 < x < L$ and zero otherwise and m is the mass of the particle. We write the eigenwave-functions of the system as

$$\Psi(t, x, \theta) = \psi(x, \theta)e^{-iEt/\hbar} = \frac{1}{\sqrt{N}} \sum_{n=-j}^{n=j} \psi_n(x)u_n(\theta)e^{-iEt/\hbar} \tag{809}$$

and inserting this form into the Schrödinger equation, one obtains

$$\sum_{n=-j}^{n=j} \left[-\frac{\hbar^2}{2m} \frac{d^2\psi_n}{dx^2} + n\hbar\omega\mathcal{P}(x)\psi_n(x) - E\psi_n(x) \right] u_n(\theta) = 0. \tag{810}$$

Therefore, we just have to solve the time-independent Schrödinger equation in presence of a rectangular barrier. Of course, this exercise is solved in any textbooks on quantum mechanics, see for instance Ref. [57]. It is straightforward to show that

$$\psi_n(x) = \begin{cases} Ae^{ikx} + Be^{-ikx}, & x < 0 \\ Ce^{ipx} + De^{-ipx}, & 0 < x < L \\ Ee^{ikx}, & x > L \end{cases} \tag{811}$$

where we have defined

$$k = \frac{\sqrt{2mE}}{\hbar} \tag{812}$$

$$p = \frac{\sqrt{2m(E - n\hbar\omega)}}{\hbar} = k\sqrt{1-\epsilon} \tag{813}$$

with $\epsilon \equiv n\hbar\omega/E$, a quantity which measures how the clock perturbs the particle. The quantities A, B, C, D and E are integration constants that are fixed by requiring the wave-function and its derivative to be continuous at the barrier. Notice that, after the barrier there is only one branch; this is of course because we assume that the particle is coming from the left, see Fig. 12. Matching ψ_n and $d\psi_n/dx$ at $x = 0$ and $x = L$, one obtains

$$A = \frac{1}{4p}\left[(k+p)e^{i(k-p)L} - (k-p)e^{i(k+p)L}\right]E + \frac{1}{4k}\left[(k+p)e^{i(k-p)L} + (k-p)e^{i(k+p)L}\right]E \tag{814}$$

$$B = \frac{1}{4p}\left[(k+p)e^{i(k-p)L} - (k-p)e^{i(k+p)L}\right]E - \frac{1}{4k}\left[(k+p)e^{i(k-p)L} + (k-p)e^{i(k+p)L}\right]E \tag{815}$$
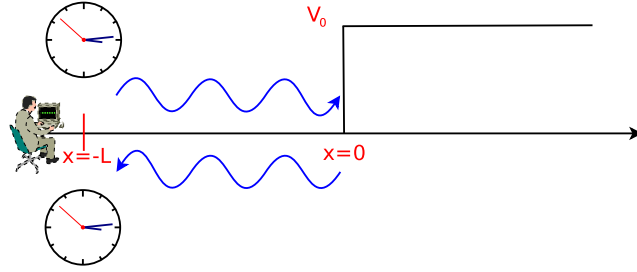
Fig. 13. Measurement of an out-and-back time of flight. An observer, located at $x = -L$ sends a particle towards a square potential step located at $x = 0$. The physicist turns on the clock when the particle leaves and turn it off when the reflected wave (i.e. the reflected particle) comes back and reaches $x = -L$.

Then, we assume that the clock does not perturb the particle too much and, as a consequence, that $\epsilon \ll 1$. In this case, $B \simeq 0$ and $A \simeq e^{i(k-p)L} E$. Indeed, we have $p = k\sqrt{1-\epsilon}$ which, at leading order, simply gives $k \simeq p$ (but, of course, one should use the next to leading order when one estimates the argument of the exponentials, see below). Then, initially, this means that the wave-function can be written as, see Eq. (809)

$$\psi(x,\theta) = \frac{1}{\sqrt{N}} \sum_{n=-j}^{n=j} A e^{ikx} u_n(\theta) = A e^{ikx} v_0(\theta), \quad x \ll 0 \tag{816}$$

because k does not depend on n, see Eq. (812). We see that, initially, the clock points to zeroth hour, as expected. After the barrier, using the previous considerations, the expression of the wave-function can be written as, see Eq. (809)

$$\psi(x,\theta) \simeq \frac{A}{\sqrt{N}} e^{ikx} \sum_{n=-j}^{n=j} e^{-i(k-p)L} u_n(\theta), \quad x \gg L \tag{817}$$

where we have used Eq. (804). But we have

$$p \simeq k\left(1 - \frac{\epsilon}{2}\right) \simeq k - \frac{n\omega}{(2E/m)^{1/2}} \tag{818}$$

which implies that (for $x \gg L$)

$$\psi(x,\theta) = \frac{A}{\sqrt{2\pi N}} e^{ikx} \sum_{n=-j}^{n=j} e^{in\theta - in\omega L/(2E/m)^{1/2}} \tag{819}$$

$$= A e^{ikx} v_0\left(\theta - \frac{\omega L}{\sqrt{2E/m}}\right) \tag{820}$$

This means, see Eq. (806), that the clock now indicates the time

$$t = \frac{L}{(2E/m)^{1/2}} = \frac{L}{v_{\text{clas}}} \tag{821}$$

where $v_{\text{clas}} \equiv (2E/m)^{1/2}$ is the classical velocity. Therefore, the clock has measured a time of flight which is nothing but the classical time of flight L/v_{clas} [137].

We now consider the situation represented in Fig. 13, see Ref. [140]. An observer, located at $x = -L$, sends a particle towards a reflecting barrier of height V_0 located at $x = 0$. When the particle leaves the Peres clock is turned on. Then the particle bounces back and when it comes back at the observer position, the Peres clock is turned off. With this experimental set up, one can measure the out-and-back time of flight. Following the same approach as before, the wave-function can be written as

$$\psi_n(x) = \begin{cases} A e^{ikx} + B e^{-ikx}, & x < 0 \\ C e^{-px}, & x > 0 \end{cases} \tag{822}$$

where we now have

$$k = \frac{\sqrt{2m(E - n\hbar\omega)}}{\hbar} = \frac{\sqrt{2mE}}{\hbar}\sqrt{1-\epsilon} \tag{823}$$

$$p = \frac{\sqrt{2m(V_0 + n\hbar\omega - E)}}{\hbar} = \frac{\sqrt{2m(V_0 - E)}}{\hbar}\left(1 + \frac{E}{V_0 - E}\epsilon\right)^{1/2} \tag{824}$$

Then, we have to match the wave-function and its derivative at $x = 0$. This gives

$$B = \frac{k - ip}{k + ip} A \equiv R A \tag{825}$$

$$C = \frac{2k}{k + ip} A \tag{826}$$

As before, we assume that, initially, the clock was is a given state, here $v_0(\theta)$. Therefore, the wave-function in the region $x < 0$ can be expressed as, see Eq. (809)

$$\psi(x, \theta) = \frac{A}{\sqrt{N}} \sum_{n=-j}^{n=j} \left(e^{ikx} + R e^{-ikx} \right) u_n(\theta) \tag{827}$$

But $|R| = 1$ and its phase is given by $-2 \arctan(p/k)$. As a consequence, one has

$$\psi(x = -L, \theta) = \frac{A}{\sqrt{N}} \sum_{n=-j}^{n=j} \left[e^{-ikL} + e^{-2i \arctan(p/k)} e^{ikL} \right] u_n(\theta) \tag{828}$$

The next step is similar to the case of a free particle, namely we expand the phases of the two wave-function branches in terms of the parameter ϵ. This gives

$$\psi(x = -L, \theta) = \frac{A}{\sqrt{2\pi N}} \sum_{n=-j}^{n=j} \exp\left[-\frac{iL}{\hbar} \sqrt{2mE} + in\left(\theta + \frac{\omega L}{v_{\text{clas}}} \right) \right]$$

$$+ \exp\left[-2i \arctan\left(\sqrt{\frac{V_0 - E}{E}} \right) + \frac{iL}{\hbar} \sqrt{2mE} + in\left(\theta - \frac{\omega L}{v_{\text{clas}}} - \frac{\hbar \omega}{\sqrt{E(V_0 - E)}} \right) \right] \tag{829}$$

$$= \frac{A}{\sqrt{N}} \exp\left(-\frac{iL}{\hbar} \sqrt{2mE} \right) v_0\left(\theta + \frac{\omega L}{v_{\text{clas}}} \right)$$

$$+ \frac{A}{\sqrt{N}} \exp\left[-2i \arctan\left(\sqrt{\frac{V_0 - E}{E}} \right) + \frac{iL}{\hbar} \sqrt{2mE} \right] v_0\left(\theta - \frac{\omega L}{v_{\text{clas}}} - \frac{\hbar \omega}{\sqrt{E(V_0 - E)}} \right) \tag{830}$$

and this allows us to directly read the time indicated by the clock. For the first branch, it is $-L/v_{\text{clas}}$ while for the second one, it is given by $L/v_{\text{clas}} + \hbar/\sqrt{E(V_0 - E)}$. As a consequence, the time of flight of the particle can be expressed as

$$\Delta t = \frac{2L}{v_{\text{clas}}} + \frac{\hbar}{\sqrt{E(V_0 - E)}} \tag{831}$$

Therefore, the time of flight is the classical one, $2L/v_{\text{clas}}$, but, this time, there is an additional contribution. This one can be understood as follows. The quantity Δt can be re-written as [140]

$$\Delta t = \frac{2(L + d)}{v_{\text{clas}}} \tag{832}$$

where $d = 1/p$, see Eq. (824) (evaluated at $\epsilon = 0$) is the penetration depth into the potential step. This means that there is an additional delay due to the tunnel effect and the fact that the particle has a non-vanishing probability to be below the barrier [140]. In this regime, the "velocity" is nevertheless given by its classical value v_{clas} [140].

This concludes this section on the Peres clock. We have seen that this is a useful device to measure times of flight. We now use it to study the motion of a quantum particle in a uniform gravitational field.

### 14.2.2. The gravitational case

In this section, we study the status of the weak equivalence principle and of the universality of the free fall in quantum mechanics [139,140,146–150].

Following Ref. [140], let us consider the situation where a particle is sent upwards in a uniform gravitational field, see Fig. 14. If z denotes the vertical coordinate, the corresponding classical situation is described by the following Lagrangian

$$L(\dot{z}, z) = \frac{m_{\text{in}}}{2} \dot{z}^2 - m_{\text{grav}} g z \tag{833}$$

where g is the gravitational field and where we have made the difference between the inertial mass m_{in} and the gravitational mass (the "gravitational charge") m_g. The solution to the classical equation of motion reads

$$z(t) = -\frac{1}{2} \frac{m_{\text{grav}}}{m_{\text{in}}} g t^2 + v_{\text{ini}} t + L \tag{834}$$
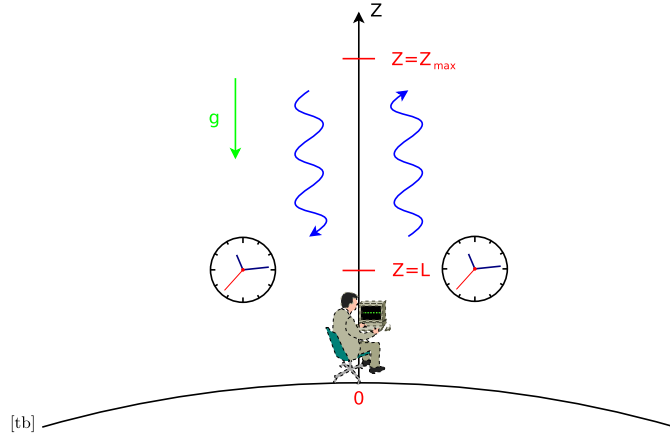
Fig. 14. Sketch of a quantum particle in a uniform gravitational field. The particle is sent upward, the initial time being recorded by a quantum Peres clock. The particle reaches a maximum altitude and then falls back to the detector at $z = L$ where the Peres clock measures this return time.

where $z = L$ is the initial altitude (at $t = 0$) and v_{ini} is the initial velocity, related to the energy E of the particle by the following expression

$$v_{\mathrm{ini}} = \sqrt{2g\frac{m_{\mathrm{grav}}}{m_{\mathrm{in}}}\left(\frac{E}{m_{\mathrm{grav}}g} - L\right)^{1/2}} \tag{835}$$

Of course, the motion of the particle only depends on the ratio $m_{\mathrm{grav}}/m_{\mathrm{in}}$. The maximum altitude, z_{max}, reached by the particle, see Fig. 14, is given by

$$z_{\mathrm{max}} = \frac{E}{m_{\mathrm{grav}}g} \tag{836}$$

which means that, from its initial position to z_{max}, the particle can rise a distance of $E/(m_{\mathrm{grav}}g) - L$. After having reached z_{max}, the particle falls back and the up-to-down classical time of flight reads

$$\Delta t = 2\sqrt{\frac{2}{g}}\left(\frac{m_{\mathrm{in}}}{m_{\mathrm{g}}}\right)^{1/2}\left(\frac{E}{m_{\mathrm{grav}}g} - L\right)^{1/2} \tag{837}$$

If $m_{\mathrm{grav}} = m_{\mathrm{in}}$, this equation can also be written as $\Delta t = 2v_{\mathrm{ini}}/g$.

Then, we would like to consider the same situation but from a quantum-mechanical point of view [140]. In particular, we would like to measure the quantum time of flight. Given the discussion of the previous section, it is clear that we must couple the quantum particle to a Peres clock located at $z = L$. Then, the next step is, of course, to solve the Schrödinger equation. Following the previous considerations, we just have to consider the equivalent of Eq. (810), but for the case of a uniform gravitational field. It reads

$$\frac{\mathrm{d}^2\psi_n}{\mathrm{d}z^2} + \left[\frac{2m_{\mathrm{in}}(E - n\hbar\omega)}{\hbar^2} - \frac{2m_{\mathrm{in}}m_{\mathrm{grav}}g}{\hbar^2}z\right]\psi_n = 0 \tag{838}$$

It is straightforward to obtain that the solution can be expressed in terms of Airy functions [48,49] Ai(z) and Bi(z). If we retain only the branch that leads to a bounded wave-function, the result is given by

$$\psi(z, \theta) = \frac{A}{\sqrt{N}}\sum_{n=-j}^{n=j}\mathrm{Ai}\left(\frac{z - b}{a}\right)u_n(\theta) \tag{839}$$

where A is constant fixed by the normalization of the wave-function and a and b are defined by

$$a = \left(\frac{\hbar^2}{2m_{\mathrm{in}}m_{\mathrm{grav}}g}\right)^{1/3} \tag{840}$$

$$b = \frac{E - n\hbar\omega}{m_{\mathrm{grav}}g} \tag{841}$$

The next step consists in evaluating the wave-function at $x = L$, with the assumption that the return point is located far from the maximum altitude. This means that $L - b$ is a negative quantity and that $|L - b|$ is large [more precisely, one has

$L - b = L - E/(m_{\text{grav}}g) + E\epsilon/(m_{\text{grav}}g)$ and the quantity $L - E/(m_{\text{grav}}g)$ is minus the distance the particle can rise, see the remark after Eq. (836). Therefore, $L - b < 0$ and, if this distance is large, $|L - b|$ is indeed a large quantity]. In this limit, one has

$$\psi(L, \theta) \simeq \frac{A}{2i\sqrt{\pi N}} \sum_{n=-j}^{n=j} \left(\frac{z-b}{a}\right)^{-1/4} \left(e^{i\xi + i\pi/4} - e^{-i\xi - i\pi/4}\right) u_n(\theta) \tag{842}$$

where $\xi \equiv (2/3)[(b-L)/a]^{3/2}$. Explicitly, this quantity can be expressed as

$$\xi = \frac{2}{3a^{3/2}} \left[\frac{E}{m_{\text{grav}}g}(1-\epsilon) - L\right]^{3/2} \tag{843}$$

$$\simeq \frac{2}{3a^{3/2}} \left(\frac{E}{m_{\text{grav}}g} - L\right)^{3/2} - \frac{1}{a^{3/2}} \left(\frac{E}{m_{\text{grav}}g} - L\right)^{1/2} \frac{n\hbar\omega}{m_{\text{grav}}g} \tag{844}$$

In order to calculate the difference between the two times indicated by the clock and corresponding to the two branches of the wave-function, it is in fact sufficient to estimate the part of the phase shift that depends on n. Therefore, it is clear that the phase shift $\pm\pi/4$ in the expression (842) will not contribute because it is n-independent. Moreover, the term $[(z-b)/a]^{-1/4}$ is n-dependent but is common to the two branches of the wave-functions. As a consequence, it will cancel out when the phase shift is determined (but would participate to the "absolute" time indicated by the clock when the particle leaves $z = L$ and returns to $z = L$). Summarizing, the up-and down time of flight is given by

$$\Delta t = \frac{2}{a^{3/2}} \left(\frac{E}{m_{\text{grav}}g} - L\right)^{1/2} \frac{\hbar}{m_{\text{grav}}g} \tag{845}$$

$$= 2\sqrt{\frac{2}{g}} \left(\frac{m_{\text{in}}}{m_{\text{grav}}}\right)^{1/2} \left(\frac{E}{m_{\text{grav}}g} - L\right)^{1/2} \tag{846}$$

that is to say exactly Eq. (837), see Ref. [140]. Several remarks are in order here. Firstly, although it is of quantum-mechanical origin, the above equation does not contain $\hbar$ that cancels out. Secondly, this result seems to indicate that the weak equivalence principle holds at the quantum level. Thirdly, it may appear surprising that the quantum Δt is not corrected by a term corresponding to the penetration depth, see Section 14.2.1, in particular Eq. (832), as it is the case for a square potential. The interpretation given in Ref. [140] is that there is a finite probability that the particle tunnels below the barrier but there is also a finite probability that the particle is scattered back before the classical turning point. If these two probabilities cancel then one obtains the classical prediction. Fourthly, the above result is valid only if the return time is measured far from the maximum altitude. What happens if, on the contrary, the particle starts from a position which is close to the turning point? To investigate this case, we re-start from Eq. (839) and write the solution as

$$\psi(z, \theta) = \frac{A}{\sqrt{N}} \sum_{n=-j}^{n=j} \frac{\sqrt{y}}{3} \Big\{ \big[e^{i\pi/3} J_{1/3}(\zeta) + e^{-i\pi/3} J_{-1/3}(\zeta)\big]$$

$$+ \big[\big(1 - e^{i\pi/3}\big) J_{1/3}(\zeta) + \big(1 - e^{-i\pi/3}\big) J_{-1/3}(\zeta)\big] \Big\} u_n(\theta) \tag{847}$$

where $y \equiv (b-z)/a > 0$ and $\zeta \equiv 2y^{3/2}/3$. In this way, we have identified the incident (first term in the squared bracket) and the reflected wave (second term in the square bracket). Indeed, using the asymptotic behavior of the Bessel functions at infinity, one can check that each branches give precisely $e^{\pm(i\xi + i\pi/4)}$ if $z = L$, see Eq. (842). If we now consider the limit $y \ll 1$, then one arrives at

$$\psi(z, \theta) = \frac{A}{\sqrt{N}} \sum_{n=-j}^{n=j} \Big\{ \Big[\frac{3^{-2/3}}{2\Gamma(2/3)} + \frac{3^{-4/3}}{2\Gamma(4/3)}y\Big] + i\sqrt{3}\Big[-\frac{3^{-2/3}}{2\Gamma(2/3)} + \frac{3^{-4/3}}{2\Gamma(4/3)}y\Big] + \Big[\frac{3^{-2/3}}{2\Gamma(2/3)} + \frac{3^{-4/3}}{2\Gamma(4/3)}y\Big]$$

$$- i\sqrt{3}\Big[-\frac{3^{-2/3}}{2\Gamma(2/3)} + \frac{3^{-4/3}}{2\Gamma(4/3)}y\Big] + \cdots \Big\} u_n(\theta) \tag{848}$$

the two first terms corresponding to the incident wave and the two last ones to the reflected wave. The two branches are of course complex conjugate to each other. Therefore, if we denotes by $\aleph$ the complex number corresponding to the two first terms in the above equation, then the wave-function takes the form

$$\psi(z, \theta) \simeq \frac{A}{\sqrt{N}} \sum_{n=-j}^{n=j} |\aleph| \big(e^{i\arg\aleph} + e^{-i\arg\aleph}\big) u_n(\theta) \tag{849}$$

where

$$\arg \aleph = \arctan\left[\sqrt{3}\frac{3^{2/3}\Gamma(2/3)y - 3^{4/3}\Gamma(4/3)}{3^{2/3}\Gamma(2/3)y + 3^{4/3}\Gamma(4/3)}\right] \tag{850}$$

The rest of the calculation proceeds as before. We first expand the phase in ϵ and, then, take the limit $E/(m_g g) - L \to 0$. This leads to

$$\arg \aleph \simeq -\frac{\pi}{3} - \frac{3^{5/6}}{6}\frac{\Gamma(2/3)}{\Gamma(4/3)}\frac{\hbar}{m_{\mathrm{grav}}ga}n\omega \tag{851}$$

As a consequence, one finds that the time of flight is given by

$$\Delta t \simeq 2\frac{3^{5/6}}{6}\frac{\Gamma(2/3)}{\Gamma(4/3)}2^{1/3}\left(\frac{m_{\mathrm{in}}}{m_{\mathrm{grav}}}\right)^{1/3}\left(\frac{\hbar}{m_{\mathrm{grav}}g^2}\right)^{1/3} \tag{852}$$

$$\simeq 1.59 \times \left(\frac{m_{\mathrm{in}}}{m_{\mathrm{grav}}}\right)^{1/3}\left(\frac{\hbar}{m_{\mathrm{grav}}g^2}\right)^{1/3} \tag{853}$$

in rough agreement with Ref. [140]. This time, we notice that $\hbar$ has not canceled in the final expression and that the time of flight is different from its classical counterpart (837). This also means, as already mentioned, that the cancellation which leads to a quantum time of flight equals to the classical one is in fact valid only if the particle starts far from the turning point.

We conclude that the weak equivalence principle (or the universality of the free fall) can be extended to the quantum regime. The linear potential seems to possess precisely the shape which leads to identical times of flight. It is important to realize that this property depends on the shape of the potential and would not be obtained with another shape, see Ref. [140]. Moreover, this seems to be true only far from the turning point. If one identifies the wave-vector to $k = (2m_{\mathrm{in}}m_{\mathrm{grav}}g/\hbar^2)^{1/3}$, see Ref. [151], the de Broglie relation $v \simeq \hbar k/m$ allows us to estimate a velocity. Then, from the quantum time of flight, see Eq. (853), we can construct a length $\simeq [\hbar^2/(m^2 g)]^{1/3}$ which typically controls the validity of the approximation just mentioned.

14.2.3. Atom trampoline

In this section, we want to briefly mention the case of the "atom trampoline" [150,151]. Let us consider again the experimental set up of the previous section but let us now assume that a reflecting wall has been installed on the ground, at $z = 0$. Then, the boundary condition $\psi_n(0) = 0$ leads to $\mathrm{Ai}(-b/a) = 0$. If z_n denotes the zero of the Airy function, this means that we now have a discrete spectrum of energy levels given by

$$E_n = -z_n\left(\frac{\hbar^2 m_{\mathrm{grav}}^2 g^2}{2m_{\mathrm{in}}}\right)^{1/3} \tag{854}$$

It is interesting to notice that this result depends on the ratio $m_{\mathrm{grav}}^2/m_{\mathrm{in}}$ and not on $m_{\mathrm{grav}}/m_{\mathrm{in}}$. Moreover, this spectrum has been observed (with ultra-cold neutrons which leads to E_n of the order of $\sim 10^{-12}$ eV) in the gravitational field of the Earth [152,153]. Therefore, this result confirms the discussion about the COW experiment described in Section 14.1. It experimentally establishes that the gravitational force "couples" to the Schrödinger equation in a standard way.

14.2.4. Schrödinger equation in an accelerated frame

It is also frequent to refer to the weak equivalence principle as the property stating that, locally, the effect of a constant gravitational field can be mimicked by an accelerating frame. Therefore, it is interesting to study whether this claim holds in quantum mechanics. This question has been studied in Refs. [146,150,151]. Here, we follow the treatment of Ref. [146].

Let us consider the free Schrödinger equation (in one dimension to simplify the problem). It reads

$$i\hbar\frac{\partial \Psi(t,z)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \Psi(t,z)}{\partial z^2} \tag{855}$$

Then, let us consider the following coordinate transformation

$$z' = z - vt - \frac{1}{2}gt^2 \tag{856}$$

$$t' = t \tag{857}$$

It is straightforward to show that the function $\varphi(t', z')$ defined according to

$$\Psi(t,z) = \varphi[t'(t), z'(t,z)]\exp\left[\frac{imv}{\hbar}\left(z' + \frac{vt'}{2}\right) + \frac{imgt'}{\hbar}\left(z' + \frac{vt'}{2} + \frac{gt'^2}{6}\right)\right] \tag{858}$$

satisfies the equation [146]

$$i\hbar\frac{\partial\varphi(t',z')}{\partial t'} = -\frac{\hbar^2}{2m}\frac{\partial^2\varphi(t',z')}{\partial z'^2} + mgz'\varphi(t',z') \tag{859}$$

The potential in the above equation is precisely of the form describing a constant gravitational field. In this sense, we find that the equivalence between a constant gravitational field and an accelerated frame propagates to quantum mechanics.

14.2.5. Falling composite quantum objects

In the previous sections, we have studied how a quantum particle falls in a uniform gravitational field. For instance, we have investigated under which circumstances one can say that the weak equivalence principle holds in quantum mechanics. Here, we would like to study the same question but in the case of a composite quantum object. In particular, it is interesting to ask whether the motion of the composite system (the atom) can depend on its internal state.

Let us start by recalling the Ehrenfest theorem [57] for a quantum particle. It states that the time derivatives of the mean position and momentum are given by (for simplicity, we consider the case of a one-dimensional system since we have in mind a particle falling down along the z-axis)

$$\frac{\mathrm{d}\langle z\rangle}{\mathrm{d}t} = \frac{\langle p_z\rangle}{m_{\mathrm{in}}} \tag{860}$$

$$\frac{\mathrm{d}\langle p_z\rangle}{\mathrm{d}t} = -\left\langle\frac{\partial V}{\partial z}\right\rangle \tag{861}$$

In the case of a linear potential, $V(z) = m_{\mathrm{grav}}gz$, the two previous equations expressing the Ehrenfest theorem can be combined in such a way that we arrive at

$$\frac{\mathrm{d}^2\langle z\rangle}{\mathrm{d}t^2} = -\frac{m_{\mathrm{grav}}}{m_{\mathrm{in}}}g \tag{862}$$

This means that the average position $\langle z\rangle(t)$ evolves exactly as the classical z given in Eq. (834). However, let us also notice that the variance of the position is not a function of the ratio $m_{\mathrm{grav}}/m_{\mathrm{in}}$ only, see in particular Refs. [146–148].

After this short reminder, let us consider a bound system of two particles interacting through a potential V which only depends on the distance between them. We also assume that this system is embedded in a region where there is an exterior field described by the potential $U(\boldsymbol{r})$. The corresponding Lagrangian reads

$$L(\dot{\boldsymbol{r}}_1, \dot{\boldsymbol{r}}_2, \boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{m_{\mathrm{in},1}}{2}\dot{\boldsymbol{r}}_1^2 + \frac{m_{\mathrm{in},2}}{2}\dot{\boldsymbol{r}}_2^2 - V(\boldsymbol{r}_1 - \boldsymbol{r}_2) - U(\boldsymbol{r}_1) - U(\boldsymbol{r}_2) \tag{863}$$

where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the positions of the two particles and $m_{\mathrm{in},1}$, $m_{\mathrm{in},2}$ their mass. If we now introduce the center of mass position vector,

$$\boldsymbol{r}_{\mathrm{G}} = \frac{m_{\mathrm{in},1}}{m_{\mathrm{in},1} + m_{\mathrm{in},2}}\boldsymbol{r}_1 + \frac{m_{\mathrm{in},2}}{m_{\mathrm{in},1} + m_{\mathrm{in},2}}\boldsymbol{r}_2 \tag{864}$$

and the relative position vector,

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2 \tag{865}$$

then Eq. (863) takes the form

$$L(\dot{\boldsymbol{r}}_{\mathrm{G}}, \dot{\boldsymbol{r}}, \boldsymbol{r}_{\mathrm{G}}, \boldsymbol{r}) = \frac{M}{2}\dot{\boldsymbol{r}}_{\mathrm{G}}^2 + \frac{\mu}{2}\dot{\boldsymbol{r}}^2 - V(\boldsymbol{r}) - U(\boldsymbol{r}_1) - U(\boldsymbol{r}_2) \tag{866}$$

where $M \equiv m_{\mathrm{in},1} + m_{\mathrm{in},2}$ is the total mass and $\mu \equiv m_{\mathrm{in},1}m_{\mathrm{in},2}/(m_{\mathrm{in},1} + m_{\mathrm{in},2})$ is the reduced one. Of course, in the above expression $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ must be viewed as function of $\boldsymbol{r}_{\mathrm{G}}$ and $\boldsymbol{r}$.[5] From the above considerations, it is straightforward to obtain the Hamiltonian of the system. We arrive at the following expression:

$$H = \frac{\boldsymbol{p}_{\mathrm{G}}^2}{2M} + \frac{\boldsymbol{p}^2}{2\mu} + V(\boldsymbol{r}) + U(\boldsymbol{r}_1) + U(\boldsymbol{r}_2) \tag{867}$$

where $\boldsymbol{p}_{\mathrm{G}} \equiv \boldsymbol{p}_1 + \boldsymbol{p}_2$ and $\boldsymbol{p} = (\mu/m_{\mathrm{in},1})\boldsymbol{p}_1 - (\mu/m_{\mathrm{in},2})\boldsymbol{p}_2$. If we now assume that the external field is a uniform gravitational field (directed along the z-axis), then it is easy to show that

---

[5] The relation between these quantities is given by

$$\boldsymbol{r}_1 = \boldsymbol{r}_{\mathrm{G}} + \frac{m_{\mathrm{in},2}}{M}\boldsymbol{r}$$

$$\boldsymbol{r}_2 = \boldsymbol{r}_{\mathrm{G}} - \frac{m_{\mathrm{in},1}}{M}\boldsymbol{r}$$

$$U(\boldsymbol{r}_1) + U(\boldsymbol{r}_2) = (m_{\mathrm{grav},1} + m_{\mathrm{grav},2})gz_{\mathrm{G}} + \frac{1}{M}(m_{\mathrm{grav},1}m_{\mathrm{in},2} - m_{\mathrm{grav},2}m_{\mathrm{in},1})gz \tag{868}$$

$$\equiv \overline{U}(z_{\mathrm{G}}) + \tilde{U}(z) \tag{869}$$

where, according to Eq. (865), $z \equiv z_1 - z_2$. As a consequence, the Hamiltonian (867) takes the form

$$H = H_{\mathrm{G}} + H_{\mathrm{rel}} \tag{870}$$

where the center of mass and relative Hamiltonians can be expressed as

$$H_{\mathrm{G}} = \frac{\boldsymbol{p}_{\mathrm{G}}^2}{2M} + \overline{U}(z_{\mathrm{G}}) \tag{871}$$

$$H_{\mathrm{rel}} = \frac{\boldsymbol{p}^2}{2\mu} + V(\boldsymbol{r}) + \tilde{U}(z) \tag{872}$$

The physical interpretation of these equations is clear. The center of mass behaves as a free particle embedded in a uniform gravitational field. On the other hand, the internal states of the composite system are determined by the potential $V(\boldsymbol{r})$ but are also affected by the gravitational field through the term $\tilde{U}(z)$. In practice, one expects that $\tilde{U}(z) \ll V(\boldsymbol{r})$ and, therefore, the perturbations of the energy levels will be either very small or even totally negligible. Moreover, because $[z, p_{\mathrm{G},z}] = [z_{\mathrm{G}}, p_z] = 0$, the two above Hamiltonians commute

$$[H_{\mathrm{G}}, H_{\mathrm{rel}}] = 0 \tag{873}$$

which implies that the total state space is in fact the tensorial product of the center of mass state space and of the relative state space.

Let us now study the behavior of $\langle z_{\mathrm{G}} \rangle$. One has

$$\frac{\mathrm{d}\langle z_{\mathrm{G}} \rangle}{\mathrm{d}t} = \frac{1}{i\hbar}\langle [z_{\mathrm{G}}, H_{\mathrm{G}} + H_{\mathrm{rel}}] \rangle = \frac{1}{i\hbar}\langle [z_{\mathrm{G}}, H_{\mathrm{G}}] \rangle = \frac{\langle p_{\mathrm{G},z} \rangle}{M} \tag{874}$$

In the same manner, one can now calculate the evolution of $\langle p_{\mathrm{G},z} \rangle$. One obtains

$$\frac{\mathrm{d}\langle p_{\mathrm{G},z} \rangle}{\mathrm{d}t} = \frac{1}{i\hbar}\langle [p_{\mathrm{G},z}, H_{\mathrm{G}} + H_{\mathrm{rel}}] \rangle = \frac{1}{i\hbar}\langle [p_{\mathrm{G},z}, H_{\mathrm{G}}] \rangle$$

$$= \frac{1}{i\hbar}\langle [p_{\mathrm{G},z}, \overline{U}(z_{\mathrm{G}})] \rangle = -(m_{\mathrm{grav},1} + m_{\mathrm{grav},2})g \tag{875}$$

Therefore, combining Eqs. (874) and (875), we arrive at the following expression

$$\frac{\mathrm{d}^2\langle z_{\mathrm{G}} \rangle}{\mathrm{d}t^2} = -\frac{m_{\mathrm{grav},1} + m_{\mathrm{grav},2}}{m_{\mathrm{in},1} + m_{\mathrm{in},2}}g \tag{876}$$

Of course, this equation is similar to Eq. (862) and this implies that the center of mass mean value of the quantum composite system will follow the classical trajectory. In particular, this means that $\langle z_{\mathrm{G}} \rangle$ does not depend in which internal energy state the system is placed. Therefore, if, for instance, we consider a falling Hydrogen atom, the behavior of the position mean value of the atom will not depend of whether the atom is in the state, say, $1s$ or $2p$. We also notice that the coefficient in front of g in Eq. (876) is the ratio of the total gravitational mass to the total inertial mass. In this sense, the weak equivalence principle is also satisfied by the composite quantum systems.

Finally, it is also interesting to repeat the gedanken experiment in Fig. 14 but with a composite system. It seems reasonable to couple the Peres clock to the center of mass of the system which means that the Hamiltonian of the system is still given by

$$H = H_{\mathrm{G}} + H_{\mathrm{rel}} \tag{877}$$

where, now, the center of mass and relative Hamiltonians can be expressed as

$$H_{\mathrm{G}} = \frac{\boldsymbol{p}_{\mathrm{G}}^2}{2M} + \overline{U}(z_{\mathrm{G}}) + \mathcal{P}(z_{\mathrm{G}})H_{\mathrm{clock}} \tag{878}$$

$$H_{\mathrm{rel}} = \frac{\boldsymbol{p}^2}{2\mu} + V(\boldsymbol{r}) + \tilde{U}(z) \tag{879}$$

Then, following Eq. (809), we write the wave-function of the system as

$$\Psi(t, \boldsymbol{r}_{\mathrm{G}}, \boldsymbol{r}, \theta) = \frac{1}{\sqrt{N}} \sum_{n=-j}^{n=j} \psi_n(\boldsymbol{r}_{\mathrm{G}})\chi_n(\boldsymbol{r})u_n(\theta)\mathrm{e}^{-iEt/\hbar} \tag{880}$$

The next step consists in inserting this wave-function in the Schrödinger equation in order to obtain the equivalent of Eq. (810). The result can be written as

$$\sum_{n=-j}^{n=j} \left\{ \chi_n(\mathbf{r}) \left[-\frac{\hbar^2}{2M} \Delta_G \psi_n(\mathbf{r}_G) + \overline{U}(z_G)\psi_n(\mathbf{r}_G) + n\hbar\omega\mathcal{P}(z_G)\psi_n(\mathbf{r}_G) - E_G\psi_n(\mathbf{r}_G) \right] \right.$$

$$\left. + \psi_n(\mathbf{r}_G) \left[-\frac{\hbar^2}{2m} \Delta \chi_n(\mathbf{r}) + \tilde{U}(z)\chi_n(\mathbf{r}) - E_{\text{rel}}\chi_n(\mathbf{r}) \right] \right\} u_n(\theta) = 0 \tag{881}$$

where $E = E_G + E_{\text{rel}}$. This implies that only the center of mass part, $\psi_n(\mathbf{r}_G)$, of the total wave-function will depend on n. As a consequence, the internal part should in fact be written as $\chi(\mathbf{r})$ instead of $\chi_n(\mathbf{r})$. This means that $\chi(\mathbf{r})$ cannot cause a rotation of the Peres clock. Therefore, the time indicated by the clock will not depend on the internal state of the composite system. We conclude that the time of flight of an atom, defined and calculated in terms of the Peres clock, does not depend in which quantum state this atom is placed.

This concludes the section on the weak equivalence principle in quantum mechanics. The result is mixed. We have shown that, very often, it is satisfied in the framework of quantum mechanics. But, on the other hand, it seems sometimes to be modified by quantum effects, see in particular Eq. (853). Maybe this indicates that its application to the more complicated case of quantum vacuum fluctuations of a field is rather suspicious?

15. Conclusions

We are now in a position where we can really formulate the cosmological constant problem. Vacuum fluctuations seem to exist in Nature and to have normal gravitational properties. Therefore, it seems natural to postulate that they participate in the value of the cosmological constant. However, when one tries to use well controlled techniques of quantum field theory to compute their energy density, one obtains a number which seems to be in contradiction with the measurement of Λ that one obtains in cosmology and with the constraints that one deduces from (for instance) the motion of the planets in our solar system. The difficulty of the problem stems from the fact that, in order to remove the above contradiction, one necessarily has to abandon something that is considered as robust, i.e. renormalization in quantum field theory, weak equivalence principle, high accuracy measurements of the expansion of the Universe, etc.

Even if this is not the purpose of this article, let us conclude this review with a few words about the solutions that have been proposed to solve the cosmological constant problem, for a complete overview of the subject, see for instance Ref. [154]. Of course, the most obvious solution is super-symmetry [19,74,75]. However, as we have discussed in Section 7.4, super-symmetry has to be broken and this destroys the "miraculous" cancellation of the various terms participating the vacuum energy. In fact, the only common point between these various contributions is gravity. Therefore, this suggests that the mechanism that cancels the cosmological constant makes use of gravity. This idea has been explored in the so-called adjustment mechanisms [155–157]. These models work reasonably well but, unfortunately, they suffer from a severe disease. Indeed, in these scenarios, the Newton constant is time-dependent and it turns out that the variation of the gravitational coupling is too strong to be compatible with the known constraints on $|\dot{G}/G|$. Therefore, although the idea is very attractive at first sight, it seems that it does not yet exist a realistic realization of it. It is also worth signaling that back-reaction mechanisms have also been studied in Refs. [158,159]. Here, the idea is that the long wavelength (super-Hubble) cosmological perturbations are described by a stress–energy tensor that exactly cancels the cosmological constant.

Another class of solution is based on quantum gravity and quantum cosmology [160–162]. It is argued that the no-boundary Hartle–Hawking wave-function is such that it peaks at $\Lambda = 0$. However, this approach suffers from various limitations among which is the fact that the path integral is not properly defined and that probabilities are not positive definite (this comes from the fact that the Wheeler–De Witt equation is in fact similar to a Klein–Gordon equation. The positivity of probabilities in quantum cosmology can be only be restored in the WKB approximation).

Recently, following the pioneered work of Ref. [163], there have been many attempts to solve the cosmological constant problem in the framework of extra dimensions, see Chapter 7 of Ref. [154] or Ref. [164]. Another popular solution, also based on high-energy physics considerations, is the landscape approach [12] in string theory [165]. This approach assumes that all the pocket universes of the landscape are populated during eternal inflation. Combining this fact with the anthropic principle [166], it is argued that the most probable value of Λ is approximately the value observed today. This solution suffers from the difficulty of defining a measure on the landscape [167] and the use of the anthropic principle in this context has been criticized in Refs. [168,169].

Finally, another possibility that has recently been investigated is to modify the Einstein equations such that they become blind to a stress–energy tensor of the form $T_{\mu\nu} \propto g_{\mu\nu}$, see Refs. [170–172].

In order to be complete, let us mention that models of analogue gravity have also been developed to study the vacuum energy problem [173–175]. In these models, one can calculate the equivalent of the cosmological constant and study precisely its origin since the physics controlling the microscopic constituents is explicitly known. These models are based on many different types of analogue physical systems such as condensed matter systems [176], super-fluid Helium [177], Fermi liquid [178], Bose–Einstein condensates [179,180], etc. Then, one can show that a simple calculation of the ground state may lead to an incorrect result, the only possibility in order to reach the correct answer being to fully take into account the

microscopic theory. This would indicate that a correct calculation of the cosmological constant must necessarily be based on the fine structure of space–time, i.e. probably on a consistent theory of quantum gravity.

As the variety of the subjects studied is this review shows, the cosmological constant problem is a very rich one. Given its difficulty, it is probable that it will have to wait for a new theory beyond the current standard model to find its resolution. This theory will have to describe the gravitational properties of the vacuum fluctuations and to regulate the ultra-violet infinities that appear in the calculation of vacuum energy density. This list of requirements almost constitutes an identity card for a theory of quantum gravity. But the most interesting aspect maybe goes the other way round. Given the fact that cosmology has enabled us to grasp observational signatures related to the vacuum energy, maybe these experimental results will help us to deduce and to establish a convincing theory of quantum gravity? This is the reason why the cosmological constant problem appears to be so important and so interesting.

Acknowledgements

References

[1] S. Weinberg, Rev. Mod. Phys. 61 (1989) 1.
[2] S.M. Carroll, W.H. Press, E.L. Turner, Ann. Rev. Astron. Astrophys. 30 (1992) 499.
[3] A.D. Dolgov, astro-ph/9708045, 1997.
[4] V. Sahni, A.A. Starobinsky, Int. J. Mod. Phys. D 9 (2000) 373, astro-ph/9904398.
[5] N. Straumann, Eur. J. Phys. 20 (1999) 419, astro-ph/9908342.
[6] S. Weinberg, astro-ph/0005265, 2000.
[7] S.M. Carroll, Living Rev. Rel. 4 (2001) 1, astro-ph/0004075.
[8] S.E. Rugh, H. Zinkernagel, Stud. Hist. Philos. Mod. Phys. (2000), hep-th/0012253.
[9] T. Padmanabhan, Phys. Rept. 380 (2003) 235, hep-th/0212290.
[10] J. Yokoyama, gr-qc/0305068, 2003.
[11] S. Sarkar, Special section of current science on the legacy of Albert Einstein, Curr. Sci. 88 (2005) 2120.
[12] J. Polchinski, hep-th/0603249, 2006, pp. 216–236.
[13] E.J. Copeland, M. Sami, S. Tsujikawa, Int. J. Mod. Phys. D 15 (2006) 1753, hep-th/0603057.
[14] P. Brax, arXiv:0912.3610, 2009.
[15] D. Sapone, Int. J. Mod. Phys. A 25 (2010) 5253, arXiv:1006.5694.
[16] M.E. Peskin, D.V. Schroeder, An Introduction to Quantum Field Theory, Addison-Wesley Publishing Company, 1995.
[17] W. Greiner, J. Reinhardt, Field Quantization, Springer, Berlin, Germany, 1996.
[18] C. Itzykson, J.B. Zuber, Quantum Field Theory, International Series in Pure and Applied Physics, McGraw-Hill, 1980.
[19] D. Bailin, A. Love, Supersymmetric Gauge Field Theory and String Theory, Graduate Student Series in Physics, IOP, 1994.
[20] M. Le Bellac, Quantum and Statistical Field Theory, Oxford, 1991.
[21] L.H. Ryder, Quantum Field Theory, Cambridge Univ. Press, 1985.
[22] F. Mandl, G. Shaw, Quantum Field Theory, Wiley-Interscience Publication, 1984.
[23] S. Perlmutter, et al., Supernova Cosmology Project, Astrophys. J. 517 (1999) 565, astro-ph/9812133.
[24] A.G. Riess, et al., Supernova Search Team, Astron. J. 116 (1998) 1009, astro-ph/9805201.
[25] A.A. Starobinsky, Phys. Lett. B 91 (1980) 99.
[26] A.H. Guth, Phys. Rev. D 23 (1981) 347.
[27] V.F. Mukhanov, G. Chibisov, JETP Lett. 33 (1981) 532.
[28] A.D. Linde, Phys. Lett. B 108 (1982) 389.
[29] V.F. Mukhanov, G. Chibisov, Sov. Phys. JETP 56 (1982) 258.
[30] A.A. Starobinsky, Phys. Lett. B 117 (1982) 175.
[31] A.H. Guth, S.Y. Pi, Phys. Rev. Lett. 49 (1982) 1110.
[32] S. Hawking, Phys. Lett. B 115 (1982) 295, revised version.
[33] A. Albrecht, P.J. Steinhardt, Phys. Rev. Lett. 48 (1982) 1220.
[34] A.D. Linde, Phys. Lett. B 129 (1983) 177.
[35] J. Martin, Braz. J. Phys. 34 (2004) 1307, astro-ph/0312492.
[36] J. Martin, Lect. Notes Phys. 669 (2005) 199, hep-th/0406011.
[37] J. Martin, Lect. Notes Phys. 738 (2008) 193, arXiv:0704.3540.
[38] G.F. Smoot, C. Bennett, A. Kogut, E. Wright, J. Aymon, et al., Astrophys. J. 396 (1992) L1.
[39] J. Martin, C. Ringeval, JCAP 0608 (2006) 009, astro-ph/0605367.
[40] P. Astier, R. Pain, Observational evidence of the accelerated expansion of the Universe, C. R. Physique 13 (2012) 521–538, in this issue; http://dx.doi.org/10.1016/j.crhy.2012.04.009.
[41] M. Kunz, The phenomenological approach to modeling the dark energy, C. R. Physique 13 (2012) 539–565, in this issue; http://dx.doi.org/10.1016/j.crhy.2012.04.007.
[42] C. Clarkson, Establishing homogeneity of the universe in the shadow of dark energy, C. R. Physique 13 (2012) 682–718, in this issue; http://dx.doi.org/10.1016/j.crhy.2012.04.005.
[43] C. de Rham, Galileons in the sky, C. R. Physique 13 (2012) 666–681, in this issue; http://dx.doi.org/10.1016/j.crhy.2012.04.006.
[44] A. Sakharov, Sov. Phys. Dokl. 12 (1968) 1040.
[45] J. Kapusta, C. Gale, Finite-Temperature Field Theory: Principles and Applications, Cambridge Univ. Press, 2006.
[46] S. Chatrchyan, et al., CMS Collaboration, arXiv:1202.1488, 2012.

[47] E.K. Akhmedov, hep-th/0204048, 2002.
[48] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, ninth ed., National Bureau of Standards, Washington, US, 1970.
[49] I.S. Gradshteyn, I.M. Ryzhik, Table of Integrals, Series, and Products, Academic Press, New York and London, 1965.
[50] G. Ossola, A. Sirlin, Eur. Phys. J. C 31 (2003) 165, hep-ph/0305050, 28 pages, 5 figures.
[51] J.F. Koksma, T. Prokopec, arXiv:1105.6296, 2011.
[52] L. Rosenberg, Phys. Rev. 129 (1963) 2786.
[53] N.D. Birrell, P.C.W. Davies, Quantum Fields in Curved Space, Cambridge Univ. Press, 1982.
[54] C.M. Bender, T.T. Wu, Phys. Rev. 184 (1969) 1231.
[55] T. Abbott, Lecture Notes (2006), web.mit.edu/tabbott/Public/8.06/current.pdf.
[56] J. Minahan, MIT Lecture Notes (2011), http://www.teorfys.uu.se/people/minahan/Courses/QFT/wholeshebang.pdf.
[57] C. Cohen-Tannoudji, B. Diu, F. Laloe, Quantum Mechanics, Wiley-Interscience, 2006.
[58] P.M. Stevenson, Phys. Rev. D 30 (1984) 1712.
[59] P.M. Stevenson, Phys. Rev. D 32 (1985) 1389.
[60] P.M. Stevenson, I. Roditi, Phys. Rev. D 33 (1986) 2305.
[61] P.M. Stevenson, Z. Phys. C 35 (1987) 467.
[62] P.M. Stevenson, B. Alles, R. Tarrach, Phys. Rev. D 35 (1987) 2407.
[63] P.M. Stevenson, R. Tarrach, Phys. Lett. B 176 (1986) 436.
[64] G. Hajj, P.M. Stevenson, Phys. Rev. D 37 (1988) 413.
[65] P.M. Stevenson, G. Hajj, J. Reed, Phys. Rev. D 34 (1986) 3117.
[66] R. Ibanez-Meier, I. Stancu, P.M. Stevenson, Z. Phys. C 70 (1996) 307, hep-ph/9207276.
[67] S.R. Coleman, E.J. Weinberg, Phys. Rev. D 7 (1973) 1888.
[68] E.J. Weinberg, A.-q. Wu, Phys. Rev. D 36 (1987) 2474.
[69] T. Curtright, C.B. Thorn, J. Math. Phys. 25 (1984) 541.
[70] J. Wess, B. Zumino, Nucl. Phys. B 70 (1974) 39.
[71] S.R. Coleman, J. Wess, B. Zumino, Phys. Rev. 177 (1969) 2239.
[72] J. Callan, G. Curtis, S.R. Coleman, J. Wess, B. Zumino, Phys. Rev. 177 (1969) 2247.
[73] J. Wess, B. Zumino, Phys. Lett. B 49 (1974) 52.
[74] I.J. Aitchison, Notes of Lectures for Graduate Students in Particle Physics, Oxford, 2004 and 2005, hep-ph/0505105.
[75] P. Binetruy, Supersymmetry: Theory, Experiment and Cosmology, Oxford Univ. Press, 2006.
[76] L. Parker, in: Cargese 1978, Proceedings, Recent Developments in Gravitation, 1978, pp. 219–273.
[77] T.S. Bunch, L. Parker, Phys. Rev. D 20 (1979) 2499.
[78] E. Poisson, A. Pound, I. Vega, Living Rev. Rel. 14 (7) (2011), arXiv:1102.0529.
[79] R. Scranton, et al., SDSS, astro-ph/0307335, 2003.
[80] P. Solevi, R. Mainini, S.A. Bonometto, astro-ph/0412054, 2004.
[81] D.N. Spergel, et al., WMAP, Astrophys. J. Suppl. 170 (2007) 377, astro-ph/0603449.
[82] M. Tegmark, et al., SDSS, Phys. Rev. D 69 (2004) 103501, astro-ph/0310723.
[83] E. Kolb, M. Turner, The Early Universe, Frontiers in Physics Series, vol. 69, Addison-Wesley Publishing Company, 1990.
[84] B. Ratra, P.J.E. Peebles, Phys. Rev. D 37 (1988) 3406.
[85] P. Brax, J. Martin, Phys. Lett. B 468 (1999) 40, astro-ph/9905040.
[86] P. Brax, J. Martin, Phys. Rev. D 61 (2000) 103502, astro-ph/9912046.
[87] P. Brax, J. Martin, A. Riazuelo, Phys. Rev. D 62 (2000) 103505, astro-ph/0005428.
[88] P. Brax, J. Martin, A. Riazuelo, Phys. Rev. D 64 (2001) 083505, hep-ph/0104240.
[89] J. Martin, M.A. Musso, Phys. Rev. D 71 (2005) 063514, astro-ph/0410190.
[90] J. Martin, C. Schimd, J.-P. Uzan, Phys. Rev. Lett. 96 (2006) 061303, astro-ph/0510208.
[91] P. Brax, J. Martin, Phys. Lett. B 647 (2007) 320, hep-th/0612208.
[92] P. Brax, J. Martin, JCAP 0611 (2006) 008, astro-ph/0606306.
[93] P. Brax, J. Martin, Phys. Rev. D 75 (2007) 083507, hep-th/0605228.
[94] P. Brax, C. van de Bruck, J. Martin, A.-C. Davis, JCAP 0909 (2009) 032, arXiv:0904.3471.
[95] C.F. Kolda, D.H. Lyth, Phys. Lett. B 458 (1999) 197, hep-ph/9811375.
[96] S.M. Carroll, Phys. Rev. Lett. 81 (1998) 3067, astro-ph/9806099.
[97] M. Doran, J. Jaeckel, astro-ph/0205206, 2002.
[98] M. Doran, J. Jaeckel, Lect. Notes Phys. 646 (2004) 273.
[99] M. Garny, Phys. Rev. D 74 (2006) 043009, hep-ph/0606120.
[100] C. Ringeval, T. Suyama, T. Takahashi, M. Yamaguchi, S. Yokoyama, Phys. Rev. Lett. 105 (2010) 121301, arXiv:1006.0368.
[101] I. Maor, R. Brustein, P.J. Steinhardt, Phys. Rev. Lett. 86 (2001) 6, astro-ph/0007297.
[102] S. Weinberg, Gravitation and Cosmology, John Wiley & Sons, 1972.
[103] E.L. Wright, astro-ph/9805292, 1998.
[104] A.W. Kerr, J.C. Hauck, B. Mashhoon, Class. Quant. Grav. 20 (2003) 2727, gr-qc/0301057.
[105] P. Jetzer, arXiv:0711.3723, 2007.
[106] L. Parker, Phys. Rev. D 22 (1980) 1922.
[107] L. Parker, L.O. Pimentel, Phys. Rev. D 25 (1982) 3180.
[108] L. Parker, Phys. Rev. D 24 (1981) 535.
[109] S. Moradi, E. Aboualizadeh, Gen. Rel. Grav. 42 (2010) 435.
[110] E. Fischbach, B.S. Freeman, W.-K. Cheng, Phys. Rev. D 23 (1981) 2157.
[111] N. Boulanger, P. Spindel, F. Buisseret, Phys. Rev. D 74 (2006) 125014, hep-th/0610207.
[112] T.A. Welton, Phys. Rev. 74 (1948) 1157.
[113] G. Plunien, B. Muller, W. Greiner, Phys. Rept. 134 (1986) 87.
[114] K.A. Milton, J. Phys. A 37 (2004) R209, hep-th/0406024.
[115] K.A. Milton, hep-th/0009173, 2000, pp. 333–349.
[116] K.A. Milton, Lect. Notes Phys. 834 (2011) 39, arXiv:1005.0031.
[117] A. Lambrecht, S. Reynaud, arXiv:1112.1301, 2011.
[118] K.A. Milton, Phys. Rev. D 68 (2003) 065020, hep-th/0210081.
[119] D. Deutsch, P. Candelas, Phys. Rev. D 20 (1979) 3063.

[120] P. Milonni, M.-L. Shih, Phys. Rev. A 45 (1992) 4241.
[121] R. Jaffe, Phys. Rev. D 72 (2005) 021301, hep-th/0503158.
[122] M. Chernodub, arXiv:1203.6588, 2012.
[123] E. Calloni, L. Di Fiore, G. Esposito, L. Milano, L. Rosa, Phys. Lett. A 297 (2002) 328, quant-ph/0109091.
[124] G. Bimonte, E. Calloni, G. Esposito, L. Rosa, Phys. Rev. D 76 (2007) 025008, hep-th/0703062.
[125] S.A. Fulling, K.A. Milton, P. Parashar, A. Romeo, K. Shajesh, et al., Phys. Rev. D 76 (2007) 025004, hep-th/0702091.
[126] K.A. Milton, S.A. Fulling, P. Parashar, A. Romeo, K. Shajesh, et al., J. Phys. A 41 (2008) 164052, arXiv:0710.3841.
[127] K. Milton, P. Parashar, J. Wagner, K. Shajesh, A. Romeo, et al., arXiv:0810.0081, 2008.
[128] R. Caldwell, astro-ph/0209312, 2002.
[129] F. Sorge, Class. Quant. Grav. 22 (2005) 5109.
[130] E. Masso, Phys. Lett. B 679 (2009) 433, arXiv:0902.4318.
[131] Y. Su, B.R. Heckel, E. Adelberger, J. Gundlach, M. Harris, et al., Phys. Rev. D 50 (1994) 3614.
[132] T. Damour, Space Sci. Rev. (2009), arXiv:0906.3174.
[133] R. Colella, A. Overhauser, S. Werner, Phys. Rev. Lett. 34 (1975) 1472.
[134] D.M. Greenberger, A. Overhauser, Rev. Mod. Phys. 51 (1979) 43.
[135] D.M. Greenberger, Rev. Mod. Phys. 55 (1983) 875.
[136] H. Salecker, E. Wigner, Phys. Rev. 109 (1958) 571.
[137] A. Peres, Am. J. Phys. 48 (1980) 552.
[138] C. Leavens, Solid Sate Commun. 86 (1993) 781.
[139] P. Davies, Class. Quant. Grav. 21 (2004) 5677.
[140] P.C.W. Davies, Class. Quant. Grav. 21 (2004) 2761, arXiv:quant-ph/0403027.
[141] N. Grot, C. Rovelli, R.S. Tate, Phys. Rev. A 54 (1996) 4679, quant-ph/9603021.
[142] Y. Aharonov, J. Oppenheim, S. Popescu, B. Reznik, W. Unruh, Phys. Rev. A 57 (1998) 4130, quant-ph/9709031.
[143] J. Oppenheim, B. Reznik, W. Unruh, Phys. Rev. A 59 (1999) 1804, quant-ph/9807043.
[144] P.C. Davies, Am. J. Phys. 73 (2005) 23.
[145] B. Mielnik, G. Torres-Vega, arXiv:1112.4198, 2011, e-prints.
[146] L. Viola, R. Onofrio, Phys. Rev. D 55 (1997) 455, quant-ph/9612039.
[147] M. Ali, A. Majumdar, D. Home, A.K. Pan, Class. Quant. Grav. 23 (2006) 6493, quant-ph/0606183.
[148] P. Chowdhury, D. Home, A. Majumdar, S. Mousavi, M. Mozaffari, et al., Class. Quant. Grav. 29 (2012) 025010.
[149] R. Onofrio, Int. J. Mod. Phys. A 25 (2010) 2260, arXiv:1101.1484.
[150] D. Giulini, arXiv:1105.0749, 2011.
[151] E. Kajari, N.L. Harshman, E.M. Rasel, S. Stenholm, G. Süßmann, W.P. Schleich, Appl. Phys. B 100 (2010) 43, arXiv:1006.1988.
[152] V. Nesvizhevsky, H. Borner, A. Gagarski, A. Petoukhov, G. Petrov, et al., Phys. Rev. D 67 (2003) 102002, hep-ph/0306198.
[153] H. Abele, S. Baessler, H. Borner, A. Gagarsky, V. Nesvizhevsky, et al., AIP Conf. Proc. 842 (2006) 793.
[154] S. Nobbenhuis, Ph.D. thesis, 2006, gr-qc/0609011 (Advisor: G. 't Hooft).
[155] A. Dolgov, in: G. Gibbons, S.W. Hawking, S.T. Tiklos (Eds.), The Very Early Universe, Cambridge University Press, Cambridge, England, 1982.
[156] L. Ford, Phys. Rev. D 35 (1987) 2339.
[157] A.D. Dolgov, F. Urban, Phys. Rev. D 77 (2008) 083503, arXiv:0801.3090.
[158] L.W. Abramo, R.H. Brandenberger, V.F. Mukhanov, Phys. Rev. D 56 (1997) 3248, gr-qc/9704037.
[159] V.F. Mukhanov, L.W. Abramo, R.H. Brandenberger, Phys. Rev. Lett. 78 (1997) 1624, gr-qc/9609026.
[160] J. Hartle, S. Hawking, Phys. Rev. D 28 (1983) 2960.
[161] E. Baum, Phys. Lett. B 133 (1983) 185.
[162] S.R. Coleman, Nucl. Phys. B 310 (1988) 643.
[163] V. Rubakov, M. Shaposhnikov, Phys. Lett. B 125 (1983) 139.
[164] C. Burgess, Annals Phys. 313 (2004) 283, paper has been updated to include developments since the conference, hep-th/0402200.
[165] B. Zwiebach, A First Course in String Theory, Cambridge Univ. Press, 2004.
[166] S. Weinberg, Phys. Rev. Lett. 59 (1987) 2607.
[167] A. Linde, M. Noorbala, JCAP 1009 (2010) 008, arXiv:1006.2170.
[168] G.D. Starkman, R. Trotta, Phys. Rev. Lett. 97 (2006) 201301, astro-ph/0607227.
[169] R. Trotta, G.D. Starkman, AIP Conf. Proc. 878 (2006) 323, astro-ph/0610330.
[170] N. Afshordi, arXiv:0807.2639, 2008.
[171] G.F. Ellis, H. van Elst, J. Murugan, J.-P. Uzan, Class. Quant. Grav. 28 (2011) 225007, arXiv:1008.1196.
[172] S. Aslanbeigi, G. Robbers, B.Z. Foster, K. Kohri, N. Afshordi, Phys. Rev. D 84 (2011) 103522, arXiv:1106.3955.
[173] G. Volovik, Phys. Rept. 351 (2001) 195, gr-qc/0005091.
[174] G. Volovik, Annalen Phys. 14 (2005) 165, gr-qc/0405012.
[175] L. Sindoni, arXiv:1110.0686, 2011, invited review for SIGMA Special Issue 'Loop Quantum Gravity and Cosmology', 54 pages.
[176] G. Volovik, gr-qc/0101111, 2001.
[177] G. Volovik, Int. Ser. Monogr. Phys. 117 (2006) 1.
[178] S.H. Alexander, G. Calcagni, Found. Phys. 38 (2008) 1148, arXiv:0807.0225.
[179] S. Finazzi, S. Liberati, L. Sindoni, Phys. Rev. Lett. 108 (2012) 071101, arXiv:1103.4841, 5 pages, 1 figures.
[180] S. Finazzi, S. Liberati, L. Sindoni, arXiv:1204.3039, 2012.