OPEN ACCESS
IOP Publishing
Journal of Physics G: Nuclear and Particle Physics
J. Phys. G: Nucl. Part. Phys. **49** (2022) 103001 (100pp)
https://doi.org/10.1088/1361-6471/ac7f24

Topical Review

False vacuum decay: an introductory review

Federica Devoto[1], Simone Devoto[2], Luca Di Luzio[3] and Giovanni Ridolfi[4,*]

[1] Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, United Kingdom
[2] Dipartimento di Fisica Aldo Pontremoli, Università di Milano, INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy
[3] Dipartimento di Fisica e Astronomia Galileo Galilei, Università di Padova, INFN, Sezione di Padova, Via Marzolo 8, I-35131 Padova, Italy
[4] Dipartimento di Fisica, Università di Genova, INFN, Sezione di Genova, Via Dodecaneso 33, I-16146 Genova, Italy

E-mail: giovanni.ridolfi@ge.infn.it

Received 9 May 2022, revised 23 June 2022
Accepted for publication 7 July 2022
Published 25 August 2022

Abstract
We review the description of tunnelling phenomena in the semi-classical approximation in ordinary quantum mechanics and in quantum field theory. In particular, we describe in detail the calculation, up to the first quantum corrections, of the decay probability per unit time of a metastable ground state. We apply the relevant formalism to the case of the standard model of electroweak interactions, whose ground state is metastable for sufficiently large values of the top quark mass. Finally, we discuss the impact of gravitational interactions on the calculation of the tunnelling rate.

Keywords: tunnelling, standard model, quantum field theory, vacuum decay

(Some figures may appear in colour only in the online journal)

* Author to whom any correspondence should be addressed.

Contents

1. Introduction

Tunnelling in quantum field theory is a fascinating subject. In many respects, it differs substantially from the analogous phenomenon in ordinary quantum mechanics with a finite number of degrees of freedom, and deserves a special treatment. In a couple of famous papers [1, 2], Coleman and Callan have shown that the semi-classical approximation, in conjunction with path integral techniques, provides a suited context to deal with quantum tunnelling with infinite degrees of freedom. The first purpose of this work is to review both the semi-classical approximation in quantum mechanics and its extension to quantum field theory, as illustrated in references [1, 2]. We will adopt a pedagogical attitude: special attention will be devoted to the derivation of the bounce formalism, the definition and calculation of functional determinants, and the role of renormalization. All steps in the derivation of these known results, some of which are omitted in the original literature, are carefully illustrated.

A remarkable feature of tunnelling in quantum field theory is that the transition from the false to the true vacuum does not take place between two spatially homogeneous field configurations, but rather through the formation of a bubble of true vacuum in a false-vacuum background. Such field configuration (the bounce) is therefore not spatially homogenous, and the gradient term in the potential energy gives a non-zero contribution to the full potential energy. This is different from the case of ordinary quantum mechanics, in which the gradient term is absent and the tunnelling rate does not depend on the physics beyond the potential barrier.

An important application of the formalism above is the study of the stability of the electroweak vacuum in the case of the standard model of electroweak interactions. The accurate measurements of the top quark and Higgs boson masses indicate that the electroweak vacuum, which determines the observed spectrum of physical particles, is not the absolute ground state of the electroweak theory. Extrapolating the Higgs effective potential at high energies, the electroweak minimum becomes metastable for Higgs field values around 10^{10} GeV. Many efforts have been devoted to establish whether the lifetime metastable vacuum of the standard model is sufficiently larger that the age of the Universe, and therefore compatible with observations, or else if physics beyond the standard model is needed in order to justify the observed electroweak spectrum.

The application of the formalism outlined in references [1, 2] to the standard model is not straightforward. At the level of the classical Lagrangian, the electroweak vacuum corresponds to a true minimum of the scalar potential, and it is only upon the inclusion of quantum corrections that the instability occurs. Furthermore, the tunnelling rate calculation in the standard model including the first quantum corrections (originally addressed in reference [3]) is complicated by the presence of many degrees of freedom with different spin, and by the approximate scale invariance of the theory at high-energy scales, relevant for the tunnelling process in the standard model.

In the second part of this work, we will review several aspects of the calculation of the electroweak vacuum lifetime. In particular, we will carefully discuss the role of approximate scale invariance for the determination of the standard model bounce, which selects energy scales of order 10^{17} GeV. Since the latter is very close to the Planck scale, it is a legitimate question to ask whether gravitational effects can become relevant in this regime. After addressing the ultraviolet sensitivity of the standard model vacuum decay rate [4], we will take gravitational effects into account within the formalism of Coleman and De Luccia [5], arguing that as long as gravity can be treated as an effective field theory in a perturbative regime [6], the standard model calculation does not get drastically affected.

The review is organized as follows: in section 2 we introduce the phenomenon of quantum tunnelling in quantum mechanics with some explicit examples. We then generalize this formalism to the case of quantum mechanics with many degrees of freedom, thus providing the tools required for dealing with tunnelling in quantum field theory, which is the main subject of section 3. Section 4 is devoted to a review of effective potential methods, a necessary step in order to introduce the problem of the instability of the Higgs potential in the standard model. The calculation of the standard model vacuum decay rate is illustrated in section 5, together with the construction on the so-called standard model phase diagram and the possible impact of non-standard physics. Section 6 is devoted to the study of gravitational corrections to the standard model vacuum decay rate.

Some technical issues are collected in the appendices, mainly with the purpose of casting well-known results in the notations adopted in the text, for the ease of the reader. Appendix A.1 is a simple review of the semi-classical approximation in ordinary quantum mechanics, which is applied in the following appendix A.2 to the calculation of the lowest-level energy splitting for a symmetric double-well potential. This is a simple exercise in quantum mechanics, which is included here for the purpose of comparison with the analogous calculation by path-integral methods presented in the text. Small differences between the two results are discussed. The Maupertuis principle is reviewed in appendix A.3, while appendix A.4 describes briefly Legendre transformations. Appendix A.5 contains a derivation of the Fubini–Lipatov bounce. In appendix A.6 we describe the techniques employed to obtain bounce solutions by numerical methods. Finally, the derivation of the gravitational field equations for a scalar field minimally coupled to gravity is described in appendix A.7.

2. Tunnelling in quantum mechanics

In this section we review the phenomenon of tunnelling through barriers due to quantum fluctuations. We will start from ordinary quantum mechanics in one space dimension, extend the results to the case of many degrees of freedom, and finally generalize to quantum field theory. We will follow closely the work by Coleman and collaborators [1, 2]. See also reference [7] for a comprehensive review.

2.1. Tunnelling in one dimension

Tunnelling is a typical phenomenon of quantum mechanics. The simplest example is the case of a beam of particles (or a wave packet) moving in one space dimension towards a potential energy barrier $U(x)$ whose maximum value is larger than the particle energy: in quantum mechanics the particles of the beam have a non-zero probability of being transmitted beyond the potential barrier, contrary to what happens in classical mechanics. The relevant quantity to be computed is the transmission coefficient, defined as the ratio between the transmitted flux and the incident flux.

If the potential barrier is sufficiently high and broad, the transmission coefficient is conveniently computed within the semi-classical or Wentzel–Kramers–Brillouin (WKB) approximation, which is reviewed in some detail in appendix A.1. It turns out that, in the leading semi-classical approximation, the transmission coefficient $\mathcal{T}$ through a potential barrier is given by [8]

$$\mathcal{T} = A_T e^{-B}; \quad B = \frac{2}{\hbar}\int_a^b |p(x)|\mathrm{d}x = \frac{2}{\hbar}\int_a^b \sqrt{2m[U(x)-E]}\mathrm{d}x, \tag{2.1}$$

where E is the energy of the incident particles, and a, b are the classical turning points, $U(a) = U(b) = E$. The overall factor A_T is not fixed in the leading semi-classical approximation.

The result equation (2.1) can also be used to estimate the decay probability per unit time γ of a metastable state, which is the case of interest to us. Indeed, if the barrier separates a local minimum of the potential energy from a deeper minimum, then the probability per unit time γ that the particle, initially localized around the local minimum with energy E_0, the fundamental level of the local minimum, penetrates the barrier, is proportional to $\mathcal{T}$. The exact proportionality coefficient cannot be computed within the lowest-order semi-classical approximation. However, it must be proportional to the number N of times the particle hits the barrier per unit time. Approximating the initial state as an oscillatory state with frequency $\omega = \frac{E_0}{\hbar}$, we have simply $N = \frac{\omega}{2\pi} = \frac{1}{t_{\mathrm{cl}}}$, where t_{cl} is the classical period of the oscillation. Hence [9]

$$\gamma \simeq \frac{\mathcal{T}}{t_{\mathrm{cl}}} \equiv A e^{-B}. \tag{2.2}$$

It is customary to define a decay width

$$\Gamma = \hbar\gamma, \tag{2.3}$$

with the dimension of an energy, whose physical meaning will be clear in a moment.

Finally, we note that, if the particle has a probability per unit time γ to escape from the local minimum, then the probability $P(t)$ to find it in the vicinity of the local minimum decreases in time according to

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = -\gamma P(t), \tag{2.4}$$

which has the solution

$$P(t) = P(0)e^{-\gamma t} = P(0)e^{-\frac{\Gamma t}{\hbar}}. \tag{2.5}$$

This suggests that we may interpret the unstable state as an approximate stationary state with complex energy eigenvalue:

$$E = E_0 - i\frac{\Gamma}{2}, \tag{2.6}$$

so that the corresponding wave function evolves in time according to

$$\Psi(x,t) = \psi(x)e^{-i\frac{E_0 t}{\hbar}}e^{-\frac{\Gamma t}{2\hbar}}, \tag{2.7}$$

consistently with equation (2.5).

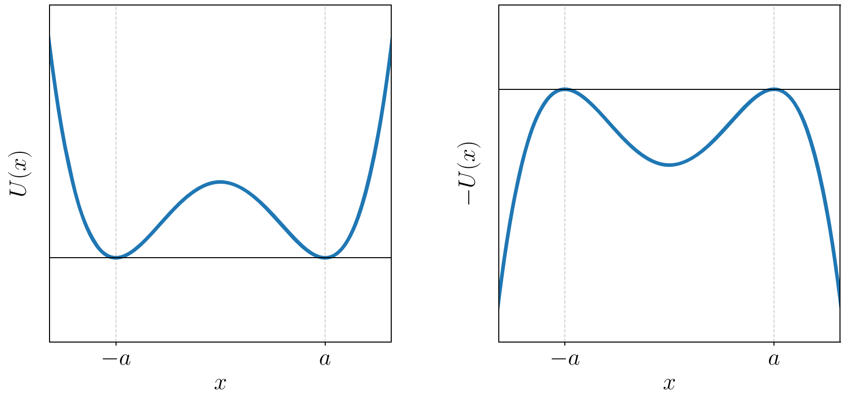

Figure 1. Left panel: double-well potential with two symmetric minima in $\pm a$. Right panel: the same potential after the change of sign.

2.2. The double-well potential

A typical application of the semi-classical approximation method is the calculation of the splitting between the two lowest energy eigenvalues of a one-dimensional Hamiltonian whose potential energy has two degenerate minima, such as the one sketched in the left panel of figure 1. The standard calculation can be found for example in reference [8], and it is reviewed in appendix A.2.

For our present purposes, it will prove useful to obtain the result of appendix A.2 in the context of the path integral formulation of quantum mechanics. In the case of one-dimensional ordinary quantum mechanics the path integral calculation turns out to be much more lengthy and complicated, but it has the advantage that it can be generalized to quantum field theory beyond the leading semi-classical approximation.

We will assume that the potential energy $U(x)$ has two degenerate minima at $x = \pm a$, and that $U''(\pm a) = m\omega^2$, where m is the mass of the particle and ω a real constant. An explicit example of such a potential is

$$U(x) = \frac{m\omega^2}{8a^2}(x^2 - a^2)^2. \tag{2.8}$$

It will often be useful to refer to the specific example in equation (2.8), for which analytic calculations are possible. Our final result, however, applies to a more general class of one-dimensional potential energies with a symmetric double well. In particular, we will keep $U_0 = U(\pm a)$ different from zero.

The first step is switching to imaginary time, $t = i\tau$. The Lagrange function reads

$$L(x, \dot{x}) = -\frac{1}{2}m\dot{x}^2 - U(x), \tag{2.9}$$

where x is the position of the particle,

$$\dot{x} = \frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau}, \tag{2.10}$$

and the conjugate momentum to the coordinate x is

$$p = \frac{\partial L(x, \dot{x})}{\partial \dot{x}} = -m\dot{x}, \tag{2.11}$$

so that the Hamiltonian reads

$$H(x, p) = p\dot{x} - L(x, \dot{x}) = -\frac{1}{2}m\dot{x}^2 + U(x). \tag{2.12}$$

This is called the Euclidean Hamiltonian, because the Minkowski space with imaginary time has a Euclidean metric (this is a slight abuse of language: the Minkowski space–time is not relevant here, since we are in a non-relativistic context).

We consider the probability amplitudes for the particle to start at $x = -a$ at $\tau = -\frac{T}{2}$ and end up at either $x = +a$ or $x = -a$ at $\tau = \frac{T}{2}$:

$$Z_{\pm}(T) = \langle \pm a| e^{-\frac{HT}{\hbar}} | -a\rangle, \tag{2.13}$$

where $|\pm a\rangle$ are position eigenstates with eigenvalues $\pm a$. Inserting a complete set $|n\rangle$ of energy eigenstates

$$H|n\rangle = E_n|n\rangle \tag{2.14}$$

we get

$$Z_{\pm}(T) = \sum_n e^{-\frac{E_n T}{\hbar}} \langle \pm a|n\rangle\langle n| - a\rangle. \tag{2.15}$$

If the potential barrier between the two minima is sufficiently high and broad, the two lowest levels E_1, E_2 differ from the lowest level in the absence of tunnelling, $E_0 = \frac{\hbar\omega}{2} + U_0$, by an amount which is much smaller than the separation of E_0 from higher levels, of order $\hbar\omega$. Hence, in the large-T limit, the sum in equation (2.15) is dominated by the two lowest eigenstates:

$$Z_{-}(T) = e^{-\frac{E_1 T}{\hbar}} \langle -a|1\rangle\langle 1| - a\rangle + e^{-\frac{E_2 T}{\hbar}} \langle -a|2\rangle\langle 2| - a\rangle \tag{2.16}$$

$$Z_{+}(T) = e^{-\frac{E_1 T}{\hbar}} \langle +a|1\rangle\langle 1| - a\rangle + e^{-\frac{E_2 T}{\hbar}} \langle +a|2\rangle\langle 2| - a\rangle, \tag{2.17}$$

up to terms that vanish more rapidly as $T \to \infty$. Furthermore, the wave functions of states $|1\rangle$ and $|2\rangle$ in the position representation are even and odd, respectively; thus

$$\langle +a|1\rangle = \langle -a|1\rangle; \qquad \langle +a|2\rangle = -\langle -a|2\rangle. \tag{2.18}$$

Finally,

$$|\langle 1| + a\rangle|^2 = |\langle 2| + a\rangle|^2. \tag{2.19}$$

As a consequence,

$$Z_{-}(T) = |\langle 1|a\rangle|^2 \left(e^{-\frac{E_1 T}{\hbar}} + e^{-\frac{E_2 T}{\hbar}}\right) \tag{2.20}$$

$$Z_{+}(T) = |\langle 1|a\rangle|^2 \left(e^{-\frac{E_1 T}{\hbar}} - e^{-\frac{E_2 T}{\hbar}}\right) \tag{2.21}$$

and therefore, for large T,

$$\frac{Z_-(T) - Z_+(T)}{Z_-(T) + Z_+(T)} = e^{-\frac{(E_2 - E_1)T}{\hbar}}, \tag{2.22}$$

or

$$E_2 - E_1 = -\hbar \lim_{T \to +\infty} \frac{1}{T} \log \frac{Z_-(T) - Z_+(T)}{Z_-(T) + Z_+(T)}. \tag{2.23}$$

Equation (2.23) allows us to compute the energy splitting between the two lowest energy levels, provided the amplitudes $Z_\pm(T)$ can be computed in an independent way.

The path integral formalism provides such an independent calculation. Indeed, by standard path integral arguments it can be shown that

$$Z_\pm(T) = \int [\mathrm{d}x]\, e^{-\frac{S[x]}{\hbar}}, \tag{2.24}$$

where $[\mathrm{d}x]$ is a functional measure, to be defined accurately below, over all paths $x(\tau)$ with

$$x\left(-\frac{T}{2}\right) = -a; \qquad x\left(\frac{T}{2}\right) = \pm a, \tag{2.25}$$

and $S[x]$ is the Euclidean action of the path:

$$S[x] = -\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}\tau\, L(x, \dot{x}) = \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}\tau \left[\frac{1}{2} m\dot{x}^2 + U(x)\right]. \tag{2.26}$$

Let us now assume that $S[x]$ has a stationary point, i.e. a path $\overline{x}(\tau)$ such that $S[\overline{x}]$ is finite and stationary upon deformations around $\overline{x}(\tau)$. In this case, the path integral in equation (2.24) can be computed by means of a generalization of the saddle-point approximation technique. This is motivated by the observation that the exponent in the integrand of equation (2.24) is weighted by a factor of $\frac{1}{\hbar}$, which is large in the semi-classical limit $\hbar \to 0$; the contributions to the integral from field configurations far from the minimum configurations are therefore exponentially suppressed in this limit.

In the leading saddle-point approximation, the functional integral is simply given by the value of the integrand computed in correspondence of the path which minimizes $S[x]$, that is, a solution of the classical equations of motions. The next-to-leading correction involves a calculation of the first non-trivial fluctuations of the exponent around the stationary point; for this reason we need the idea of a functional Taylor expansion, and therefore of functional differentiation. Functional derivatives are defined by

$$\frac{\delta x(\tau)}{\delta x(\tau')} = \delta(\tau - \tau'), \tag{2.27}$$

together with the usual rules of differentiations of sums and products of ordinary functions. We find

$$\begin{aligned}\frac{\delta S[x]}{\delta x(\tau)} &= \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}\tau' \left[m \frac{\mathrm{d}x}{\mathrm{d}\tau'} \frac{\mathrm{d}}{\mathrm{d}\tau'} \delta(\tau - \tau') + U'(x)\delta(\tau - \tau')\right] \\ &= -m\ddot{x}(\tau) + U'(x(\tau)),\end{aligned} \tag{2.28}$$

where we have used

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}\tau\, f(\tau)\frac{\mathrm{d}^n}{\mathrm{d}\tau^n}\delta(\tau) = (-1)^n \frac{\mathrm{d}^n f(0)}{\mathrm{d}\tau^n}. \tag{2.29}$$

The second functional derivative of the action is given by

$$\begin{aligned}\frac{\delta^2 S[x]}{\delta x(\tau')\delta x(\tau)} &= -m\,\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\delta(\tau-\tau') + U''(x)\delta(\tau-\tau')\\ &= \left[-m\,\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + U''(x)\right]\delta(\tau-\tau') \equiv S''[x]\delta(\tau-\tau').\end{aligned} \tag{2.30}$$

We now take the functional Taylor expansion of $S[x]$ in the vicinity of a solution $\overline{x}(\tau)$ of the classical equation of motion:

$$\left.\frac{\delta S[x]}{\delta x(\tau)}\right|_{x=\overline{x}} = -m\ddot{\overline{x}} + U'(\overline{x}) = 0 \tag{2.31}$$

with the boundary conditions equation (2.25). We find

$$\begin{aligned}S[x] &= S[\overline{x}] + \frac{1}{2}\int_{-\frac{T}{2}}^{\frac{T}{2}}\mathrm{d}\tau\int_{-\frac{T}{2}}^{\frac{T}{2}}\mathrm{d}\tau'\,\delta x(\tau)S''[\overline{x}]\delta(\tau-\tau')\delta x(\tau') + O((\delta x)^3)\\ &= S[\overline{x}] + \frac{1}{2}\int_{-\frac{T}{2}}^{\frac{T}{2}}\mathrm{d}\tau\,\delta x(\tau)S''[\overline{x}]\delta x(\tau) + O((\delta x)^3),\end{aligned} \tag{2.32}$$

where $\delta x(\tau) = x(\tau) - \overline{x}(\tau)$.

As in the case of a finite-dimensional Gaussian integral, it is convenient to introduce a set of eigenfunctions of $S''[\overline{x}]$:

$$S''[\overline{x}]\psi_n(\tau) = \left[-m\,\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + U''(\overline{x})\right]\psi_n(\tau) = \lambda_n\psi_n(\tau), \tag{2.33}$$

which can be taken to be orthogonal and normalized:

$$\int_{-\frac{T}{2}}^{\frac{T}{2}}\mathrm{d}\tau\,\psi_n(\tau)\psi_m(\tau) = \delta_{nm};\quad \psi_n\left(\pm\frac{T}{2}\right) = 0. \tag{2.34}$$

A generic path can be parametrized as

$$x(\tau) = \overline{x}(\tau) + \sum_n c_n\psi_n(\tau). \tag{2.35}$$

The integration over all possible paths is therefore identified with an integration over all possible values of the coefficients c_n. This suggests that the functional measure $[\mathrm{d}x]$ be defined by

$$[\mathrm{d}x] = N\prod_n \frac{\mathrm{d}c_n}{\sqrt{2\pi\hbar}}. \tag{2.36}$$

The factors of $\frac{1}{\sqrt{2\pi\hbar}}$ and the overall factor N have been introduced for later convenience. The expansion of $S[x]$ around $\overline{x}$ to second order, equation (2.32), takes the form

$$S[x] = S[\overline{x}] + \frac{1}{2}\sum_n \lambda_n c_n^2 + O(c_n^3), \tag{2.37}$$

and the path integral equation (2.24) can be computed as an infinite product of Gaussian integrals:

$$\begin{aligned} Z_{\pm}(T) &= N e^{-\frac{S[\bar{x}]}{\hbar}} \prod_n \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}c_n \, e^{-\frac{1}{2\hbar}\lambda_n c_n^2} \\ &= N e^{-\frac{S[\bar{x}]}{\hbar}} \prod_n \frac{1}{\sqrt{\lambda_n}} \end{aligned} \tag{2.38}$$

(the dependence of the rhs of equation (2.38) on the $\pm$ sign in the lhs is hidden in the boundary values of the stationary path $\bar{x}$ and in the eigenvalues λ_n). We now formally define the determinant of an operator on a space of functions, denoted by the symbol Det, as the product of its eigenvalues:

$$\mathrm{Det}\; S''[\bar{x}] = \prod_n \lambda_n, \tag{2.39}$$

so that, at least formally,

$$Z_{\pm}(T) = \frac{N}{\sqrt{\mathrm{Det}\; S''[\bar{x}]}} e^{-\frac{S[\bar{x}]}{\hbar}}. \tag{2.40}$$

A few comments are in order.

- The result equation (2.40) holds provided all eigenvalues λ_n are strictly positive. We will see that this is not the case in some relevant situations, including the present one. We will show later in this section how to deal with this difficulty.
- The determinant in equation (2.40) is divergent, since it is the infinite product of positive and growing numbers. This infinity is usually absorbed by a suitable definition of the constant N in the functional integration measure. However, we shall not need an explicit expression of N, because it cancels in the calculation of the splitting between the two lowest energy eigenvalues.
- If there are more paths which minimize the Euclidean action, $x_n(\tau), n = 0, 1, \ldots$, the path integral receives one contribution equation (2.40) from each of them:

$$Z_{\pm}(T) = N \sum_n \frac{1}{\sqrt{\mathrm{Det}\; S''[\bar{x}_n]}} e^{-\frac{S[\bar{x}_n]}{\hbar}}. \tag{2.41}$$

In order to complete the calculation of the energy splitting $E_2 - E_1$ from equation (2.23), we must therefore (i) find all the minimum paths for the Euclidean action, and (ii) learn how to compute determinants of operators on spaces of functions. Both steps are not trivial, and we will carefully go through the details.

2.2.1. Stationary paths. We first address the problem of finding all paths $\bar{x}(\tau)$ which minimize the Euclidean action. In order to help intuition, it is useful to view equation (2.31) as the equation of motion of a particle under the effect of the potential $-U(x)$ in real time (see the right panel of figure 1).

As far as $Z_{-}(T)$ is concerned, for which $x(T/2) = -a$, one solution is obviously

$$\bar{x}(\tau) = x_0(\tau) = -a \tag{2.42}$$

for all times τ. We have

$$S[x_0] = U_0 T; \qquad S''[x_0] = -m\,\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + m\omega^2, \tag{2.43}$$

and the corresponding contribution to the path integral is

$$I[x_0] = \frac{N}{\sqrt{\mathrm{Det}\ S''[x_0]}} e^{-\frac{S[x_0]}{\hbar}}. \tag{2.44}$$

The action $S[x_0]$ vanishes if we set the zero of the potential at $x = \pm a$, as in the case of equation (2.8). This is the choice adopted in figure 1 and in appendix A.2. For the time being we keep however $U_0 \neq 0$ for greater generality. Note that in this case $S[x_0] = U_0 T$ is only finite for finite T.

In the case $x(T/2) = +a$, the particle starts from $x = -a$ at time $\tau = -T/2$ with zero velocity, and reaches $x = a$ at time $\tau = T/2$. This solution, which we will denote by $x_1(\tau)$, is usually called an *instanton*. The explicit form of $x_1(\tau)$ depends on the shape of the potential energy $U(x)$, but its main features follow from general considerations.

(a) Because $U(x)$ is an even function, and the equation of motion is time-reversal invariant, we have $x_1(-\tau) = -x_1(\tau)$, and therefore $x_1(0) = 0$.

(b) From energy conservation,

$$-\frac{1}{2}m\dot{x}_1^2 + U(x_1) = U_0 \tag{2.45}$$

we get

$$\frac{\mathrm{d}x_1(\tau)}{\mathrm{d}\tau} = \sqrt{\frac{2[U(x_1) - U_0]}{m}}, \tag{2.46}$$

where we have selected the positive sign because $x_1(\tau)$ is an increasing function of τ. Equation (2.46) implies that the particle spends most of the time close to either the initial position $x = -a$ or the final position $x = a$; the transition between these two values occurs within a short time interval of order $1/\omega$ (the *instanton size*) around $\tau = 0$ (the *instanton centre*).

(c) When $T \to +\infty$, the instanton centre can be located anywhere; hence, in this limit we must consider all solutions $x_1(\tau - \tau_1)$ with $-\frac{T}{2} < \tau_1 < \frac{T}{2}$. The one-instanton contribution is obtained by summing over all possible locations of the instanton center, i.e. by integrating over τ_1 in this range:

$$I[x_1] = \int_{-T/2}^{T/2} \mathrm{d}\tau_1\, \frac{N}{\sqrt{\mathrm{Det}\ S''[x_1]}} e^{-\frac{S[x_1]}{\hbar}} = \frac{N}{\sqrt{\mathrm{Det}\ S''[x_1]}} T e^{-\frac{S[x_1]}{\hbar}}. \tag{2.47}$$

(d) Using again equation (2.45), we find that the instanton action $S[x_1]$ can be expressed either in terms of the potential energy:

$$\begin{aligned} S[x_1] &= \int_{-T/2}^{T/2} \mathrm{d}\tau\, 2U(x_1) - U_0 T \\ &= \int_{-T/2}^{T/2} \mathrm{d}\tau\, 2[U(x_1) - U_0] + U_0 T \\ &= \int_{-a}^{a} \mathrm{d}x\, \sqrt{2m[U(x) - U_0]} + S[x_0], \end{aligned} \tag{2.48}$$

or in terms of the kinetic energy:

$$S[x_1] = \int_{-T/2}^{T/2} \mathrm{d}\tau\, m\dot{x}_1^2 + S[x_0]. \tag{2.49}$$

Both expressions will prove useful. Note that $S[x_0]$ (and therefore $S[x_1]$) is only finite for finite values of T, while the difference $S[x_1] - S[x_0]$ is finite even in the limit $T \to +\infty$.

We can check the above features of the instanton solution in the case of the potential equation (2.8). In this case, equation (2.46) can be integrated analytically:

$$x_1(\tau) = a\frac{e^{\omega\tau} - 1}{e^{\omega\tau} + 1}. \tag{2.50}$$

In this explicit example, $x_1(\tau)$ is close $\pm a$ for $|\tau|$ larger than a few times $\frac{1}{\omega}$; indeed $x_1(\pm 4/\omega) \simeq \pm 0.96a$. Hence, as anticipated, $x_1(\tau)$ differs sizeably from $\pm a$ only in an interval $-\delta < \tau < \delta$, with δ of order a few times ω^{-1}. Furthermore, the instanton action can be computed analytically in this case. Using equation (2.48) and $S[x_0] = 0$ we find

$$S[x_1] = \frac{m\omega}{2a}\int_{-a}^{a} \mathrm{d}x\,(a^2 - x^2) = \frac{2ma^2\omega}{3}. \tag{2.51}$$

As anticipated, in the limit $T \to \infty$ the operator $S''[x_1]$ has a *zero mode*, that is, an eigenfunction with zero eigenvalue:

$$\psi_0(\tau) = C\,\frac{\mathrm{d}x_1(\tau)}{\mathrm{d}\tau}, \tag{2.52}$$

where

$$C = \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}\tau \left|\frac{\mathrm{d}x_1(\tau)}{\mathrm{d}\tau}\right|^2\right]^{-\frac{1}{2}} = \sqrt{\frac{m}{S[x_1] - S[x_0]}} \tag{2.53}$$

by equation (2.49). Indeed,

$$\left[-m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + U''(x_1)\right]\frac{\mathrm{d}x_1}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left[-m\ddot{x}_1 + U'(x_1)\right] = 0 \tag{2.54}$$

by the equation of motion. The presence of a zero mode implies that the integration over the corresponding coefficient c_0 leads to an infinity. Luckily, we do not need to perform this integration, because an infinitesimal change in c_0 is equivalent to an infinitesimal shift in the instanton centre τ_1:

$$\frac{\mathrm{d}x(\tau)}{\mathrm{d}c_0} = \psi_0(\tau) = \sqrt{\frac{m}{S[x_1] - S[x_0]}}\,\frac{\mathrm{d}x_1}{\mathrm{d}\tau} \tag{2.55}$$

$$\frac{\mathrm{d}x(\tau)}{\mathrm{d}\tau_1} = \frac{\mathrm{d}x_1}{\mathrm{d}\tau_1} = -\frac{\mathrm{d}x_1}{\mathrm{d}\tau}, \tag{2.56}$$

or equivalently

$$\frac{\mathrm{d}c_0}{\sqrt{2\pi\hbar}} = -\sqrt{\frac{S[x_1] - S[x_0]}{2\pi\hbar m}}\,\mathrm{d}\tau_1. \tag{2.57}$$

Since we have already performed an integration over the instanton centre τ_1, we may simply omit the integration over c_0, which amounts to omitting the zero eigenvalue in the determinant, and multiply the result with the jacobin factor $\sqrt{\frac{S[x_1]-S[x_0]}{2\pi\hbar m}}$:

$$I[x_1] = \frac{N}{\sqrt{\mathrm{Det}'\, S''[x_1]}}\sqrt{\frac{S[x_1]-S[x_0]}{2\pi\hbar m}}\,T e^{-\frac{S[x_1]}{\hbar}} = I[x_0] K T e^{-\frac{S[x_1]-S[x_0]}{\hbar}}, \quad (2.58)$$

where $I[x_0]$ is given in equation (2.44), we have defined

$$K = \sqrt{\frac{S[x_1]-S[x_0]}{2\pi\hbar m}}\sqrt{\frac{\mathrm{Det}\; S''[x_0]}{\mathrm{Det}'\, S''[x_1]}} \quad (2.59)$$

and Det' stands for a determinant with the zero eigenvalue removed.

More approximate solutions of the equation of motion can be built out of the following observation. Given an instanton solution $x_1(\tau)$, then, by time reversal invariance,

$$x_1(-\tau) = -x_1(\tau) \quad (2.60)$$

is also a solution, with boundary conditions interchanged, and the same action as the instanton. This is called an *anti-instanton*. It follows that any sequence of an even [odd] number n of alternate instanton and anti-instanton solutions contributes to the saddle-point estimate of $Z_-(T)$ $[Z_+(T)]$. We denote these solutions by $x_n(\tau)$.

In order to compute $S[x_n]$ we observe that, as in the single-instanton case, we may use energy conservation to obtain

$$S[x_n] = \int_{-T/2}^{T/2} \mathrm{d}\tau\, m\dot{x}_n^2 + S[x_0]. \quad (2.61)$$

The function $x_n(\tau)$ is approximately equal to $\pm a$ almost everywhere in time, except for n small intervals, of order $1/\omega$ in size, where it behaves as the single instanton or anti-instanton solutions. We may therefore approximate it by

$$x_n(\tau) \simeq x_0 + \sum_{k=1}^{n} (-1)^{k+1}[x_1(\tau-\tau_k) - x_0], \quad (2.62)$$

which gives

$$\dot{x}_n^2 = \left[\sum_{k=1}^{n} (-1)^{k+1}\dot{x}_1(\tau-\tau_k)\right]^2 \simeq \sum_{k=1}^{n} \dot{x}_1^2(\tau-\tau_k), \quad (2.63)$$

where in the last step we have used the fact that $x_1(\tau-\tau_k)$ is approximately constant for τ far from τ_k. Hence, in the large-T limit,

$$S[x_n] = n\,(S[x_1] - S[x_0]) + S[x_0], \quad (2.64)$$

and

$$I[x_n] = e^{-\frac{S[x_0]}{\hbar}}\left(e^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right)^n \int [\mathrm{d}x] \times \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2} \mathrm{d}\tau\, x(\tau) S''[x_n] x(\tau)\right]. \quad (2.65)$$

In order to compute the functional integral, we observe that $S''[x_n]$ is approximately given by

$$S''[x_n] \simeq \begin{cases} S''[x_1(\tau-\tau_k)] & \tau_k-\delta<\tau<\tau_k+\delta \\ S''[x_0] & \text{elsewhere,} \end{cases} \tag{2.66}$$

where $k=1,\ldots,n$ and $\delta\sim\omega^{-1}$. This is a consequence of the fact that $x_n(\tau)$ is a multi-instanton solution in which the different instantons and anti-instantons are well separated from each other (this is usually called a dilute instanton gas; the validity of this assumption must be verified, which we will do in a moment).

Let us now denote by $[\mathrm{d}x]_k$ the functional integration measure over paths $x(\tau)$ that are sizably different from zero only in the range $\tau_k-\delta<\tau<\tau_k+\delta$, where $\tau_k, k=1,\ldots,n$ are the centres of the n instantons and anti-instantons. Using the approximation equation (2.66) we obtain

$$\begin{aligned}
&\int[\mathrm{d}x]\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_n]x(\tau)\right]\\
&\quad\simeq\int[\mathrm{d}x]\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_0]x(\tau)\right]\\
&\qquad\times\prod_{k=1}^{n}\frac{\int[\mathrm{d}x]_k\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_1(\tau-\tau_k)]x(\tau)\right]}{\int[\mathrm{d}x]_k\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_0])]x(\tau)\right]}.
\end{aligned}\tag{2.67}$$

The ratio under the product sign is in fact independent of k. Hence

$$\begin{aligned}
&\prod_{k=1}^{n}\frac{\int[\mathrm{d}x]_k\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_1(\tau-\tau_k)]x(\tau)\right]}{\int[\mathrm{d}x]_k\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_0]x(\tau)\right]}\\
&\quad=\left[\frac{\int[\mathrm{d}x]_k\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_1(\tau-\tau_k)]x(\tau)\right]}{\int[\mathrm{d}x]_k\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_0]x(\tau)\right]}\right]^n\\
&\quad=\left[\frac{\int[\mathrm{d}x]\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_1(\tau-\tau_k)]x(\tau)\right]}{\int[\mathrm{d}x]\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_0]x(\tau)\right]}\right]^n.
\end{aligned}\tag{2.68}$$

In the last step the functional integrals have been extended to all paths, not restricted by the condition of being different from zero only around τ_k. This is allowed in the ratio, because $S''[x_1(\tau-\tau_k)]$ and $S''[x_0]$ essentially coincide for τ not too close to τ_k. Hence, leaving aside for the moment the problem of the zero mode of $S''[x_1]$, we find

$$\int[\mathrm{d}x]\ \exp\left[-\frac{1}{2}\int_{-T/2}^{T/2}\mathrm{d}\tau\, x(\tau)S''[x_n]x(\tau)\right]\simeq\frac{N}{\sqrt{\mathrm{Det}\ S''[x_0]}}\left[\sqrt{\frac{\mathrm{Det}\ S''[x_0]}{\mathrm{Det}\ S''[x_1]}}\right]^n. \tag{2.69}$$

To complete the calculation of the n-instanton contribution, we must remove the zero eigenvalue from each factor of $\det S''[x_1]$, include n jacobian factors

$$\sqrt{\frac{S[x_1]-S[x_0]}{2\pi\hbar m}} \tag{2.70}$$

and integrate over all possible locations $\tau_1, \ldots, \tau_n$ of the instanton centres. This yields a factor of

$$\begin{aligned}&\left[\sqrt{\frac{S[x_1]-S[x_0]}{2\pi\hbar m}}\right]^n \int_{-T/2}^{T/2} \mathrm{d}\tau_1 \int_{\tau_1}^{T/2} \mathrm{d}\tau_2 \ldots \int_{\tau_{n-1}}^{T/2} \mathrm{d}\tau_n \\ &\quad = \left[\sqrt{\frac{S[x_1]-S[x_0]}{2\pi\hbar m}}\right]^n \frac{T^n}{n!}.\end{aligned} \tag{2.71}$$

So finally

$$\begin{aligned} I[x_n] &= \frac{N}{\sqrt{\mathrm{Det}\ S''[x_0]}} \frac{K^n T^n}{n!} e^{-\frac{n(S[x_1]-S[x_0])}{\hbar}} e^{-\frac{S[x_0]}{\hbar}} \\ &= I[x_0] \frac{1}{n!}\left[KTe^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right]^n \end{aligned} \tag{2.72}$$

with K defined in equation (2.59). Note that our previous results equations (2.44) and (2.58) are recovered for $n = 0, 1$ respectively.

We are now in a position to check the validity of the dilute instanton gas assumption. When all contributions $I[x_n]$ are summed up, the dominant contribution to the sum comes from values of n for which

$$\frac{1}{n!}\left[KTe^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right]^n \sim \frac{1}{(n-1)!}\left[KTe^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right]^{n-1} \tag{2.73}$$

or

$$\frac{n}{T} \sim Ke^{-\frac{S[x_1]-S[x_0]}{\hbar}}. \tag{2.74}$$

We see that the instanton density $\frac{n}{T}$ for the relevant contributions is exponentially suppressed, as long as $S[x_1] - S[x_0] \gg \hbar$, which is the condition of validity of the semi-classical approximation (we will be able to check this statement in the explicit example of equation (2.8) after computing K). This is an *a posteriori* confirmation of the reliability of our assumption that instantons and anti-instantons are well separated from each other in the relevant multi-instanton configurations.

2.2.2. Functional determinants. We now turn to the calculation of the ratio

$$\frac{\mathrm{Det}\ S''[x_0]}{\mathrm{Det}'\ S''[x_1]} \tag{2.75}$$

which appears in the definition of K, equation (2.59). The procedure is illustrated in reference [10]; we will reproduce here the argument presented there, including some details which are omitted in reference [10]. We first show how to compute the ratio

$$\frac{\mathrm{Det}\ O_1}{\mathrm{Det}\ O_2}, \tag{2.76}$$

where

$$O_i = -m\,\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + W_i(\tau); \quad i = 1, 2 \tag{2.77}$$

and $W_i(\tau)$ are arbitrary functions of τ in the range $-T/2 \leqslant \tau \leqslant T/2$. Then, we remove the contribution of the minimum eigenvalue, which is different from zero as long as T is kept finite, from the denominator of equation (2.76). Finally, we take the limit $T \to +\infty$.

Let us consider solutions $\phi_\lambda^{(i)}(\tau)$ of the Cauchy problem

$$O_i \phi_\lambda^{(i)}(\tau) = \lambda \phi_\lambda^{(i)}(\tau) \tag{2.78}$$

$$\phi_\lambda^{(i)}\left(-\frac{T}{2}\right) = 0 \tag{2.79}$$

$$\dot{\phi}_\lambda^{(i)}\left(-\frac{T}{2}\right) = 1. \tag{2.80}$$

The important point here is that the function $\phi_\lambda^{(i)}(\tau)$, which certainly exists, because it is the unique solution to the Cauchy problem equations (2.78), (2.79) and (2.80), is not necessarily an eigenvector of O_i, because $\phi_\lambda^{(i)}(T/2)$ is in general different from zero. Also note that $\phi_\lambda^{(i)}(\tau)$ is not, in general, normalized to one.

We now define

$$F(\lambda) = \frac{\mathrm{Det}\,(O_1 - \lambda)}{\mathrm{Det}\,(O_2 - \lambda)}; \qquad G(\lambda) = \frac{\phi_\lambda^{(1)}\left(\frac{T}{2}\right)}{\phi_\lambda^{(2)}\left(\frac{T}{2}\right)}. \tag{2.81}$$

The function $F(\lambda)$ has a simple zero whenever λ is an eigenvalue $\lambda_n^{(1)}$ of O_1, and a simple pole whenever λ is an eigenvalue $\lambda_n^{(2)}$ of O_2. The function $G(\lambda)$ has exactly the same poles and zeros. By Liouville's theorem, the ratio $\frac{F(\lambda)}{G(\lambda)}$ is therefore a constant. Since both $F(\lambda)$ and $G(\lambda)$ tend to 1 as $\lambda \to \infty$ in any direction except the real axis, where both functions have simple poles, the constant is equal to one. Hence

$$F(\lambda) = G(\lambda) \tag{2.82}$$

in the whole complex plane λ. In particular, for $\lambda \to 0$,

$$\frac{\mathrm{Det}\ O_1}{\mathrm{Det}\ O_2} = \frac{\phi_0^{(1)}\left(\frac{T}{2}\right)}{\phi_0^{(2)}\left(\frac{T}{2}\right)}, \tag{2.83}$$

which is the desired result.

We are interested in the case

$$O_1 = S''[x_0] = -m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + m\omega^2 \qquad O_2 = S''[x_1] = -m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + U(x_1). \tag{2.84}$$

We need the solutions of the Cauchy problems

$$\left(-m\,\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}+m\omega^2\right)\phi_0(\tau)=0;\qquad \phi_0\left(-\frac{T}{2}\right)=0;\qquad \dot{\phi}_0\left(-\frac{T}{2}\right)=1 \tag{2.85}$$

$$\left(-m\,\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}+U''(x_1)\right)\psi_0(\tau)=0;\qquad \psi_0\left(-\frac{T}{2}\right)=0;\qquad \dot{\psi}_0\left(-\frac{T}{2}\right)=1. \tag{2.86}$$

The solution of equation (2.85) is easily found:

$$\phi_0(\tau)=\frac{e^{\omega(\tau+T/2)}-e^{-\omega(\tau+T/2)}}{2\omega}=\frac{1}{\omega}\,\sinh\,\omega\left(\tau+\frac{T}{2}\right). \tag{2.87}$$

Equation (2.86) requires some work. We have

$$\psi_0(\tau)=c_1f(\tau)+c_2g(\tau), \tag{2.88}$$

where $f(\tau)$ and $g(\tau)$ are two independent solutions with constant Wronskian determinant $w(\tau)$. We may normalize them so that

$$w(\tau)=f(\tau)\dot{g}(\tau)-\dot{f}(\tau)g(\tau)=1. \tag{2.89}$$

The initial conditions

$$\psi_0\left(-\frac{T}{2}\right)=c_1f\left(-\frac{T}{2}\right)+c_2g\left(-\frac{T}{2}\right)=0 \tag{2.90}$$

$$\dot{\psi}_0\left(-\frac{T}{2}\right)=c_1\dot{f}\left(-\frac{T}{2}\right)+c_2\dot{g}\left(-\frac{T}{2}\right)=1 \tag{2.91}$$

give

$$c_1=-g\left(-\frac{T}{2}\right)\qquad c_2=f\left(-\frac{T}{2}\right) \tag{2.92}$$

with the choice equation (2.89). Since we are interested in $\psi_0(T/2)$ for large T, we only need the asymptotic behaviours of $f(\tau)$ and $g(\tau)$. We already know that $f(\tau)=\dot{x}_1(\tau)$. The asymptotic behaviour of $f(\tau)$ can be obtained from equation (2.46) and from the observation that $x_1\to a$ for $\tau\to+\infty$. In this limit

$$\frac{\mathrm{d}x_1(\tau)}{\mathrm{d}\tau}=\sqrt{\frac{2}{m}[U(x_1)-U_0]}\simeq-\omega(x_1-a)\quad\text{for }\tau\to+\infty \tag{2.93}$$

which gives, in the same limit,

$$a-x_1(\tau)\simeq\hat{a}e^{-\omega\tau};\qquad f(\tau)=\dot{x}_1(\tau)\simeq\omega\hat{a}e^{-\omega\tau}, \tag{2.94}$$

where $\hat{a}$ is a positive constant, which is determined as follows. We have

$$\tau = \sqrt{\frac{m}{2}} \int_0^{x_1} \frac{\mathrm{d}x}{\sqrt{U(x) - U_0}} \tag{2.95}$$

from equation (2.46), and

$$\tau = -\frac{1}{\omega} \log \frac{\epsilon}{\hat{a}} + O(\epsilon^2) \tag{2.96}$$

from equation (2.94), where $\epsilon = a - x_1$. Equating equations (2.95) and (2.96) we get

$$\hat{a} = \lim_{\epsilon \to 0} \epsilon \exp\left[m\omega \int_0^{a-\epsilon} \frac{\mathrm{d}x}{\sqrt{2m[U(x) - U_0]}}\right]. \tag{2.97}$$

The integral in the exponent in equation (2.97) is logarithmically divergent as $\epsilon \to 0$, due to the behaviour of the integrand at the upper integration bound:

$$U(x) - U_0 = \frac{1}{2} m\omega^2 (x - a)^2 + O((x - a)^3)\big). \tag{2.98}$$

We have

$$\begin{aligned} m\omega \int_0^{a-\epsilon} & \frac{\mathrm{d}x}{\sqrt{2m[U(x) - U_0]}} \\ &= \int_0^{a-\epsilon} \mathrm{d}x \left[\frac{m\omega}{\sqrt{2m[U(x) - U_0]}} - \frac{1}{a - x}\right] + \int_0^{a-\epsilon} \frac{\mathrm{d}x}{a - x} \\ &= \int_0^{a-\epsilon} \mathrm{d}x \left[\frac{m\omega}{\sqrt{2m[U(x) - U_0]}} - \frac{1}{a - x}\right] + \log \frac{a}{\epsilon}, \end{aligned} \tag{2.99}$$

and the integral is now convergent as $\epsilon \to 0$. Therefore

$$\hat{a} = a \exp \int_0^{a} \mathrm{d}x \left[\frac{m\omega}{\sqrt{2m[U(x) - U_0]}} - \frac{1}{a - x}\right]. \tag{2.100}$$

In the case of the potential in equation (2.8) we find

$$\sqrt{2m[U(x) - U_0]} = \frac{m\omega}{2a}(a^2 - x^2); \quad \hat{a} = 2a. \tag{2.101}$$

The asymptotic behaviour of $g(\tau)$ can now be obtained from equation (2.89), which can be rewritten

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \frac{g(\tau)}{f(\tau)} = \frac{1}{f^2(\tau)}. \tag{2.102}$$

The solution for $\tau \to +\infty$ is

$$g(\tau) = \frac{1}{2\hat{a}\omega^2} e^{\omega\tau} \quad \text{for } \tau \to +\infty. \tag{2.103}$$

Hence, in the large-T limit,

$$\psi_0\left(\frac{T}{2}\right) = -f\left(\frac{T}{2}\right)g\left(-\frac{T}{2}\right) + f\left(-\frac{T}{2}\right)g\left(\frac{T}{2}\right)$$
$$= \frac{1}{2\omega}\left[e^{-\frac{\omega T}{2}}e^{\frac{\omega T}{2}} + e^{\frac{\omega T}{2}}e^{-\frac{\omega T}{2}}\right] = \frac{1}{\omega}. \tag{2.104}$$

Using equation (2.87) we finally obtain

$$\frac{\mathrm{Det}\ S''[x_0]}{\mathrm{Det}\ S''[x_1]} = \frac{\phi_0\left(\frac{T}{2}\right)}{\psi_0\left(\frac{T}{2}\right)} = \sinh\,\omega T. \tag{2.105}$$

The next step is the removal of the smallest eigenvalue from the denominator of equation (2.105). To this purpose, we consider the differential equation

$$-m\ddot{\psi}_\lambda(\tau) + U''(\overline{x}_1)\psi_\lambda(\tau) = \lambda\psi_\lambda(\tau); \quad \psi_\lambda\left(-\frac{T}{2}\right) = 0; \quad \dot{\psi}_\lambda\left(-\frac{T}{2}\right) = 1 \tag{2.106}$$

and we take λ to be small, because we know that the smallest eigenvalue is zero for large T:

$$\psi_\lambda(\tau) = \psi_0(\tau) + \lambda\eta(\tau) + O(\lambda^2). \tag{2.107}$$

Replacing in equation (2.106) and expanding to first order in λ we get

$$-m\ddot{\eta}(\tau) + U''(x_1)\eta(\tau) = \psi_0(\tau); \qquad \eta\left(-\frac{T}{2}\right) = \dot{\eta}\left(-\frac{T}{2}\right) = 0. \tag{2.108}$$

This differential equation can be turned into an integral equation

$$\eta(\tau) = -\frac{1}{m}\int_{\frac{T}{2}}^{\tau}\mathrm{d}\tau''\int_{-\frac{T}{2}}^{\tau''}\mathrm{d}\tau'\,\left[\psi_0(\tau') - U''(x_1)\eta(\tau')\right]$$
$$= -\frac{1}{m}\int_{-\frac{T}{2}}^{\tau}\mathrm{d}\tau'(\tau-\tau')\left[\psi_0(\tau') - U''(x_1)\eta(\tau')\right] \tag{2.109}$$

and solved by iteration. The result has the form

$$\eta(\tau) = \int_{-\frac{T}{2}}^{\tau}\mathrm{d}\tau'\,k(\tau,\tau')\psi_0(\tau') \tag{2.110}$$

where the kernel $k(\tau,\tau')$ is a solution of

$$\left(-m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + U''(x_1)\right)k(\tau,\tau') = 0; \quad k(\tau,\tau) = 0; \quad \left[\frac{\mathrm{d}}{\mathrm{d}\tau}k(\tau,\tau')\right]_{\tau'=\tau} = -\frac{1}{m}, \tag{2.111}$$

as one can check directly by replacing equation (2.110) in equation (2.108). Hence, $k(\tau,\tau')$ is a linear combination of $f(\tau)$ and $g(\tau)$ with τ'-dependent coefficients. We find

$$k(\tau,\tau') = \frac{1}{m}\left[f(\tau)g(\tau') - g(\tau)f(\tau')\right], \tag{2.112}$$

which obeys the relevant initial conditions thanks to the choice equation (2.89). So finally

$$\psi_\lambda(\tau) = \psi_0(\tau) + \frac{\lambda}{m}\int_{-\frac{T}{2}}^{\tau} \mathrm{d}\tau' \left[f(\tau)g(\tau') - g(\tau)f(\tau')\right]\psi_0(\tau') + O(\lambda^2). \quad (2.113)$$

The smallest eigenvalue is determined by the condition $\psi_\lambda(T/2) = 0$, which gives

$$\frac{1}{\omega} + \frac{\lambda}{m}\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}\tau' \left[f\left(\frac{T}{2}\right)g(\tau') - g\left(\frac{T}{2}\right)f(\tau')\right]\psi_0(\tau') = 0, \quad (2.114)$$

where we have used equation (2.104). For large T the integrand reads

$$\begin{aligned}
&\left[f\left(\frac{T}{2}\right)g(\tau') - g\left(\frac{T}{2}\right)f(\tau')\right]\left[-g\left(-\frac{T}{2}\right)f(\tau') + f\left(-\frac{T}{2}\right)g(\tau')\right] \\
&\simeq \left[\omega\hat{a}e^{-\frac{\omega T}{2}}g(\tau') - \frac{1}{2\hat{a}\omega^2}e^{\frac{\omega T}{2}}f(\tau')\right]\left[\frac{1}{2\hat{a}\omega^2}e^{\frac{\omega T}{2}}f(\tau') + \omega\hat{a}e^{-\frac{\omega T}{2}}g(\tau')\right] \\
&\simeq -\frac{1}{4\omega^4\hat{a}^2}f^2(\tau')e^{\omega T},
\end{aligned} \quad (2.115)$$

and therefore in this limit equation (2.114) becomes

$$1 - \frac{\lambda}{m\omega^3}\frac{1}{4\hat{a}^2}e^{\omega T}\int_{-\frac{T}{2}}^{\frac{T}{2}} \mathrm{d}\tau_1\, f^2(\tau_1) = 1 - \frac{\lambda}{m\omega^3}\frac{1}{4\hat{a}^2}e^{\omega T}\frac{S[x_1] - S[x_0]}{m} = 0, \quad (2.116)$$

where we have used equation (2.52). Thus

$$\lambda = \frac{4\hat{a}^2 m^2\omega^3}{S[x_1] - S[x_0]}e^{-\omega T}. \quad (2.117)$$

Observe that the smallest eigenvalue tends to 0 as $T \to \infty$, as expected. We conclude that

$$\frac{\mathrm{Det}\ S''[x_0]}{\mathrm{Det}'\, S''[x_1]} = \lambda\frac{\mathrm{Det}\ S''[x_0]}{\mathrm{Det}\ S''[x_1]} = \frac{4\hat{a}^2 m^2\omega^3}{S[x_1] - S[x_0]}e^{-\omega T}\ \sinh\,\omega T, \quad (2.118)$$

which is finite and T independent as $T \to +\infty$:

$$\lim_{T\to+\infty}\frac{\mathrm{Det}\ S''[x_0]}{\mathrm{Det}'\, S''[x_1]} = \frac{2\hat{a}^2 m^2\omega^3}{S[x_1] - S[x_0]}. \quad (2.119)$$

2.2.3. Energy splitting of the lowest energy levels. Collecting all our results, we find

$$\begin{aligned}
Z_-(T) &= \sum_{n\ \text{even}} I[x_n] = I[x_0]\sum_{n\ \text{even}}\frac{1}{n!}\left(KTe^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right)^n \\
&= I[x_0]\cosh\left(KTe^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right)
\end{aligned} \quad (2.120)$$

$$\begin{aligned}
Z_+(T) &= \sum_{n\ \text{odd}} I[x_n] = I[x_0]\sum_{n\ \text{odd}}\frac{1}{n!}\left(KTe^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right)^n \\
&= I[x_0]\sinh\left(KTe^{-\frac{S[x_1]-S[x_0]}{\hbar}}\right),
\end{aligned} \quad (2.121)$$

where

$$K=\sqrt{\frac{S[x_1]-S[x_0]}{2\pi\hbar m}}\sqrt{\frac{\mathrm{Det}\ S''[x_0]}{\mathrm{Det}'\ S''[x_1]}}=\hat{a}\omega\sqrt{\frac{m\omega}{\pi\hbar}}. \tag{2.122}$$

The energy splitting E_2-E_1 is given by equation (2.23), which reads

$$E_2-E_1=\hbar\omega\left[2\hat{a}\sqrt{\frac{m\omega}{\pi\hbar}}\right]e^{-\frac{S[x_1]-S[x_0]}{\hbar}}, \tag{2.123}$$

where

$$S[x_1]-S[x_0]=\int_{-a}^{a}\mathrm{d}x\ \sqrt{2m[U(x)-U_0]}. \tag{2.124}$$

Note that the result is independent of the constant N, which appears in the functional integration measure, and of $\mathrm{Det}\,S''[x_0]$, because the factors of $I[x_0]$, cancel in the ratio.

Recalling the expression equation (2.74) for the instanton density in the dominant multi-instanton solutions, and the expression equation (2.51) for the instanton action, we obtain, in the case of the potential equation (2.8),

$$\frac{n}{T}=2\omega\sqrt{\frac{ma^2\omega}{\pi\hbar}}\ \exp\left(-\frac{2ma^2\omega}{3\hbar}\right), \tag{2.125}$$

which is exponentially small for $ma^2\omega\gg\hbar$, thus justifying the use of the dilute instanton gas approximation.

2.3. Decay of a metastable state

Although admittedly cumbersome, the procedure described in the previous subsection lends itself to an extension to the case of the decay probability per unit time of an unstable state. Let us consider a one-dimensional potential energy $U(x)$ with a narrow local minimum in $x=a$, with $U(a)=U_0$ and $U''(a)=m\omega^2$, separated from a lower, much broader minimum by a potential barrier. An example of such a potential is sketched in the left panel of figure 2. As in the case of the double potential well, if the barrier were infinitely broad and high, there would be two distinct sets of energy eigenstates. When tunnelling is taken into account, we expect the eigenstates of the potential well with a local minimum in $x=a$ to become unstable, or equivalently, the energy eigenvalues to acquire an imaginary part. We shall see that the path integral formulation of the problem confirms this expectation, and also allows us to compute the imaginary part of the lowest eigenvalue, which is related to the decay probability per unit time of the metastable ground state.

The calculation is essentially the same as in the case of the double well, with a few modifications. We consider the amplitude

$$Z(T)=\langle a|e^{-\frac{HT}{\hbar}}|a\rangle=\sum_n e^{-\frac{E_nT}{\hbar}}|\langle a|n\rangle|^2, \tag{2.126}$$

where $|n\rangle$ is a complete set of eigenstates of the Hamiltonian with tunnelling neglected. We expect E_0, the energy of the lowest-lying metastable state, to acquire a negative imaginary part when tunnelling is turned on:

$$\mathrm{Im}\,E_0=-\frac{\Gamma}{2} \tag{2.127}$$

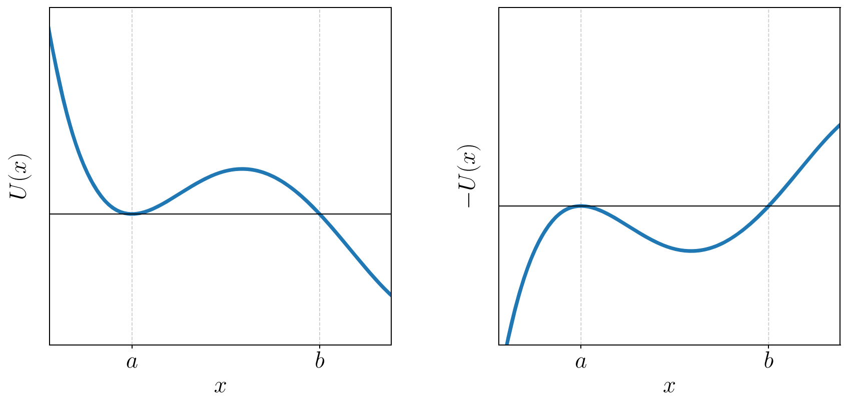

Figure 2. Left panel: potential of a metastable state. Right panel: the same potential after the change of sign.

related to the decay probability γ of the metastable ground state per unit time by $\Gamma = \hbar\gamma$.

The contribution to the amplitude equation (2.126) of the state with minimum real part of E_n is isolated by taking the limit $T \to +\infty$. We find

$$E_0 = -\hbar \lim_{T\to+\infty} \frac{1}{T} \log Z(T); \qquad \gamma = 2\,\mathrm{Im} \lim_{T\to+\infty} \frac{1}{T} \log Z(T). \tag{2.128}$$

As in the previous case, the amplitude $Z(T)$ is given by

$$Z(T) = \int [\mathrm{d}x]\, e^{-\frac{S[x]}{\hbar}}, \tag{2.129}$$

and the functional integral can be computed by the saddle-point technique.

We therefore look for stationary points of the action, i.e. solutions $\overline{x}(\tau)$ of the classical equation of motion, with $\overline{x}(-T/2) = \overline{x}(T/2) = a$. One of them is the constant solution $x_0(\tau) = a$. Other solutions with the same boundary conditions are sequences of the so-called *bounce* solution, namely, a solution $x_b(\tau)$ which starts at $x = a$ in the remote Euclidean past $\tau = -T/2$ with zero velocity, reaches the point $x = b$ beyond the potential barrier with $U(b) = U(a)$ at time $\tau = 0$, and bounces off to reach again $x = a$ at $\tau = T/2$. The path integral is given by the sum of saddle-point contributions around all multibounce solutions, sequences of n bounces located at different instants, this time with no constraints on the parity of n, integrated over all possible locations of the n bounces, in analogy to the case of the double potential well.

The details of the bounce solution obviously depend on the specific form of the potential energy $U(x)$; however, as in the case of the instanton, its general features are independent of the details. As in the case of the double-well, the behaviour of the bounce solution is better understood if the Euclidean equation of motion

$$\frac{\mathrm{d}^2 x}{\mathrm{d}\tau^2} = U'(x) \tag{2.130}$$

is viewed as the real-time equation of motion n a potential $-U(x)$ (see right panel of figure 2), which has therefore a maximum at $x = a$ and a well for $a < x < b$.

(a) The bounce $x_b(\tau)$ actually equals a only in the limit $\tau \to -\infty$. Indeed, since the bounce energy is $U(a)$, in the vicinity of $x = a$ we have

$$\frac{\mathrm{d}x_b}{\mathrm{d}\tau} = \sqrt{\frac{2[U(x_b) - U(a)]}{m}} \simeq \omega(x_b - a) \tag{2.131}$$

with the solution

$$x_b(\tau) = a + Ce^{\omega\tau}, \tag{2.132}$$

which approaches a at $\tau \to -\infty$.

(b) For the same reason, the bounce has zero velocity at time $\tau = 0$, when it reaches the point $x = b$, where the potential energy has another zero:

$$\left.\frac{\mathrm{d}x_b(\tau)}{\mathrm{d}\tau}\right|_{\tau=0} = \sqrt{\frac{2[U(b) - U(a)]}{m}} = 0. \tag{2.133}$$

(c) In the limit $T \to +\infty$, the bounce can reach $x = b$ at any time τ_1. An integration over all possible values of τ_1 is therefore necessary, in analogy with the integration over all possible instanton centres in the previous case.

Using the results of the previous section, we get

$$\begin{aligned} Z(T) &= \frac{N}{\sqrt{\mathrm{Det}\ S''[x_0]}} e^{-\frac{S[x_0]}{\hbar}} \sum_{n=0}^{\infty} \frac{K^n T^n}{n!} e^{-\frac{n(S[x_b]-S[x_0])}{\hbar}} \\ &= \frac{N}{\sqrt{\mathrm{Det}\ S''[x_0]}} e^{-\frac{S[x_0]}{\hbar}} \exp\left[KTe^{-\frac{S[x_b]-S[x_0]}{\hbar}}\right]. \end{aligned} \tag{2.134}$$

We now turn to a careful evaluation of K. A naive expectation would be an expression like equation (2.122), with the instanton solution $x_1(\tau)$ replaced by the bounce $x_b(\tau)$. There is however an important difference: because the bounce has zero velocity at $\tau = 0$, the zero mode $\dot{x}_b(\tau)$ has a node, and therefore cannot be the eigenvector of $S''[x_b]$ with smallest eigenvalue (the eigenvalue equation for $S''[x_b]$ is formally the same as a time-independent Schrödinger equation in one dimension, and we can rely on well known results in that context). This in turn implies that $S''[x_b]$ has an eigenvector with negative eigenvalue, and therefore its determinant (with the zero eigenvalue removed) is negative. This is expected, because as a consequence K is purely imaginary, as appropriate for a metastable state. The contribution of the negative mode to the Gaussian integral, however, requires some care.

Let us investigate in some more detail the origin of the negative eigenvalue of $S''[x_b]$. The presence of a negative eigenvalue means that x_b is not a minimum of the Euclidean action in the space of paths, but rather a saddle point: there is a direction in the space of configurations along which the action $S[x]$ has a maximum at $x(\tau) = x_b(\tau)$. It is not difficult to identify this direction. Let us consider the value of the Euclidean action in correspondence of the sequence of paths $x(\tau)$ with $x(\pm\infty) = a$ and a maximum at $\tau = 0$, and let us parametrise them by their value at the maximum, $c = x(0)$. Members of this family are the constant solution, $x_0(\tau) = a$, for which $c = a$, and the bounce, for which $c = b$ (we assume $b > a$ for definiteness). Other values of c correspond to paths that either go back to $x = a$ without reaching the classically allowed region beyond the barrier (those with $c < b$) or go beyond the turning point, $c > b$.

The Euclidean action, restricted to this family of paths, is an ordinary function $S(c)$ of c. We have

$$S(a) = S[x_0] = U(a)T, \tag{2.135}$$

which is a local minimum, because increasing c with respect to a increases both terms in the action. The action increases monotonically until $c = b$ (the bounce) is reached; in $c = b$ the action has a stationary point,

$$\left.\frac{\mathrm{d}S(c)}{\mathrm{d}c}\right|_{c=b} = 0, \tag{2.136}$$

because the bounce is another solution of the equation of motion. If we now increase c further, the action starts decreasing down to $-\infty$, because the path spends more and more time in the region where $U(x) < U(a)$. Hence, the path with $c = b$ (the bounce) must be a maximum, because there is no other stationary point. We conclude that the bounce is a minimum configuration among those paths which turn back at $x = b$, but a maximum in the direction of paths which turn back at different values (either larger or smaller than b). This clarifies the origin of the negative mode.

Since, however, the action goes to $-\infty$ as c increases beyond $c = b$, the contribution of this direction to the Gaussian integral is divergent. It must be defined by an analytic continuation from the case when $x = a$ is a global minimum, to the case of interest. Such an analytic continuation can be performed restricting ourselves to the direction in the configuration space which corresponds to a local maximum. Following reference [11], we may cast this contribution in the form of an integral over the parameter c:

$$J = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{d}c\, e^{-\frac{S(c)}{\hbar}}. \tag{2.137}$$

As long as the potential energy has an absolute minimum in $x = a$, J is convergent and real. If we now continuously deform the potential energy so that a new, lower minimum appears for $x > b$, the integral is finite and real between $-\infty$ and b, but diverges between b and $+\infty$ because $S(c) \to -\infty$. We may regularize the integral by deforming the integration path (the real c axis) away from the real axis for $\mathrm{Re}\, c > b$. This gives J an imaginary part, which can be evaluated by the steepest descent method:

$$\begin{aligned}\mathrm{Im}\, J &= \mathrm{Im}\, \frac{1}{\sqrt{2\pi\hbar}} \int_{b}^{b+i\infty} \mathrm{d}c\, e^{-\frac{S(c)}{\hbar}} = \frac{1}{\sqrt{2\pi\hbar}} \int_{0}^{+\infty} \mathrm{d}c\, e^{-\frac{S(b+ic)}{\hbar}} \\ &\simeq \frac{e^{-\frac{S(b)}{\hbar}}}{\sqrt{2\pi\hbar}} \int_{0}^{+\infty} \mathrm{d}c\, e^{\frac{S''(b)}{2\hbar}c^2} = \frac{1}{2} e^{-\frac{S(b)}{\hbar}} \frac{1}{\sqrt{|S''(b)|}},\end{aligned} \tag{2.138}$$

where $S''(b) < 0$. Note the factor of $\frac{1}{2}$, arising from the Gaussian integration over one half of the Gaussian peak. Taking this into account, the value of K in the case of an unstable minimum is given by

$$K = \frac{i}{2} \sqrt{\frac{S[x_b] - S[x_0]}{2\pi\hbar m}} \frac{\sqrt{\mathrm{Det}\ S''[x_0]}}{\sqrt{|\mathrm{Det}'\, S''[x_b]|}}. \tag{2.139}$$

We are now in a position to use equation (2.128) to compute the decay probability per unit time of the metastable ground state:

$$\begin{aligned}\gamma &= 2\,\mathrm{Im}\lim_{T\to+\infty}\frac{1}{T}\,\log\,Z(T)\\ &= 2\,\mathrm{Im}\lim_{T\to+\infty}\left[\frac{1}{T}\,\log\left(\frac{N}{\sqrt{\det\,S''[x_0]}}e^{-\frac{S[x_0]}{\hbar}}\right)+Ke^{-\frac{S[x_b]-S[x_0]}{\hbar}}\right]\\ &= \sqrt{\frac{S[x_b]-S[x_0]}{2\pi\hbar m}}\frac{\sqrt{\mathrm{Det}\,S''[x_0]}}{\sqrt{|\mathrm{Det}'\,S''[x_b]|}}e^{-\frac{S[x_b]-S[x_0]}{\hbar}}\\ &\equiv Ae^{-B}.\end{aligned} \tag{2.140}$$

Note that the prefactor A has the correct dimension of an inverse time:

$$[A] = \mathrm{mass}^{-\frac{1}{2}}\times\left(\frac{\mathrm{mass}}{\mathrm{time}^2}\right)^{\frac{1}{2}} = \frac{1}{\mathrm{time}}. \tag{2.141}$$

The exponent

$$B = \frac{S[x_b]-S[x_0]}{\hbar} \tag{2.142}$$

in equation (2.140) can be computed in terms of the potential energy as in the case of the previous section, using

$$-\frac{1}{2}m\left(\frac{\mathrm{d}x_b}{\mathrm{d}\tau}\right)^2 + U(x_b) = U(a). \tag{2.143}$$

We find

$$S[x_0] = \int_{-T/2}^{T/2}\mathrm{d}\tau\left[\frac{1}{2}m\left(\frac{\mathrm{d}x_0}{\mathrm{d}\tau}\right)^2+U(x_0)\right] = U(a)T \tag{2.144}$$

$$\begin{aligned}S[x_b]-S[x_0] &= \int_{-T/2}^{T/2}\mathrm{d}\tau\left[\frac{1}{2}m\left(\frac{\mathrm{d}x_b}{\mathrm{d}\tau}\right)^2+U(x_b)-U(a)\right]\\ &= \int_{-T/2}^{T/2}\mathrm{d}\tau\;2[U(x_b)-U(a)].\end{aligned} \tag{2.145}$$

The integral in the last line of equation (2.145) can be turned to a space integral, using equation (2.143) again, to obtain

$$B = \frac{2}{\hbar}\int_a^b\mathrm{d}x\,\sqrt{2m[U(x)-U(a)]}. \tag{2.146}$$

Equations (2.140) and (2.146) are the main result of this section.

As a useful consistency check, we may compute the real part of E_0 by the same technique; we expect it to be $U(a)+\frac{\hbar\omega}{2}$, as appropriate for a potential well with a minimum in $x = a$ and

$U''(a) = m\omega^2$. We find

$$\begin{aligned} \mathrm{Re}\, E_0 &= -\hbar\, \mathrm{Re} \lim_{T\to\infty} \frac{1}{T} \log Z(T) \\ &= -\hbar \lim_{T\to+\infty} \frac{1}{T}\left[\log N - \frac{1}{2}\log \mathrm{Det}\, S''[x_0] - \frac{S[x_0]}{\hbar}\right]. \end{aligned} \tag{2.147}$$

The constant N is T-independent, and disappears from this expression in the large-T limit, while $S[x_0] = U(a)T$. Finally, $\mathrm{Det}\, S''[x_0]$ is proportional to $\phi_0(T)$, given in equation (2.87). Hence

$$\begin{aligned} \mathrm{Re}\, E_0 &= \mathrm{U}(a) + \frac{\hbar}{2} \lim_{T\to+\infty} \frac{1}{T} \log \frac{\sinh \omega T}{\omega} \\ &= U(a) + \frac{\hbar\omega}{2} \end{aligned} \tag{2.148}$$

as expected.

2.4. Tunnelling with many degrees of freedom

The extension of the above results to quantum systems with more than one degree of freedom was developed in references [12, 13]. We consider a generic quantum system with N degrees of freedom. The relevant generalized coordinates are collected in an N-dimensional vector $\mathbf{x}$. The Lagrangian of the system is

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}\left|\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}\right|^2 - U(\mathbf{x}) = \frac{1}{2}|\dot{\mathbf{x}}(t)|^2 - U(\mathbf{x}) \tag{2.149}$$

(the generalized coordinates $\mathbf{x}$ are not necessarily Cartesian coordinates, hence the absence of a factor of m in the kinetic term).

We assume that the potential energy $U(\mathbf{x})$ has a local minimum in $\mathbf{x} = \mathbf{a}$, and that the system is initially in the ground state of the potential well in $\mathbf{x} = \mathbf{a}$, whose energy E differs from $U(\mathbf{a})$ by an amount of order $\hbar$. By tunnell effect, the system has a finite probability per unit time γ to penetrate the potential barrier, escape from the local minimum, and emerge at some point $\mathbf{x} = \mathbf{b}$ on a surface Σ in the configuration space, defined implicitly by $U(\mathbf{x}) = U(\mathbf{a})$ for all $\mathbf{x}$ on Σ. It is shown in reference [12] that

$$\gamma = Ae^{-B}, \tag{2.150}$$

where the prefactor A is not fixed in the leading semi-classical approximation, while the exponent B can be written in analogy with the one-dimensional case:

$$B = \frac{2}{\hbar}\int_0^{s_f} \mathrm{d}s\, \sqrt{2[U(\mathbf{x}(s)) - E]}, \tag{2.151}$$

where $\mathbf{x}(s)$ is a path in configuration space which starts at the local minimum $\mathbf{x} = \mathbf{a}$ and ends at a point $\mathbf{b}$ on the surface Σ:

$$\mathbf{x}(0) = \mathbf{a}; \qquad \mathbf{x}(s_f) = \mathbf{b}, \tag{2.152}$$

normalized so that

$$\left|\frac{\mathrm{d}\mathbf{x}(s)}{\mathrm{d}s}\right|^2 = \sum_i \frac{\mathrm{d}x_i(s)}{\mathrm{d}s}\frac{\mathrm{d}x_i(s)}{\mathrm{d}s} = 1. \tag{2.153}$$

The path $\mathbf{x}(s)$, including its endpoint $\mathbf{b}$, must be such that the exponent B is minimum, that is, tunnelling takes place along the path (or paths) with maximum decay probability. The problem is therefore reduced to the variational problem of finding those particular paths.

An analogous problem arises in classical mechanics, when one is interested in finding the trajectory of motion, with no reference to the time dependence of the generalized coordinates. The argument is reviewed in appendix A.3; it is based on a modified version of Hamilton's principle, sometimes called the Maupertuis principle. Here, we recall the relevant result. The action of the system is considered as a function of the configuration $\mathbf{x}(t)$ at time t, and t itself:

$$S(\mathbf{x}(t), t) = \int_{t_0}^{t} L(\mathbf{x}(t'), \dot{\mathbf{x}}(t'))\mathrm{d}t', \tag{2.154}$$

where $\mathbf{x}(t')$ is a solution of the equations of motion with $\mathbf{x}(t_0) = \mathbf{a}$. If the energy E of the system is conserved, one finds

$$S(\mathbf{x}(t), t) = S_0(\mathbf{x}(t)) - E(t - t_0), \tag{2.155}$$

where

$$S_0(\mathbf{x}) = \sum_i \int_{a_i}^{x_i} p_i(\mathbf{x}')\mathrm{d}x_i' \tag{2.156}$$

and

$$p_i(\mathbf{x}) = \frac{\partial L}{\partial x_i} \tag{2.157}$$

is the conjugate momentum of x_i. It can be shown that $S_0(\mathbf{x}(t))$, sometimes called the reduced action, is stationary upon variations of $\mathbf{x}(t')$ around a solution of the equations of motion with $\mathbf{a}$ and $\mathbf{x}(t)$ kept fixed, but t allowed to vary. Furthermore, it can be shown that

$$S_0(\mathbf{x}(t)) = \int_0^{s_f} \sqrt{2[E - U(\mathbf{x}(s))]}\,\mathrm{d}s, \tag{2.158}$$

where $\mathbf{x}(s)$ is a parametrization of the physical trajectory normalized as in equation (2.153), with $\mathbf{b} = \mathbf{x}(s_f) = \mathbf{x}(t)$.

This variational problem differs from the one we would like to solve, namely the minimization of the exponent B, equation (2.151), in two respects. First, the quantities under square root in equations (2.151) and (2.158) have opposite signs. This is expected, since the physical solution of the equations of motion refers to the classically allowed region $E \geqslant U(\mathbf{x})$, while the exponent B depends on a path which is defined in the classically forbidden region $E \leqslant U(\mathbf{x})$. Second, the minimization of the exponent B in equation (2.151) should be performed not only with respect to the path $\mathbf{x}(s)$, but also with respect to its endpoint $\mathbf{b}$, while the endpoint $\mathbf{x}(t)$ is kept fixed in the minimization of $S_0(\mathbf{x}(t))$, equation (2.158).

Leaving aside for a while the problem of the endpoint, the two variational problems become exactly the same switching to imaginary (or Euclidean) time

$$t = i\tau. \tag{2.159}$$

The Lagrangian is transformed into

$$L(\mathbf{x}, \dot{\mathbf{x}}) = -\frac{1}{2}\sum_i \frac{\mathrm{d}x_i}{\mathrm{d}\tau}\frac{\mathrm{d}x_i}{\mathrm{d}\tau} - U(\mathbf{x}) = -\frac{1}{2}|\dot{\mathbf{x}}(\tau)|^2 - U(\mathbf{x}). \tag{2.160}$$

Correspondingly,

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = -\dot{x}_i \tag{2.161}$$

and therefore

$$H(\mathbf{p}, \mathbf{x}) = \sum_i p_i \dot{x}_i - L(\mathbf{x}, \dot{\mathbf{x}}) = -\frac{1}{2}\sum_i \frac{\mathrm{d}x_i}{\mathrm{d}\tau}\frac{\mathrm{d}x_i}{\mathrm{d}\tau} + U(\mathbf{x}) = E. \tag{2.162}$$

The Euclidean action is defined by

$$S(\mathbf{x}(\tau), \tau) = -\int_{\tau_0}^{\tau} L(\mathbf{x}(\tau'), \dot{\mathbf{x}}(\tau'))\mathrm{d}\tau'. \tag{2.163}$$

By the same procedure described in appendix A.3 one finds

$$S(\mathbf{x}(\tau), \tau) = S_0(\mathbf{x}(\tau)) + E(\tau - \tau_0), \tag{2.164}$$

where $S_0(\mathbf{x}(\tau))$ is the reduced action:

$$S_0(\mathbf{x}(\tau)) = -\sum_i \int_{a_i}^{x_i(\tau)} p_i\, \mathrm{d}x_i = \int_0^{s_f} \sqrt{2[U(\mathbf{x}(s)) - E]}\, \mathrm{d}s = \frac{\hbar}{2}B. \tag{2.165}$$

Hence, the paths which minimize the exponent B are the solutions of the equation of motion in Euclidean time [1]

$$\frac{\mathrm{d}^2 x_i}{\mathrm{d}\tau^2} = +\frac{\partial U(\mathbf{x})}{\partial x_i} \tag{2.166}$$

with $\mathbf{x}(\tau_0) = \mathbf{a}$ and $\mathbf{x}(\tau) = \mathbf{b}$, where $\mathbf{b}$ is some point such that $U(\mathbf{b}) = U(\mathbf{a})$.

We now turn to the problem of minimizing the path with respect to its endpoint. We note that the reduced action, and hence the exponent B, is automatically minimized with respect to variations of the endpoint $\mathbf{b} = \mathbf{x}(\tau)$ of the path $\mathbf{x}(s)$. Indeed,

$$\frac{\partial}{\partial b_i} S_0(\mathbf{b}) = -\sqrt{2[U(\mathbf{b}) - E]}\frac{\mathrm{d}x_i(s)}{\mathrm{d}s}\bigg|_{s=s_f} \tag{2.167}$$

which is zero for all points $\mathbf{b}$ on the surface Σ.

We now observe that equation (2.166) can be interpreted as an ordinary equation of motion (that is, in ordinary time) in a potential $\tilde{U}(\mathbf{x}) = -U(\mathbf{x})$ which has a local maximum in $\mathbf{x} = \mathbf{a}$. Of course, the explicit form of the solution of the equation of motion depends on the potential $U(\mathbf{x})$. Some of its features can however be obtained on general grounds:

(a)The solution $\mathbf{x}(\tau)$ can only start at $\mathbf{a}$ at $\tau_0 = -\frac{T}{2} \to -\infty$. Indeed, the assumption that $U(\mathbf{x})$ has a minimum in $\mathbf{x} = \mathbf{a}$ implies that

$$U(\mathbf{x}) = E + \frac{1}{2}\sum_{i,j} U_{ij}(\mathbf{a})(x_i - a_i)(x_j - a_j) + O(|\mathbf{x} - \mathbf{a}|^3), \tag{2.168}$$

where the matrix $U_{ij}(\mathbf{a})$ has positive eigenvalues ω_i^2. Hence, after a suitable rotation in the configuration space, the equations of motion for $\mathbf{x}$ close to $\mathbf{a}$ read

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}(x_i - a_i) = \omega_i^2 (x_i - a_i), \tag{2.169}$$

which has the general solution

$$x_i(\tau) - a_i = A_i e^{\omega_i \tau} + B_i e^{-\omega_i \tau}. \tag{2.170}$$

Hence, $x_i(\tau) \to 0$ for either $A_i = 0, \tau \to +\infty$ or $B_i = 0, \tau \to -\infty$. Since τ_0 is the initial time, the only possibility is $\tau_0 \to -\infty$.

(b) Since $\tau_0 = -\infty$, we may perform a time translation such that $\mathbf{x} = \mathbf{b}$ at $\tau = 0$.

(c) The solution reaches $\mathbf{x} = \mathbf{b}$ with zero velocity. Indeed, from equation (2.162),

$$\left| \frac{\mathrm{d}\mathbf{x}(\tau)}{\mathrm{d}\tau} \right|^2 = 2[U(\mathbf{x}) - E], \tag{2.171}$$

and $U(\mathbf{x}) \to E$ for $\mathbf{x} \to \mathbf{b}$.

The exponent B can be related to the action of a solution of the equation of motion computed over the whole Euclidean time range $-\infty < \tau < +\infty$. Indeed, by time reversal invariance, the time-reversed configuration $\mathbf{x}(-\tau)$ is also a solution of the equation of motion, with the same action. Hence,

$$\mathbf{x}_b(\tau) = \mathbf{x}(\tau) + \mathbf{x}(-\tau) \tag{2.172}$$

is a solution with

$$\mathbf{x}_b(\pm\infty) = \mathbf{a}; \qquad \mathbf{x}_b(0) = \mathbf{b}; \qquad \dot{\mathbf{x}}_b(0) = 0 \tag{2.173}$$

and Euclidean action

$$S[x_b] = 2S[x] = 2S_0[x] + ET. \tag{2.174}$$

In analogy with the one-dimensional case, such a solution is usually called a bounce, because it starts in the remote past $\tau = -\frac{T}{2}$ at $\mathbf{x}_b = \mathbf{a}$, reaches $\mathbf{x}_b = \mathbf{b}$ at $\tau = 0$, and bounces back to $\mathbf{x}_b = \mathbf{a}$ at $\tau = +\frac{T}{2}$.

The constant solution

$$\mathbf{x}_0(\tau) = \mathbf{a} \tag{2.175}$$

is also a solution of the equations of motion, with the same boundary conditions. The corresponding Euclidean action is

$$S[x_0] = ET. \tag{2.176}$$

Hence

$$\lim_{T\to+\infty} (S[x_B] - S[x_0]) = \lim_{T\to+\infty} 2S_0[x] = \hbar B. \tag{2.177}$$

Alternatively, one may choose the zero of the potential energy such that $U(\mathbf{a}) = E = 0$; in this case, $S = S_0$ for all solutions with zero energy, and $S[x_B] = \hbar B$, which is consistent with equation (2.177) since $S[x_0] = 0$ in this case.

3. Tunnelling in quantum field theory

After reviewing barrier penetration in ordinary quantum mechanics, we now consider its generalization to quantum field theory. We will follow closely the approach of reference [1]. We

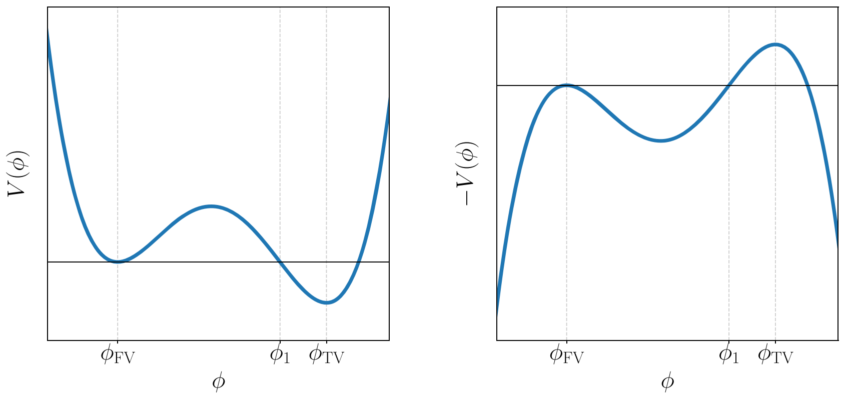

Figure 3. Left panel: asymmetric double-well potential with a local minimum in $\phi_{\rm FV}$ and a global one in $\phi_{\rm TV}$. Right panel: the same potential after the change of sign. ϕ_1 labels the turning point.

will start by simply translating the results obtained in the previous section in the language of quantum field theory, and we will then focus on the physical interpretation of tunnelling processes. Specifically, we will pay special attention to the generalization of the concept of a potential barrier in quantum field theory. Unless explicitly stated, we will adopt the natural unit convention $\hbar = c = 1$ from now on.

3.1. From quantum mechanics to quantum field theory

We consider the quantum theory of a real scalar field ϕ, characterized by a scalar potential $V(\phi)$ with a local minimum in $\phi = \phi_{\rm FV}$ (the false vacuum) and a deeper, absolute minimum in $\phi = \phi_{\rm TV}$ (the true vacuum) as shown in the left panel of figure 3. In the example of figure 3 the scalar potential is defined so that $V(\phi_{\rm FV}) = 0$; this is a convenient choice in many respects, but we will keep $V(\phi_{\rm FV}) \neq 0$ in our discussion for greater generality.

Our goal is the computation of the probability per unit time γ that the system, initially in the false vacuum configuration $\phi_{\rm FV}$, decays to the true vacuum configuration $\phi_{\rm TV}$. To this purpose, we consider our field theory as an ordinary quantum theory with an infinite number of degrees of freedom, labelled by points in three-dimensional space: $x_i(\tau) \rightarrow \phi_{\mathbf{x}}(\tau) = \phi(\tau, \mathbf{x})$. The formalism of section 2.4 is readily generalized, to conclude that

$$\gamma = A e^{-B}, \tag{3.1}$$

where the overall factor A is undetermined in the leading semi-classical approximation, while the exponent B can be computed by a generalization of the argument presented in the previous section. In particular, we need bounce solutions $\phi_b(\tau, \mathbf{x})$ to the Euclidean field equation

$$\left(\frac{\partial^2}{\partial \tau^2} + \boldsymbol{\nabla}^2\right)\phi_b(\tau, \mathbf{x}) = V'(\phi_b), \tag{3.2}$$

with the boundary conditions

$$\lim_{\tau\to\pm\infty} \phi_b(\tau, \mathbf{x}) = \phi_{\mathrm{FV}} \tag{3.3}$$

$$\left.\frac{\partial\phi_b(\tau, \mathbf{x})}{\partial\tau}\right|_{\tau=0} = 0. \tag{3.4}$$

The exponent B is then given by

$$B = S[\phi_b] - S[\phi_{\mathrm{FV}}], \tag{3.5}$$

where $S[\phi]$ is the Euclidean action of the theory,

$$S[\phi] = \int \mathrm{d}^4x \left[\frac{1}{2}(\partial_\mu\phi)^2 + V(\phi)\right]. \tag{3.6}$$

By construction, the bounce solution ϕ_b approaches ϕ_{FV}, and hence $V(\phi_b) \to V(\phi_{\mathrm{FV}})$, at infinite Euclidean time, see equation (3.3). By inspection of equation (3.5) we conclude that ϕ_b must approach ϕ_{FV} also at space infinity, in order to keep the exponent B finite:

$$\lim_{|\vec{x}|\to\infty} \phi_b(\tau, \mathbf{x}) = \phi_{\mathrm{FV}}. \tag{3.7}$$

The exponent B can be written in a more compact form by using the virial theorem. Since the action, and hence B, is stationary in correspondence of solutions of the field equations, it will be stationary, in particular, upon variations of the following type:

$$\phi_b(x) \to \phi_b(ax). \tag{3.8}$$

After such transformations,

$$\begin{aligned} B \to B_a &= \int \mathrm{d}^4x \left[\frac{1}{2}\frac{\partial\phi_b(ax)}{\partial x_\mu}\frac{\partial\phi_b(ax)}{\partial x_\mu} + V(\phi_b(ax)) - V(\phi_{\mathrm{FV}})\right] \\ &= \frac{1}{a^4}\int \mathrm{d}^4y \left[\frac{a^2}{2}\frac{\partial\phi_b(y)}{\partial y_\mu}\frac{\partial\phi_b(y)}{\partial y_\mu} + V(\phi_b(y)) - V(\phi_{\mathrm{FV}})\right], \end{aligned} \tag{3.9}$$

with the change of integration variable $y = ax$. The stationarity condition

$$\left.\frac{\mathrm{d}B_a}{\mathrm{d}a}\right|_{a=1} = 0 \tag{3.10}$$

provides a relationship between the potential and the derivative terms in B, when computed at stationary points:

$$\int \mathrm{d}^4x\, \frac{1}{2}\left(\partial_\mu\phi_b\right)^2 = -2\int \mathrm{d}^4x\, [V(\phi_b) - V(\phi_{\mathrm{FV}})] \tag{3.11}$$

and therefore

$$B = -\int \mathrm{d}^4x\, [V(\phi_b) - V(\phi_{\mathrm{FV}})]. \tag{3.12}$$

Before looking for an explicit bounce solution of the Euclidean field equation, it will be useful to illustrate its physical meaning. In ordinary quantum mechanics, the system initially occupies the false ground state; after an average time $T = 1/\gamma$, the system simply makes a quantum jump and appears at the escape point beyond the barrier with zero kinetic energy (that is, the bounce configuration at time $\tau = 0$). Afterwards, it evolves classically.

A similar picture holds in quantum field theory: at $\tau = 0$ the field makes a quantum jump from the state in which it is uniformly equal to its false vacuum configuration, to the one described by the bounce solution, then it evolves classically. Therefore, the tunnelling process does not take place between two spatially uniform configurations, but rather from a spatially uniform one (the false vacuum) to a space-dependent one (the bounce), which, as we will see, has the shape of a bubble of true vacuum $\phi_b = \phi_{\mathrm{TV}}$ of finite radius, in a false vacuum $\phi_b = \phi_{\mathrm{FV}}$ background. Indeed, as we will show in more detail in section 3.4, the tunnelling between two spatially homogeneous configurations would require penetration through an infinitely high potential energy barrier; the amplitude for such a process is zero. Instead we have a non-zero amplitude for the tunnelling between a spatially homogeneous configuration and one with a region of approximate true vacuum surrounded by a false vacuum background. This configuration then evolves classically, and eventually converts the false vacuum in the true vacuum everywhere in space. Since the bubble of true vacuum can appear anywhere, we expect the decay rate to be proportional to the volume of three-dimensional space $\mathcal{V}_3$.

The result, as described in reference [1], is closely similar to the boiling processes in a superheated fluid: the false vacuum corresponds to the superheated fluid state, while the true vacuum to the vapour state. Due to thermodynamical fluctuations, bubbles of vapour state appear at different points: when a bubble is large enough that the vapour pressure inside the bubble exceeds the external pressure plus the contribution of the surface tension, the bubble grows until it converts all the fluid into vapour, otherwise it shrinks back. By substituting thermodynamical fluctuations with quantum ones, this picture describes the decay of the false vacuum: once a bubble of true vacuum appears, large enough that its growth is energetically favourable, then it expands throughout the Universe, thereby achieving the transition to the true vacuum.

Let us go back to the problem of searching for a bounce solution of equation (3.2). It was shown in reference [14] that, if the field equation has bounce solutions, there always exists one bounce which is $O(4)$ invariant, i.e. a function of

$$r = \sqrt{\tau^2 + |\mathbf{x}|^2} \tag{3.13}$$

alone, whose Euclidean action is smaller than that of any $O(4)$ non-invariant bounce. We may therefore restrict our search to bounce solutions with $O(4)$ invariance, $\phi_b(r)$. This simplifies considerably our task. Indeed, the field equation (3.2) becomes an ordinary differential equation:

$$\frac{\mathrm{d}^2\phi_b}{\mathrm{d}r^2} + \frac{3}{r}\frac{\mathrm{d}\phi_b}{\mathrm{d}r} = V'(\phi_b), \tag{3.14}$$

while the two boundary conditions equations (3.3) and (3.7) combine into a single one:

$$\lim_{r\to\infty} \phi_b(r) = \phi_{\mathrm{FV}}. \tag{3.15}$$

Requiring that the solution is regular at $\mathbf{x} = 0$ we get the further boundary condition

$$\left.\frac{\mathrm{d}\phi_b(r)}{\mathrm{d}r}\right|_{r=0} = 0. \tag{3.16}$$

3.2. Existence of a bounce solution

We now prove that a bounce solution with the correct boundary conditions actually exists in the case of an asymmetric double-well scalar potential $V(\phi)$ with a local minimum in $\phi = \phi_{\mathrm{FV}}$ and a global minimum in $\phi = \phi_{\mathrm{TV}}$. The proof was originally presented in reference [1]; here we reproduce the same argument almost word by word.

It is useful to view the field $\phi(r)$ as the position of a classical particle of unit mass in one dimension, as a function of the time r. Then the field equation, equation (3.14), can be regarded as the classical equation of motion of such a particle in a potential $-V$, which is further subject to a damping force with a time-dependent coefficient which decreases as $1/r$. The potential $-V$ is shown in the right panel of figure 3. We are looking for a solution $\phi_b(r)$ that starts at rest at $r = 0$ from some initial value $\phi_b(0)$ and reaches $\phi_b = \phi_{\mathrm{FV}}$ with zero velocity at $r \to \infty$. We will demonstrate that such a solution exists by showing that if the initial position is too close to the left of ϕ_{TV} then the particle passes through ϕ_{FV} with non-zero velocity (the particle overshoots); if instead the initial position is too far to the left of from ϕ_{TV}, the particle does not reach ϕ_{FV} (the particle undershoots). As a consequence, there must be an intermediate initial position such that the particle comes at rest exactly at ϕ_{FV}.

It is easy to demonstrate the undershoot: if the initial condition $\phi(0)$ is smaller than ϕ_1, then the particle's energy is smaller than the potential energy at $\phi = \phi_{\mathrm{FV}}$, which therefore cannot be reached. The presence of the viscous term just makes things worse by further lowering the energy; indeed, the energy E is not conserved because of the viscous force, and

$$\frac{\mathrm{dE}}{\mathrm{d}r} = \frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{1}{2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}r}\right)^2 - V(\phi)\right] = -\frac{3}{r}\left(\frac{\mathrm{d}\phi}{\mathrm{d}r}\right)^2 < 0 \tag{3.17}$$

by the field equation.

The overshoot would also be obvious if we could neglect the damping force: the energy of a particle starting at rest with $\phi_1 < \phi(0) < \phi_{\mathrm{TV}}$ would be larger than the potential energy at ϕ_{FV}, and would therefore reach ϕ_{FV} with non-zero kinetic energy. The effect of the viscous term can be taken into account as follows. Let us take $\phi(0)$ to be close to ϕ_{TV}. For ϕ close to ϕ_{TV} we may expand the potential in powers of $\phi - \phi_{\mathrm{TV}}$ up to second order:

$$V(\phi) = V(\phi_{\mathrm{TV}}) + \frac{1}{2}\mu^2(\phi - \phi_{\mathrm{TV}})^2 + O((\phi - \phi_{\mathrm{TV}})^3), \tag{3.18}$$

where $\mu^2 = V''(\phi_{\mathrm{TV}}) > 0$. In this limit, equation (3.14) takes the form

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{3}{r}\frac{\mathrm{d}}{\mathrm{d}r} - \mu^2\right)[\phi(r) - \phi_{\mathrm{TV}}] = 0. \tag{3.19}$$

It is easy to check that the solution of this equation is

$$\phi(r) - \phi_{\mathrm{TV}} = 2[\phi(0) - \phi_{\mathrm{TV}}]\frac{I_1(\mu r)}{\mu r} \tag{3.20}$$

and $I_1(x) = -iJ_1(ix)$, where $J_1(x)$ is the Bessel function of order 1, solution of the differential equation

$$J_1''(x) + \frac{J_1'(x)}{x} + \left(1 - \frac{1}{x^2}\right)J_1(x) = 0. \tag{3.21}$$

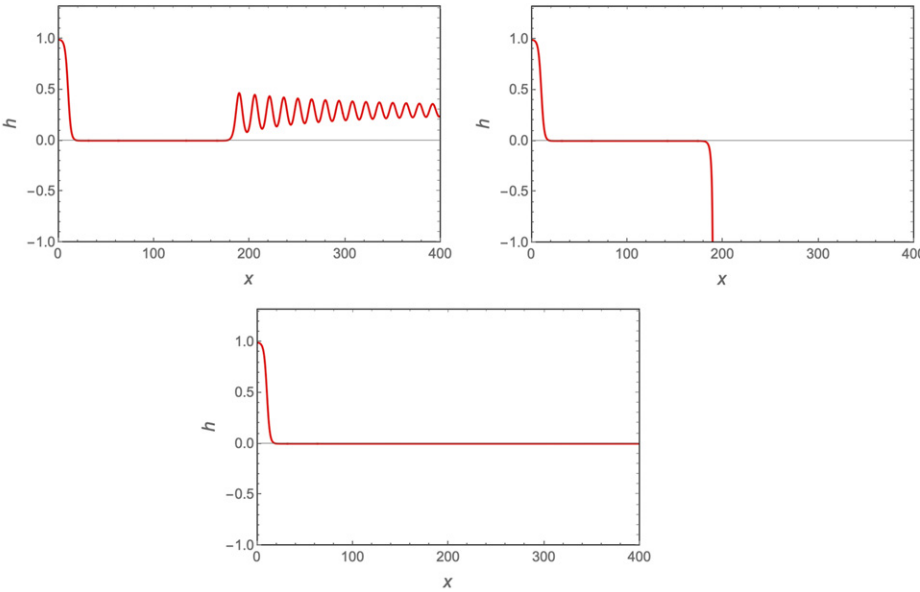

Figure 4. An example of the overshoot–undershoot mechanism.

We see that, provided $\phi(0)$ is sufficiently close to ϕ_{TV}, the particle will spend an arbitrarily large amount of time in the vicinity of ϕ_{TV}.[5] But after an arbitrarily large time r the viscous term, which decreases as r^{-1}, becomes negligible, and in the absence of a damping force the particle overshoots. The announced result is therefore proved.

As an example, we have employed the overshoot–undershoot technique to find the bounce solution in the case of the scalar potential

$$V(\phi) = g\left[\frac{1}{4}\phi^4 - \frac{1}{3}\phi^3(M + \phi_{\mathrm{TV}}) + \frac{1}{2}\phi^2 M\phi_{\mathrm{TV}}\right] \tag{3.22}$$

$$g > 0; \qquad M = 0.3\,\phi_{\mathrm{TV}}, \tag{3.23}$$

which has a local minimum for $\phi = 0$ and an absolute minimum for $\phi = \phi_{\mathrm{TV}}$. In figure 4 we plot $h(x) = \frac{\phi(r)}{\phi_{\mathrm{TV}}}$ as a function of $x = \sqrt{g}r\phi_{\mathrm{TV}}$ for three different choices of $\phi(0)$. The first case is an undershoot: $\phi(0)$ is too much smaller than ϕ_{TV}, the particle gets close the false vacuum but then turns back and oscillates around $\phi = M$. Overshoot is shown in the second panel: the particle has too much energy to come at rest at $\phi = 0$; it reaches the false vacuum with positive velocity, and goes on to negative values of the coordinate. In the third panel, $\phi(0)$ has the right value to reach the false-vacuum value $\phi = 0$ at large r with zero velocity.

The case when the two minima are nearly degenerate (in the above example, the case when M is close to $\phi_{\mathrm{TV}}/2$) is especially interesting. This case corresponds to the so-called thin-wall approximation: the bounce is essentially a constant, equal to the true vacuum field value, for r smaller than some value R, which is a function of the parameters in the theory. Then, within a small region around $r = R$ the bounce drops rapidly to the false vacuum field value. To an

[5] Note that the same conclusion holds even in the absence of the damping force; in such case we would have found $\phi(r) - \phi_{\mathrm{TV}} = [\phi(0) - \phi_{\mathrm{TV}}]\cosh\,\mu r$.

excellent approximation, the bounce is therefore a four-dimensional sphere of radius R where the field takes the true-vacuum value, while $\phi = \phi_{\mathrm{FV}}$ in the rest of Euclidean space–time. In the thin-wall approximation, the calculation of the tunnelling rate can be carried on analytically; the leading semi-classical tunnelling rate (that is, the exponent B) was computed in the original Coleman's paper [1], while a closed-form calculation at one loop in this limit has been presented recently [15].

To summarize, the decay probability per unit time of a metastable state in the leading semi-classical approximation is given by

$$\gamma = Ae^{-B}, \tag{3.24}$$

where

$$\begin{aligned} B &= S[\phi_b] - S[\phi_{\mathrm{FV}}] \\ &= \int \mathrm{d}^4x \left[\frac{1}{2}\partial_\mu\phi_b\partial_\mu\phi_b + V(\phi_b) - V(\phi_{\mathrm{FV}})\right] \\ &= -\int \mathrm{d}^4x\,[V(\phi_b) - V(\phi_{\mathrm{FV}})], \end{aligned} \tag{3.25}$$

where $\phi_b(r)$ is the $O(4)$-invariant bounce. The overall factor A is undetermined in the leading semi-classical limit; we expect it to be proportional to the three-dimensional volume $\mathcal{V}_3$ as a consequence of integration over the position of the true-vacuum bubble. This expectation will be confirmed by the inclusion of the first quantum corrections.

3.3. Time evolution of the true-vacuum bubble

The decay of the false vacuum takes place as illustrated in section 3.1: at some Euclidean time, say $\tau = 0$, and at some point $\mathbf{x}$ of three-dimensional space, an $O(4)$-symmetric bubble of true vacuum is formed in the false-vacuum background by quantum fluctuations:

$$\begin{aligned} &\phi_b(|\mathbf{x}|) = \phi_{\mathrm{TV}}; \\ &\frac{\partial}{\partial\tau}\phi_b\left(\sqrt{|\mathbf{x}|^2+\tau^2}\right)\bigg|_{\tau=0} = 0; \\ &\lim_{\tau\to\infty}\phi_b\left(\sqrt{|\mathbf{x}|^2+\tau^2}\right) = \phi_{\mathrm{FV}}. \end{aligned} \tag{3.26}$$

Then, the bubble evolves in real (i.e. Minkowski space) time according to the classical equations of motion. The qualitative behaviour of the bubble is easy to obtain by analytical continuation of the bounce to Minkowski space:

$$\phi_b(\sqrt{|\mathbf{x}|^2+\tau^2}) \;\to\; \phi_b(\sqrt{|\mathbf{x}|^2-t^2}). \tag{3.27}$$

To visualize the time evolution of the bubble, we may study the time dependence of the three-dimensional surface in space–time separating the true vacuum region from the false vacuum background, which in the thin-wall limit is the hyperbole

$$|\mathbf{x}|^2 - t^2 = R^2. \tag{3.28}$$

We find that the bubble expands in three-dimensional space with velocity

$$v = \frac{1}{\sqrt{1 + \frac{R^2}{t^2}}}, \tag{3.29}$$

which approches the velocity of light as $t \to \infty$ quite rapidly: for example, we shall see that $R \sim 10^{-16}\ \mathrm{GeV}^{-1} \sim 10^{-40}$ s in the case of the standard model.

3.4. Potential barriers in quantum field theory

This section is devoted to a discussion of a feature of tunnelling phenomena in quantum field theory which may appear as counter-intuitive, and deserves some comments. In some cases, tunnelling from a metastable ground state to a lower minimum configuration takes place even if the scalar potential of the theory, considered as an ordinary function of the scalar fields, has no barrier in the usual sense.

This apparent paradox can be explained by observing that the tunnelling process in quantum field theory does not take place between two spatially homogeneous field configurations, but rather through the formation of a bubble of true vacuum in a false-vacuum background. Such field configuration (the bounce) is therefore not spatially homogenous, and the gradient term in the potential energy gives a non-zero contribution to the full potential energy.

In order to identify a quantity with the same properties as a potential barrier in ordinary quantum mechanics, it will be useful to cast the exponent in the semi-classical decay rate

$$B = S[\phi_b] = \int \mathrm{d}^4 x \left[\frac{1}{2}\left(\partial_\mu \phi_b\right)^2 + V(\phi_b)\right] \tag{3.30}$$

(we are now taking $V(\phi_{\mathrm{FV}}) = 0$) in a form similar to the WKB formula for the exponential factor in quantum mechanics, equation (2.165),

$$B_{\mathrm{QM}} = 2\int_0^{s_f} \mathrm{d}s\,\sqrt{2U(\mathbf{x}(s))}, \tag{3.31}$$

where $U(\mathbf{x})$ is the potential energy and the integral is taken over a suitable path $\mathbf{x} = \mathbf{x}(s)$ in configuration space, which connects the local minimum to a point beyond the potential barrier. This section is devoted to the identification of such a quantity, following an argument originally presented by Lee and Weinberg in reference [16].

Our starting point is the Lagrangian density of a scalar theory in Euclidean time

$$\mathcal{L} = -\frac{1}{2}\left(\frac{\partial\phi}{\partial\tau}\right)^2 - \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 - V(\phi) \tag{3.32}$$

and the corresponding Hamiltonian density

$$\mathcal{H} = -\frac{1}{2}\left(\frac{\partial\phi}{\partial\tau}\right)^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi). \tag{3.33}$$

We assume that the total energy

$$E[\phi] = \int \mathrm{d}^3 x\, \mathcal{H} \tag{3.34}$$

has a local minimum (the false vacuum) for a constant and uniform configuration of the field $\phi = \phi_{\mathrm{FV}}$. We define the potential energy $U[\phi]$ as the term in the total energy that does not contain time derivatives of the field:

$$U[\phi] = \int \mathrm{d}^3x \left[\frac{1}{2}(\nabla\phi)^2 + V(\phi)\right], \tag{3.35}$$

which is a functional of the field ϕ and an ordinary function of the Euclidean time τ. Then the Euclidean action can be written as

$$S[\phi] = -\int \mathrm{d}^3x\,\mathrm{d}\tau\,\mathcal{L} = \int \mathrm{d}^3x \left\{\int_{-\infty}^{\infty}\mathrm{d}\tau\,\frac{1}{2}\left(\frac{\partial\phi}{\partial\tau}\right)^2 + \int_{-\infty}^{\infty}\mathrm{d}\tau\,U[\phi]\right\}. \tag{3.36}$$

For $\phi = \phi_b$, this formula can be simplified by energy conservation: the total energy of the system in the false vacuum configuration must be the same as the energy in the bounce configuration:

$$E[\phi_{\mathrm{FV}}] = E[\phi_b], \tag{3.37}$$

which is zero by our choice of the additive constant in the scalar potential. Hence

$$E[\phi_b] = -\frac{1}{2}\int \mathrm{d}^3x \left(\frac{\partial\phi_b}{\partial\tau}\right)^2 + U[\phi_b] = 0 \tag{3.38}$$

which gives

$$U[\phi_b] = \frac{1}{2}\int \mathrm{d}^3x \left(\frac{\partial\phi_b}{\partial\tau}\right)^2. \tag{3.39}$$

Thus,

$$S[\phi_b] = \int_{-\infty}^{\infty}\mathrm{d}\tau\,2U[\phi_b]. \tag{3.40}$$

To complete the analogy with equation (3.31), we perform a change of integration variable in equation (3.40) from τ to s, defined by

$$\mathrm{d}s^2 = \int \mathrm{d}^3x\,[\mathrm{d}\phi_b(x)]^2, \tag{3.41}$$

in analogy with the infinitesimal path length defined in the case of quantum mechanics with many degrees of freedom,

$$\sum_i \frac{\mathrm{d}x_i}{\mathrm{d}s}\frac{\mathrm{d}x_i}{\mathrm{d}s} = 1; \qquad \mathrm{d}s^2 = \sum_i \mathrm{d}x_i^2. \tag{3.42}$$

We have therefore

$$\left(\frac{\mathrm{d}s}{\mathrm{d}\tau}\right)^2 = \int \mathrm{d}^3x \left(\frac{\partial\phi_b}{\partial\tau}\right)^2 = 2U[\phi_b], \tag{3.43}$$

where in the second step we have used equation (3.39). This allows us to write equation (3.40) as

$$S[\phi_b] = \int_{-\infty}^{\infty} \mathrm{d}\tau\, 2U[\phi_b] = 2\int_0^{\infty} \mathrm{d}\tau\, 2U[\phi_b] = 2\int_\gamma \mathrm{d}s\, \frac{\mathrm{d}\tau}{\mathrm{d}s}\, 2U[\phi_b]$$
$$= 2\int_\gamma \mathrm{d}s\, \sqrt{2U[\phi_b]}, \tag{3.44}$$

where γ is the path in configuration space. We have therefore achieved our goal: we have written the Euclidean action as a line integral over a path in configuration space. Correspondingly, $U[\phi_b]$ can be considered as the quantum field theory generalization of the potential barrier of ordinary quantum mechanics, as one can check by comparison with equation (3.31). It is important to emphasize that the integral is independent of the parametrization chosen for the path, which is the one which minimizes the Euclidean action.

The potential energy $U[\phi_b]$ is the sum of a potential term $U_V[\phi_b]$ and a gradient term $U_\nabla[\phi_b]$:

$$U[\phi_b] = U_{\boldsymbol{\nabla}}[\phi_b] + U_V[\phi_b] \tag{3.45}$$

$$U_{\boldsymbol{\nabla}}[\phi_b] = \int \mathrm{d}^3x\, \frac{1}{2}(\boldsymbol{\nabla}\phi_b)^2 \tag{3.46}$$

$$U_V[\phi_b] = \int \mathrm{d}^3x\, V(\phi_b). \tag{3.47}$$

The identification of $U[\phi_b]$ as the barrier in quantum field theory allows us to draw some interesting consequences.

- It is now clear why tunnelling in quantum field theory requires bounce solutions and cannot directly take place as a transition between spatially homogeneous configurations: the barrier associated with such a tunnelling process would be infinitely high. Specifically, in such cases we would find $U_\nabla[\phi_b] = 0$ and $U_V[\phi_b]$ proportional to the space volume $\mathcal{V}_3$.
- The gradient and the potential terms may have opposite signs: even if the scalar potential has no barrier separating different minima, a potential energy barrier may appear because of the gradient term. An important examples of this class is the standard model, as we shall see in section 5.6.
- The barrier penetration integral $B = S[\phi_b]$ is calculated as an integral over the three-dimensional space. In particular, the potential term is obtained by integrating the scalar potential, computed at the bounce configuration. Hence, the value of $S[\phi_b]$ is affected by all energy scales between 0 and $\phi_b(0)$. We will come back to this issue in section 5.6, when discussing the impact of non-standard physics on the metastability of the electroweak ground state.

3.5. Quantum corrections

The first quantum corrections to the leading semi-classical approximation of the decay rate of an unstable ground state are computed [2] by a generalization to quantum field theory of the analogous calculation, presented in section 2.3, of the tunnelling rate in the case of an ordinary quantum theory with many degrees of freedom. This generalization is rather straightforward, except for a few differences, which must be taken into proper account.

3.5.1. Zero modes. In the case of ordinary quantum mechanics, a factor of $\sqrt{\frac{S[x_b]-S[x_0]}{2\pi\hbar m}}$ appears as a consequence of the presence of a zero mode in the spectrum of the second derivative of the action due time translation invariance. Here we have four such factors (with $m=1$), corresponding to one zero mode for translations in each space–time direction. Correspondingly, the integration over all possible bounce centers, which yields a factor of T in the case of ordinary quantum mechanics in one space dimension, is now replaced by an integration over all possible bounce centers in space–time, yielding a factor of $T\mathcal{V}_3$, where $\mathcal{V}_3$ is the three-dimensional volume.

This is a good place for a more general discussion of zero modes of the second derivative of the action, which will turn out to play an important role in the application of the formalism to the standard model. In general, zero modes of $S''[\phi_b]$ appear whenever the theory has an invariance property under a class of transformations of the field which we define in infinitesimal form as

$$\phi(x) \to \phi'(x) = \phi(x) + \epsilon\Delta\phi(x). \tag{3.48}$$

In such case,

$$\Delta\phi_b(x) = \frac{\partial\phi'_b(x)}{\partial\epsilon} \tag{3.49}$$

is a zero mode of $S''[\phi_b]$. Here is the proof. As a consequence of the invariance of the theory upon the transformations equation (3.48), $\phi'_b(x)$ is also a bounce solution, and $S[\phi'_b] = S[\phi_b]$. Therefore

$$S''[\phi_b]\Delta\phi_b = S''[\phi'_b]\frac{\partial\phi'_b}{\partial\epsilon} = (-\partial^2 + V''(\phi'_b))\frac{\partial\phi'_b}{\partial\epsilon} = \frac{\partial}{\partial\epsilon}(-\partial^2 + V'(\phi'_b))\phi'_b = 0 \tag{3.50}$$

by the field equation. For example, under an infinitesimal translation

$$x_\mu \to x_\mu + \epsilon_\mu \tag{3.51}$$

the bounce transforms as

$$\phi_b(x) \to \phi'_b(x) = \phi_b(x+\epsilon) = \phi_b(x) + \epsilon_\mu\partial_\mu\phi_b(x). \tag{3.52}$$

If the theory is invariant under translations, then

$$\phi_\mu(x) = N\partial_\mu\phi_b(x) \tag{3.53}$$

are four zero modes of $S''[\phi_b]$. The normalization factor N is finite: equations (3.5), (3.11) and (3.12) give

$$\int d^4x\,(\partial_\mu\phi_b)^2 = 4(S[\phi_b] - S[\phi_{\rm FV}]), \tag{3.54}$$

so

$$N = \frac{1}{\sqrt{S[\phi_b] - S[\phi_{\rm FV}]}}. \tag{3.55}$$

From our experience in quantum mechanics, we know how to deal with these zero modes in the computation of the path integral: the calculation of the path integral is performed by integrating over all possible values of the coefficients c_i of the expansion of the scalar field in the basis of orthonormal eigenfunctions of $S''[\phi]$, which includes normalized zero-modes:

$$\phi(x) = \phi_b(x) + \sum_i c_i\phi_i(x) = \phi_b(x) + c_\mu N\partial_\mu\phi_b(x) + \cdots . \tag{3.56}$$

As a consequence, integrating over c_μ is equivalent to integrating over all possible locations of the bounce centre:

$$\frac{\mathrm{d}c_\mu}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}}\frac{\mathrm{d}\epsilon_\mu}{N} = \sqrt{\frac{S[\phi_b] - S[\phi_{\mathrm{FV}}]}{2\pi}}\,\mathrm{d}\epsilon_\mu. \tag{3.57}$$

These integrations contribute to the path integral a factor of

$$\left[\sqrt{\frac{S[\phi_b] - S[\phi_{\mathrm{FV}}]}{2\pi}}\right]^4 \mathcal{V}_3 T, \tag{3.58}$$

where $\mathcal{V}_3 T$ is the space–time volume; note however that the integration in the directions of zero-modes can be converted into an integration over a collective coordinate (the bounce center in this case) only provided that the zero-modes are normalizable.

One might wonder whether analogous zero modes arise from Lorentz transformations, which in Euclidean space–time correspond to four-dimensional rotations. The six functions

$$\phi_{\mu\nu}(x) = \left(x_\mu\partial_\nu - x_\nu\partial_\mu\right)\phi_b(x) \tag{3.59}$$

are in fact eigenfunction of $S''[\phi_b]$ with zero eigenvalues, as one can check directly. However, they vanish because of the above-mentioned $O(4)$ invariance of the bounce.

To conclude this section we mention two more possible origins of zero modes; both will be discussed in more detail in section 5. The first one is the presence of an internal symmetry. The simplest example is a complex scalar theory, invariant under $U(1)$ transformations

$$\phi(x) \to \phi'(x) = \phi(x) + i\epsilon\phi(x). \tag{3.60}$$

In this case and $\Delta\phi_b(x) = i\phi_b(x)$ is a zero mode. The corresponding contribution to the path integral is obtained integrating over all possible $U(1)$ transformations:

$$\frac{\mathrm{d}c}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}}\frac{\mathrm{d}\epsilon}{N} = \frac{\mathrm{d}\epsilon}{\sqrt{2\pi}}\left[\int \mathrm{d}^4x\,|\phi_b(x)|^2\right]^{-\frac{1}{2}}. \tag{3.61}$$

The integration contributes to the path integral a factor of

$$\frac{\mathcal{V}}{\sqrt{2\pi}}\left[\int \mathrm{d}^4x\,|\phi_b(x)|^2\right]^{-\frac{1}{2}}, \tag{3.62}$$

where $\mathcal{V} = 2\pi$ is the volume of the symmetry group. In this case, the bounce itself must be normalizable. A full discussion of the impact on the calculation of the path integral of internal symmetry zero modes is presented in reference [17]; quite obviously, they play a role in the case of the standard model.

As a final example, let us assume that the theory is invariant upon scale transformations, defined by

$$\phi(x) \to \phi'(x) = e^{\epsilon d}\phi(e^\epsilon x), \tag{3.63}$$

where d is the mass dimension of the field ($d = 1$ for boson fields in four space–time dimensions). In infinitesimal form

$$\phi(x) \to \phi'(x) = \phi(x) + \epsilon(1 + x_\mu \partial_\mu)\phi(x). \tag{3.64}$$

Then, by the argument given above,

$$\phi_s(x) = (1 + x_\mu \partial_\mu)\phi_b(x) \tag{3.65}$$

is a zero mode of $S''[\phi_b]$. This example is especially interesting, because the only scale-invariant theories of one scalar fields in four dimensions are specified by the scalar potential

$$V(\phi) = \frac{1}{4}\lambda\phi^4, \tag{3.66}$$

with λ a constant, which is an excellent approximation of the standard model scalar potential at large field values. Furthermore, the field equation for an $O(4)$-invariant bounce,

$$\frac{\mathrm{d}^2\phi_b(r)}{\mathrm{d}r^2} + \frac{3}{r}\frac{\mathrm{d}\phi_b(r)}{\mathrm{d}r} = \lambda\phi_b^3(r) \tag{3.67}$$

has an analytical solution for $\lambda < 0$:

$$\phi_b(r) = \frac{\phi_b(0)}{1 + \frac{|\lambda|}{8}\phi_b^2(0)r^2}, \tag{3.68}$$

usually referred to as the Fubini–Lipatov bounce, originally presented in references [18, 19] (alternative derivations of equation (3.68) can be found in reference [16] and in reference [20]; the latter is reviewed in appendix A.5).

Unfortunately, the zero mode of scale transformations

$$\phi_s(r) = (1 + x_\mu \partial_\mu)\phi_b(x) = \phi_b(0)\frac{1 - \frac{|\lambda|}{8}\phi_b^2(0)r^2}{\left(1 + \frac{|\lambda|}{8}\phi_b^2(0)r^2\right)^2} \tag{3.69}$$

is not normalizable: $\phi_s^2(r) \sim r^{-4}$ for $r \to \infty$. It follows that the corresponding contribution to the path integral cannot be replaced by an integration over a collective coordinate, as in the case of translation zero modes.

3.5.2. Renormalization. A second important difference between quantum field theory and ordinary quantum mechanics is the appearance of ultraviolet divergences in the computation of the action beyond the semi-classical approximation. Ultraviolet divergences require the usual renormalization procedure; specifically, the classical action must be expressed in terms of renormalized fields and parameters, and supplemented by appropriate counterterms:

$$S[\phi] \to S[\phi] + \hbar S_{\mathrm{ct}}[\phi], \tag{3.70}$$

where $S_{\mathrm{ct}}[\phi]$ is the action induced by one-loop counterterms, and we have (temporarily) restored the dependence on $\hbar$ for later convenience. In general, the presence of the counterterms will modify the bounce solution:

$$\phi_b \to \phi_b + \hbar\phi_{\mathrm{ct}}, \tag{3.71}$$

where ϕ_b is a bounce solution for $S[\phi]$:

$$\left.\frac{\delta S[\phi]}{\delta \phi}\right|_{\phi=\phi_b} = 0. \tag{3.72}$$

However, to one-loop order, the explicit expression of the correction ϕ_{ct} is not needed. Indeed,

$$\begin{aligned} S[\phi_b + \hbar\phi_{\mathrm{ct}}] &+ \hbar S_{\mathrm{ct}}[\phi_b + \hbar\phi_{\mathrm{ct}}] \\ &= S[\phi_b] + \hbar \int \mathrm{d}^4x\, \phi_{\mathrm{ct}} \left.\frac{\delta S[\phi]}{\delta\phi}\right|_{\phi=\phi_b} + \hbar S_{\mathrm{ct}}[\phi_b] + O(\hbar^2) \\ &= S[\phi_b] + \hbar S_{\mathrm{ct}}[\phi_b] + O(\hbar^2). \end{aligned} \tag{3.73}$$

Taking these differences into proper account, and setting back $\hbar = 1$, we obtain the following result for the decay probability per unit time and per unit volume $\gamma/\mathcal{V}_3$ of a metastable state:

$$\frac{\gamma}{\mathcal{V}_3} = \frac{(S[\phi_b] - S[\phi_{\mathrm{FV}}])^2}{(2\pi)^2} \sqrt{\frac{\mathrm{Det}\, S''[\phi_{\mathrm{FV}}]}{|\mathrm{Det}'\, S''[\phi_b]|}}\, e^{-(S[\phi_b]-S[\phi_{\mathrm{FV}}])} e^{-(S_{\mathrm{ct}}[\phi_b]-S_{\mathrm{ct}}[\phi_{\mathrm{FV}}])}. \tag{3.74}$$

Equation (3.74) is dimensionally correct: since there are four zero modes, the pre-exponential factor has the dimension of the square root of $S''[\phi]$ (an inverse length) to the fourth power.

4. The effective potential

The appropriate tool to determine the ground state of a quantum field theory beyond tree level is the effective potential [21, 22], since it automatically encodes radiative corrections while retaining the advantage of the semi-classical approximation, namely the possibility of simultaneously surveying all the vacua of the theory by minimizing a function in analogy to the tree-level potential. Furthermore, we have seen in the previous sections that functional methods are extremely useful in the study of the lifetime of metastable states in quantum field theory.

In this section we briefly review the effective action formalism and apply it to the case of the standard model of electroweak interactions, in order to study the possible instability of the ground state at large field values.

4.1. Effective action and effective potential

Let us consider the theory of a single real scalar field Φ described by the Lagrangian density $\mathcal{L}(\Phi, \partial_\mu\Phi)$ with a linear coupling of $\Phi(x)$ to an external source $J(x)$:

$$\mathcal{L} \to \mathcal{L} + J(x)\Phi(x). \tag{4.1}$$

The vacuum-to-vacuum transition amplitude $Z[J]$ is usually written in terms of a functional $W[J]$ as

$$Z[J] = \langle 0|0\rangle_J \equiv \mathrm{e}^{\mathrm{i}W[J]} = \int \mathcal{D}\Phi\, \mathrm{e}^{\mathrm{i}\int \mathrm{d}^4x\, [\mathcal{L}(\Phi,\partial_\mu\Phi)+J\Phi]}, \tag{4.2}$$

where in the last step we have employed the standard path integral representation. Functional derivatives of $W[J]$ with respect to iJ at $J = 0$ give the connected Green's functions of the theory; for this reason, $W[J]$ is called the connected generating functional. $W[J]$ admits a functional Taylor expansion

$$W[J] = \sum_n \frac{i^n}{n!} \int \mathrm{d}^4 x_1 \ \ldots \ \mathrm{d}^4 x_n G^{(n)}(x_1, \ldots, x_n) J(x_1) \ldots J(x_n), \tag{4.3}$$

where $G^{(n)}$ denotes the n-point connected Green's functions. These, in turn, are computed in perturbation theory as the sum of all connected Feynman diagrams with n external lines.

Next, one then defines the *classical field* $\phi_c(x)$ as

$$\phi_c(x) \equiv \frac{\delta W[J]}{\delta J(x)} = \frac{\langle 0|\Phi(x)|0\rangle_J}{\langle 0|0\rangle_J}, \tag{4.4}$$

where the last expression follows from the path integral representation in equation (4.2). Note that $\phi_c(x)$ is by definition a functional of J; we are implicitly assuming that this functional is single-valued. This is related to the convexity issue of the effective potential, to be discussed in section 4.6. We now introduce the *effective action* $\Gamma[\phi_c]$, defined by the functional Legendre transformation

$$\Gamma[\phi_c] = W[J] - \int \mathrm{d}^4 x \ J(x)\phi_c(x). \tag{4.5}$$

The effective action can be expanded in powers of the classical fields, in a way similar to equation (4.3), yielding

$$\Gamma[\phi_c] = \sum_{n=0}^{\infty} \frac{1}{n!} \int \mathrm{d}x_1 \ldots \mathrm{d}x_n \Gamma^{(n)}(x_1, \ldots, x_n)\phi_c(x_1) \ldots \phi_c(x_n). \tag{4.6}$$

The coefficients $\Gamma^{(n)}$ can be shown to correspond to the one-particle-irreducible (1PI) Green's functions of the theory (see for example reference [11]).

The functional $\Gamma[\phi_c]$ is the appropriate tool to study spontaneous symmetry breaking. Indeed, the condition for spontaneous symmetry breaking is that ϕ_c is different from zero even when the source J is equal to zero, as can be read off equation (4.4). On the other hand, one has

$$\frac{\delta\Gamma[\phi_c]}{\delta\phi_c(x)} = \int \mathrm{d}^4 y \ \frac{\delta W[J]}{\delta J(y)} \frac{\delta J(y)}{\delta\phi_c(x)} - J(x) - \int \mathrm{d}^4 y \ \frac{\delta J(y)}{\delta\phi_c(x)} \phi_c(y) = -J(x), \tag{4.7}$$

or in other words

$$\left.\frac{\delta\Gamma[\phi_c]}{\delta\phi_c(x)}\right|_{J=0} = 0. \tag{4.8}$$

We conclude that spontaneous symmetry breaking takes place when the classical field that minimizes the effective action is different from zero.

The effective action can be computed in perturbation theory by the usual technique of Feynman diagram. In particular, it can be shown that in the tree-level approximation the effective action coincides with the classical action:

$$\Gamma[\phi_c] = \int \mathrm{d}^4 x \left[\mathcal{L}(\phi_c, \partial_\mu\phi_c) + J\phi_c\right] + \text{loop corrections}. \tag{4.9}$$

In some cases, however, an expansion in powers of external momenta is more useful. To define such an expansion, we consider the Fourier transforms of the functions $\Gamma^{(n)}(x_1, \ldots, x_n)$:

$$\Gamma^{(n)}(x_1, \ldots, x_n) = \int \frac{\mathrm{d}^4 p_1}{(2\pi)^4} \ldots \frac{\mathrm{d}^4 p_n}{(2\pi)^4} \mathrm{e}^{\mathrm{i}(p_1 x_1 + \cdots + p_n x_n)} \\ \times (2\pi)^4 \delta(p_1 + \cdots + p_n) \tilde{\Gamma}^{(n)}(p_1, \ldots, p_n), \tag{4.10}$$

and expand $\tilde{\Gamma}^{(n)}$ in powers of momenta around $p_i = 0$:

$$\tilde{\Gamma}^{(n)}(p_1, \ldots, p_n) = \tilde{\Gamma}^{(n)}(0) + \cdots . \tag{4.11}$$

The effective action becomes

$$\begin{aligned} \Gamma[\phi_c] = & \sum_{n=0}^{\infty} \frac{1}{n!} \int \mathrm{d}x_1 \ldots \mathrm{d}x_n \, \phi_c(x_1) \ldots \phi_c(x_n) \\ & \times \int \frac{\mathrm{d}^4 p_1}{(2\pi)^4} \ldots \frac{\mathrm{d}^4 p_n}{(2\pi)^4} \mathrm{e}^{\mathrm{i}(p_1 x_1 + \cdots + p_n x_n)} \int \mathrm{d}^4 x \, \mathrm{e}^{-\mathrm{i}x(p_1 + \cdots + p_n)} \\ & \times \left[\tilde{\Gamma}^{(n)}(0) + \cdots \right] \\ = & \int \mathrm{d}^4 x \sum_{n=0}^{\infty} \frac{1}{n!} \tilde{\Gamma}^{(n)}(0) \phi_c^n(x) + \cdots \end{aligned} \tag{4.12}$$

The first term in this expansion is usually written as

$$-\int \mathrm{d}^4 x \, V_{\mathrm{eff}}(\phi_c), \tag{4.13}$$

where

$$V_{\mathrm{eff}}(\phi_c) = -\sum_{n=0}^{\infty} \frac{1}{n!} \tilde{\Gamma}^{(n)}(0) \phi_c^n \tag{4.14}$$

is an ordinary function of ϕ_c, usually called the *effective potential* of the theory, since it does not contain derivatives of the classical field. The following terms, originating from higher powers of momenta in the expansion of $\tilde{\Gamma}^{(n)}$, contain instead two or more derivatives of ϕ_c.

If we require translational invariance of the vacuum state, then $\partial_\mu \phi_c = 0$ for $J = 0$, and the minimum condition equation (4.8) reduces to

$$\frac{\mathrm{d}V_{\mathrm{eff}}(\phi_c)}{\mathrm{d}\phi_c} = 0. \tag{4.15}$$

The advantage of equation (4.15) is that the ground state of the theory can be determined, including quantum corrections, via the minimization of a single ordinary function, the effective potential, in analogy with the tree-level case.

4.2. The background field method

The discussion of the previous section shows that efficient techniques for the computation of the effective potential are needed. A compact way to compute the loop-expanded effective potential[6] is given by the *background field method* [23], which we review here for the one-loop case. Let us consider the Lagrangian density for a real scalar field Φ:

$$\mathcal{L}(\Phi, \partial_\mu \Phi) = \frac{1}{2}\partial_\mu \Phi \partial^\mu \Phi - V(\Phi), \tag{4.16}$$

with $V(\Phi)$ denoting the tree-level scalar potential. We perform the following change of variable in the functional integral equation (4.2):

$$\Phi(x) = \phi_c(x) + \phi(x), \tag{4.17}$$

where $\phi_c(x)$ is a background field which we identify with the classical field of the previous section, while $\phi(x)$ is a dynamical (quantum) field. Organized in powers of ϕ (up to quadratic terms), the exponent of the rhs of equation (4.2) takes the form

$$\begin{aligned}
&\int \mathrm{d}^4x\, [\mathcal{L}(\phi_c + \phi) + J(\phi_c + \phi)] \\
&\quad = \int \mathrm{d}^4x\, [\mathcal{L}(\phi_c) + J\phi_c] + \int \mathrm{d}^4x\, \phi(x) \left[\frac{\delta \mathcal{L}}{\delta \Phi}\bigg|_{\Phi=\phi_c} + J \right] \\
&\qquad + \frac{1}{2} \int \mathrm{d}^4x\, \mathrm{d}^4y\, \phi(x) i\mathscr{D}^{-1}(\phi_c, x-y)\phi(y) + \mathcal{O}(\phi^3),
\end{aligned} \tag{4.18}$$

where

$$\begin{aligned}
i\mathscr{D}^{-1}(\phi_c, x-y) &= \frac{\delta}{\delta\Phi(x)} \frac{\delta}{\delta\Phi(y)} \int \mathrm{d}^4z\, [\mathcal{L}(\Phi) + J(z)\Phi(z)] \bigg|_{\Phi=\phi_c} \\
&= \left[-\partial^2 - V''(\phi_c(x)) \right] \delta^{(4)}(x-y)
\end{aligned} \tag{4.19}$$

is the inverse propagator in configuration space evaluated at $\Phi = \phi_c$. We can now explicitly perform the path integral in the Gaussian approximation:

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}W[J]} &\simeq \mathrm{e}^{\mathrm{i}\int \mathrm{d}^4x\, [\mathcal{L}(\phi_c) + J\phi_c]} \int \mathcal{D}\phi\, \exp\left(\frac{i}{2} \int \mathrm{d}^4x\, \mathrm{d}^4y\, \phi(x) i\mathscr{D}^{-1}(\phi_c, x-y)\phi(y) \right) \\
&= \mathcal{N} \mathrm{e}^{\mathrm{i}\int \mathrm{d}^4x\, [\mathcal{L}(\phi_c) + J\phi_c]} \left[\det\, i\mathscr{D}^{-1}(\phi_c, x-y) \right]^{-\frac{1}{2}},
\end{aligned} \tag{4.20}$$

up to an irrelevant normalization constant $\mathcal{N}$. The result equation (4.20) follows from the invariance of the path integral measure, $\mathcal{D}\Phi = \mathcal{D}\phi$, and the classical field equation $\frac{\delta\mathcal{L}}{\delta\Phi}\big|_{\phi_c} + J = 0$. The latter approximation is sufficient as long as we are interested in the one-loop effective

[6] The loop expansion corresponds to an expansion in $\hbar$ which does not affect the shift of the fields, since $\hbar$ multiplies the total Lagrangian density. In order to show this let us restore $\hbar$ in the definition of the Lagrangian, $\mathcal{L}(\phi, \partial_\mu\phi) \to \frac{1}{\hbar}\mathcal{L}(\phi, \partial_\mu\phi)$, and denote with P the power of $\hbar$ associated with a given Feynman diagram. Each vertex carries a factor $\hbar^{-1}$, while each propagator (being the inverse of the quadratic part of $\mathcal{L}$) carries a factor $\hbar$. Denoting by I the number of propagators and by V the number of vertices of a graph, we have $P = I - V$. On the other hand, the number of loops L in a diagram corresponds to the number of independent momenta: every internal line contribute to one integration momentum, while every vertex to a δ function which reduces by one unit the number of integration momenta (except for an overall δ which is left for energy–momentum conservation), which yields $L = I - V + 1$. We can hence establish an equivalence between the $\hbar$ and loop expansions: $P = L - 1$.

action. After replacing equation (4.20) in the definition of the effective action, equation (4.5), we obtain

$$\Gamma[\phi_c] = S[\phi_c] + \frac{i}{2}\,\log\,\mathrm{Det}\; i\mathscr{D}^{-1}(\phi_c, x-y), \tag{4.21}$$

where

$$S[\phi_c] = \int \mathrm{d}^4x\,\mathcal{L}(\phi_c, \partial_\mu\phi_c) \tag{4.22}$$

is the classical action. The calculation of one-loop corrections to the effective action involves the usual renormalization procedure: although not explicitly shown in equation (4.21), it is understood that the final result involves the choice of a regularization procedure and of a renormalization scheme, and the introduction of suitable counterterms. Typically, this results in the dependence of the renormalized parameters on an arbitrary energy scale, to compensate the explicit dependence of the correction term on the same scale.

The effective potential is now immediately obtained by evaluating equation (4.21) for a space–time-independent classical field configuration ϕ_c (see equation (4.12)). We get

$$-\mathcal{V}_4\;V_{\mathrm{eff}}(\phi_c) = -\mathcal{V}_4\;V_0(\phi_c) + \frac{i}{2}\,\log\,\mathrm{Det}\; i\mathscr{D}^{-1}(\phi_c, x-y), \tag{4.23}$$

where $\mathcal{V}_4 \equiv \int \mathrm{d}^4x$ denotes the space–time volume, originating from the fact that ϕ_c is space–time independent. The homogeneity of ϕ_c also allows for a simple evaluation of the functional determinant, through the identity

$$\log\,\mathrm{Det}\; i\mathscr{D}^{-1}(\phi_c, x-y) = \mathrm{tr}\;\log\, i\mathscr{D}^{-1}(\phi_c, x-y). \tag{4.24}$$

The functional trace, denoted by tr, is taken by setting $x = y$ and integrating over the space–time. This gives

$$\log\,\mathrm{Det}\; i\mathscr{D}^{-1}(\phi_c, x-y) = \mathcal{V}_4\,\log\, i\mathscr{D}^{-1}(\phi_c, 0) = \mathcal{V}_4 \int \frac{\mathrm{d}^4k}{(2\pi)^4}\,\log\, i\tilde{\mathscr{D}}^{-1}(\phi_c, k), \tag{4.25}$$

where $\tilde{\mathscr{D}}^{-1}(\phi_c, k)$ denotes the four-dimensional Fourier transform of $\mathscr{D}^{-1}(\phi_c, x-y)$.

An explicit example is provided by a renormalizable scalar theory, with

$$V_0(\Phi) = \frac{1}{2}m^2\Phi^2 + \frac{1}{4}\lambda\Phi^4. \tag{4.26}$$

In this case

$$i\mathscr{D}^{-1}(\phi_c, x-y) = \left(-\partial_\mu\partial^\mu - m^2 - 3\lambda\phi_c^2\right)\delta^{(4)}(x-y) \tag{4.27}$$

and

$$i\tilde{\mathscr{D}}^{-1}(\phi_c, k) = k^2 - m^2 - 3\lambda\phi_c^2. \tag{4.28}$$

The background field method can be consistently extended to include higher orders in the loop expansion (see e.g. reference [24] for a state-of-the-art calculation of the three-loop effective potential). In this review we will limit ourselves to the explicit calculation of the effective potential at one loop.

4.3. The one-loop effective potential of the standard model

The results of the previous section can be readily generalized to the case of a generic quantum field interacting with Φ, thereby obtaining the following closed form for the one-loop effective potential:

$$V_{\text{eff}}(\phi) = V_0(\phi) + i\sum_n \eta_n \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \,\mathrm{Tr}\, \log\, i\tilde{\mathscr{D}}_n^{-1}(\phi, k), \tag{4.29}$$

where the sum is extended to all fields in the theory, labelled by the index n, and $i\tilde{\mathscr{D}}_n^{-1}(\phi_c, k)$ denotes the corresponding inverse propagators of the dynamical fields after the background field ϕ_c shift in Φ. The trace acts on all internal indices (e.g. Lorentz or gauge); it is therefore the conventional trace in finite-dimension linear spaces, and we denote it by the usual symbol Tr to distinguish it from the functional trace, denoted by tr. The factor η_n is the power of the functional determinant due to the Gaussian path integral; it takes the values $-\frac{1}{2}$ for bosonic fields, and $+1$ for fermionic fields (matter fermions or ghosts). Finally, we have denoted the classical field ϕ_c by ϕ to simplify notations.

We are now ready to apply the background field method to the calculation of the one-loop effective potential of the standard model of electroweak interactions. In order to set the notation, we split the classical Lagrangian density of the electroweak sector into a gauge term, a Higgs term and Yukawa term, the latter including only the leading contribution from the top quark:

$$\mathcal{L}_{\mathrm{C}} = \mathcal{L}_{\mathrm{YM}} + \mathcal{L}_{\mathrm{H}} + \mathcal{L}_{\mathrm{top}}, \tag{4.30}$$

with

$$\mathcal{L}_{\mathrm{YM}} = -\frac{1}{4}\left(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon^{abc} W_\mu^b W_\nu^c\right)^2 - \frac{1}{4}\left(\partial_\mu B_\nu - \partial_\nu B_\mu\right)^2, \tag{4.31}$$

$$\mathcal{L}_{\mathrm{H}} = \left(D_\mu H\right)^\dagger (D^\mu H) - V(H), \tag{4.32}$$

$$\mathcal{L}_{\mathrm{top}} = \overline{Q}_{\mathrm{L}} i \not{D} Q_{\mathrm{L}} + \bar{t}_{\mathrm{R}} i \not{D} t_{\mathrm{R}} + \left(-y_t \overline{Q}_{\mathrm{L}} \epsilon H^* t_{\mathrm{R}} + \text{h.c.}\right), \tag{4.33}$$

where W_μ^a $(a = 1, 2, 3)$ and B_μ are the $SU(2)_L$ and $U(1)_Y$ gauge fields, H is the standard model Higgs doublet, $\epsilon = i\sigma^2$ is the 2×2 antisymmetric matrix, and

$$Q_{\mathrm{L}} = \begin{pmatrix} t_{\mathrm{L}} \\ b_{\mathrm{L}} \end{pmatrix} \tag{4.34}$$

is the left-handed quark doublet of the third generation. QCD indices are suppressed in the quark sector. The covariant derivative is defined by

$$D_\mu = \partial_\mu - igT^a W_\mu^a - ig' Y B_\mu, \tag{4.35}$$

where T^a are the $SU(2)_L$ generators in the relevant representation ($T^a = \sigma^a/2$ with σ^a the three Pauli matrices for doublets, $T^a = 0$ for singlets), and Y is the hypercharge of the standard model fields. We have $Y = Q - T^3$, where Q is the electric charge in units of the proton charge, and hence $Y(H) = 1/2$, $Y(Q_{\mathrm{L}}) = 1/6$, $Y(t_{\mathrm{R}}) = 2/3$. The Higgs potential is

$$V(H) = m^2 H^\dagger H + \lambda (H^\dagger H)^2, \tag{4.36}$$

with $m^2 < 0$. Gauge invariance allows us to perform the shift of the Higgs doublet, required by the background-field-method calculation of the effective potential, in a specific direction in the $SU(2)_L \times U(1)_Y$ space:

$$H(x) \to \frac{1}{\sqrt{2}} \begin{pmatrix} \chi^1(x) + i\chi^2(x) \\ \phi + h(x) + i\chi^3(x) \end{pmatrix}, \tag{4.37}$$

where ϕ denotes the classical background field, h the physical Higgs field and χ^a ($a = 1, 2, 3$) the Goldstone fields. At tree-level the effective potential simply reads

$$V_0(\phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4, \tag{4.38}$$

while in order to compute the one-loop quantum correction, $V_1(\phi)$, according to equation (4.29) one needs to work out the inverse propagators of the dynamical fields in the shifted standard model Lagrangian, including the gauge fixing term. Although the calculation is most easily carried out in the Landau gauge, it is instructive to discuss a more general gauge. A convenient choice is the so-called Fermi gauge in which the gauge fixing term is

$$\mathcal{L}^{\text{Fermi}}_{\text{g.f.}} = -\frac{1}{2\xi_W}\left(\partial^\mu W^a_\mu\right)^2 - \frac{1}{2\xi_B}\left(\partial^\mu B_\mu\right)^2. \tag{4.39}$$

The Landau gauge is obtained for $\xi_W = \xi_B = 0$. We start from the determination of the quadratic (ϕ-dependent) term of the Lagrangian, $\mathcal{L}^{(2)} + \mathcal{L}^{\text{Fermi}}_{\text{g.f.}}$, after the shift in equation (4.37). A straightforward calculation yields

$$\mathcal{L}^{(2)}_{\text{YM}} = \frac{1}{2}W^a_\mu\left(\partial^2 g^{\mu\nu} - \partial^\mu\partial^\nu\right)\delta^{ab}W^b_\nu + \frac{1}{2}B_\mu\left(\partial^2 g^{\mu\nu} - \partial^\mu\partial^\nu\right)B_\nu, \tag{4.40}$$

$$\begin{aligned}\mathcal{L}^{(2)}_{\text{H}} = {} & \frac{1}{2}h\left(-\partial^2 - \overline{m}^2_h\right)h + \frac{1}{2}\chi^a\left(-\partial^2 - \overline{m}^2_\chi\right)\delta^{ab}\chi^b \\ & + \frac{1}{2}\overline{m}^2_W W^a_\mu W^{a\mu} + \frac{1}{2}\overline{m}^2_B B_\mu B^\mu - \overline{m}_W\overline{m}_B W^3_\mu B^\mu \\ & - \overline{m}_W\partial_\mu\chi^1 W^{2\mu} - \overline{m}_W\partial_\mu\chi^2 W^{1\mu} + \overline{m}_W\partial_\mu\chi^3 W^{3\mu} - \overline{m}_B\partial_\mu\chi^3 B^\mu,\end{aligned} \tag{4.41}$$

$$\mathcal{L}^{(2)}_{\text{top}} = \overline{t}(i\,\partial\!\!\!/ - \overline{m}_t)t, \tag{4.42}$$

where we have defined ϕ-dependent masses

$$\overline{m}^2_h = m^2 + 3\lambda\phi^2, \tag{4.43}$$

$$\overline{m}^2_\chi = m^2 + \lambda\phi^2, \tag{4.44}$$

$$\overline{m}_W = \frac{1}{2}g\phi, \tag{4.45}$$

$$\overline{m}_B = \frac{1}{2}g'\phi, \tag{4.46}$$

$$\overline{m}_t = \frac{y_t}{\sqrt{2}}\phi. \tag{4.47}$$

Note that $\mathcal{L}_{\text{g.f.}}^{\text{Fermi}}$ is already quadratic in the gauge boson fields, while bilinear ghost terms are ϕ-independent[7]. Hence, in the Fermi gauge the ghost contribution decouples from the one-loop effective potential.

A slight complication in the Fermi gauge arises in the Goldstone–gauge boson mixing in equation (4.41). To deal with it we define an extended field vector

$$X^{\text{T}} = \left(V_\mu^{\text{T}}, \chi^{\text{T}}\right), \tag{4.48}$$

where

$$V_\mu^{\text{T}} = \left(W_\mu^1, W_\mu^2, W_\mu^3, B_\mu\right), \qquad \chi^{\text{T}} = \left(\chi^1, \chi^2, \chi^3\right). \tag{4.49}$$

The quadratic part of the Goldstone–gauge sector can be cast into

$$\frac{1}{2} X^{\text{T}}\left(i\mathscr{D}_X^{-1}\right)X = \frac{1}{2}\left(V_\mu^{\text{T}}, \chi^{\text{T}}\right)\begin{pmatrix} i\left(\mathscr{D}_V^{-1}\right)^\mu_{\ \nu} & \overline{m}_{\text{mix}}^{\text{T}}\,\partial^\mu \\ -\overline{m}_{\text{mix}}\,\partial_\nu & i\mathscr{D}_\chi^{-1}\end{pmatrix}\begin{pmatrix} V^\nu \\ \chi \end{pmatrix}, \tag{4.50}$$

with

$$\overline{m}_{\text{mix}} = \begin{pmatrix} 0 & -\overline{m}_W & 0 & 0 \\ -\overline{m}_W & 0 & 0 & 0 \\ 0 & 0 & \overline{m}_W & -\overline{m}_B \end{pmatrix}. \tag{4.51}$$

After Fourier transformation ($\partial_\mu \to ik_\mu$) the inverse propagator matrix becomes

$$i\tilde{\mathscr{D}}_X^{-1} = \begin{pmatrix} i(\tilde{\mathscr{D}}_V^{-1})^\mu_\nu & ik^\mu \overline{m}_{\text{mix}}^{\text{T}} \\ -ik_\nu \overline{m}_{\text{mix}} & i\tilde{\mathscr{D}}_\chi^{-1} \end{pmatrix}, \tag{4.52}$$

where $(\tilde{\mathscr{D}}_V^{-1})^\mu_\nu$ can be conveniently split into a transversal and a longitudinal part

$$(\tilde{\mathscr{D}}_V^{-1})^\mu_\nu = i\tilde{\mathscr{D}}_T^{-1}\,(\Pi_T)^\mu_\nu + i\tilde{\mathscr{D}}_L^{-1}\,(\Pi_L)^\mu_\nu, \tag{4.53}$$

where

$$(\Pi_T)^\mu_\nu = g^\mu_\nu - \frac{k^\mu k_\nu}{k^2}, \qquad (\Pi_L)^\mu_\nu = \frac{k^\mu k_\nu}{k^2}, \tag{4.54}$$

and

$$i\tilde{\mathscr{D}}_T^{-1} = \begin{pmatrix} -k^2 + \overline{m}_W^2 & 0 & 0 & 0 \\ 0 & -k^2 + \overline{m}_W^2 & 0 & 0 \\ 0 & 0 & -k^2 + \overline{m}_W^2 & -\overline{m}_W \overline{m}_B \\ 0 & 0 & -\overline{m}_W \overline{m}_B & -k^2 + \overline{m}_B^2 \end{pmatrix}, \tag{4.55}$$

[7] Recall that the ghost Lagrangian is given by $\mathcal{L}_{\text{ghost}} = \sum_{\alpha\beta} \eta_\alpha^\dagger \frac{\delta F^\alpha}{\delta\theta^\beta} \eta_\beta$, where $F_W^a = \partial^\mu W_\mu^a$ and $F_B = \partial^\mu B_\mu$ are gauge-fixing functionals, $\delta W_\mu^a = g\epsilon^{abc}\theta^b W_\mu^c - \partial_\mu \theta^a$ and $\delta B_\mu = -\partial_\mu \theta^B$ are gauge transformations, θ^a and θ^B the corresponding gauge parameters.

$$i\tilde{\mathscr{D}}_L^{-1} = \begin{pmatrix} -\xi_W^{-1}k^2 + \overline{m}_W^2 & 0 & 0 & 0 \\ 0 & -\xi_W^{-1}k^2 + \overline{m}_W^2 & 0 & 0 \\ 0 & 0 & -\xi_W^{-1}k^2 + \overline{m}_W^2 & -\overline{m}_W\overline{m}_B \\ 0 & 0 & -\overline{m}_W\overline{m}_B & -\xi_B^{-1}k^2 + \overline{m}_B^2 \end{pmatrix}. \tag{4.56}$$

The Goldstone boson, Higgs and top quark inverse propagators are

$$i\tilde{\mathscr{D}}_\chi^{-1} = \begin{pmatrix} k^2 - \overline{m}_\chi^2 & 0 & 0 \\ 0 & k^2 - \overline{m}_\chi^2 & 0 \\ 0 & 0 & k^2 - \overline{m}_\chi^2 \end{pmatrix}, \tag{4.57}$$

$$i\tilde{\mathscr{D}}_h^{-1} = k^2 - \overline{m}_h^2, \tag{4.58}$$

$$i\tilde{\mathscr{D}}_t^{-1} = \not{k} - \overline{m}_t. \tag{4.59}$$

The next step is the evaluation of log det $i\tilde{\mathscr{D}}_n^{-1}$, for $n = X, h, t$. Let us express the determinant of the block matrix in equation (4.52) as

$$\begin{aligned} \det\, i\tilde{\mathscr{D}}_X^{-1} &= \det\, i\tilde{\mathscr{D}}_\chi^{-1} \det\left(i(\tilde{\mathscr{D}}_V^{-1})^\mu_\nu - k^\mu k_\nu \overline{m}_{\text{mix}}^{\text{T}} \left(i\tilde{\mathscr{D}}_\chi^{-1} \right)^{-1} \overline{m}_{\text{mix}} \right) \\ &= \det\, i\tilde{\mathscr{D}}_\chi^{-1} \det\Bigg(i\tilde{\mathscr{D}}_{\text{T}}^{-1} (\Pi_T)^\mu_\nu \\ &\quad + \left(i\tilde{\mathscr{D}}_L^{-1} - k^2 \overline{m}_{\text{mix}}^{\text{T}} \left(i\tilde{\mathscr{D}}_\chi^{-1} \right)^{-1} \overline{m}_{\text{mix}} \right) (\Pi_L)^\mu_\nu \Bigg), \end{aligned} \tag{4.60}$$

where in the last step we used the decomposition in equation (4.53). It is convenient to perform a Lorentz transformation in d space–time dimensions, $k_\mu \to (k_0, 0, 0, 0, \ldots)$, such that $(\Pi_L)^\mu_\nu \to \text{diag}(1, 0, 0, 0, \ldots)$ and $(\Pi_T)^\mu_\nu \to \text{diag}(0, 1, 1, 1, \ldots)$. Then, using the Lorentz invariance of the determinant, we obtain

$$\begin{aligned} \log\, \det\, i\tilde{\mathscr{D}}_X^{-1} &= (d-1) \log\det \tilde{\mathscr{D}}_T^{-1} \\ &\quad + \log\, \det\, i\tilde{\mathscr{D}}_\chi^{-1} \det\left(i\tilde{\mathscr{D}}_L^{-1} - k^2 \overline{m}_{\text{mix}}^{\text{T}} \left(i\tilde{\mathscr{D}}_\chi^{-1} \right)^{-1} \overline{m}_{\text{mix}} \right). \end{aligned} \tag{4.61}$$

The two terms in the rhs of equation (4.61) yield respectively

$$\log\, \det\, i\tilde{\mathscr{D}}_T^{-1} = 2\, \log\left(-k^2 + \overline{m}_W^2 \right) + \log\left(-k^2 + \overline{m}_Z^2 \right) + \cdots \tag{4.62}$$

and

$$\begin{aligned} &\log\, \det\, i\tilde{\mathscr{D}}_\chi^{-1} \det\left(i\tilde{\mathscr{D}}_L^{-1} - k^2 \overline{m}_{\text{mix}}^{\text{T}} \left(i\tilde{\mathscr{D}}_\chi^{-1} \right)^{-1} \overline{m}_{\text{mix}} \right) \\ &\quad = 2\, \log\left(k^4 - k^2 \overline{m}_\chi^2 + \overline{m}_\chi^2 \xi_W \overline{m}_W^2 \right) \\ &\qquad + \log\left(k^4 - k^2 \overline{m}_\chi^2 + \overline{m}_\chi^2 (\xi_W \overline{m}_W^2 + \xi_B \overline{m}_B^2) \right) + \cdots \\ &\quad = 2\, \log\left(k^2 - \overline{m}_{A^+}^2 \right) + 2\, \log\left(k^2 - \overline{m}_{A^-}^2 \right) \\ &\qquad + \log\left(k^2 - \overline{m}_{B^+}^2 \right) + \log\left(k^2 - \overline{m}_{B^-}^2 \right) + \cdots, \end{aligned} \tag{4.63}$$

where the ellipses stand for ϕ-independent terms, and we have defined the additional ϕ-dependent masses

$$\overline{m}_Z^2 = \overline{m}_W^2 + \overline{m}_B^2, \tag{4.64}$$

$$\overline{m}_{A\pm}^2 = \frac{1}{2}\overline{m}_\chi\left(\overline{m}_\chi \pm \sqrt{\overline{m}_\chi^2 - 4\xi_W\overline{m}_W^2}\right), \tag{4.65}$$

$$\overline{m}_{B\pm}^2 = \frac{1}{2}\overline{m}_\chi\left(\overline{m}_\chi \pm \sqrt{\overline{m}_\chi^2 - 4(\xi_W\overline{m}_W^2 + \xi_B\overline{m}_B^2)}\right). \tag{4.66}$$

For the evaluation of the fermionic determinant of equation (4.59) it is sufficient to employ a naive treatment of γ_5 in dimensional regularization, i.e. $\{\gamma_5, \gamma_\mu\} = 0$ in d dimensions, and make the standard choice $\mathrm{Tr}\,\mathbf{1}_{\mathrm{Dirac}} = 4$ in d dimensions[8]. Explicitly, one has

$$\begin{aligned}\log\,\det(\not{k} - \overline{m}_t) &= \mathrm{Tr}\ \log(\not{k} - \overline{m}_t) = \mathrm{Tr}\ \log\,\gamma^5(\not{k} - \overline{m}_t)\gamma^5 = \mathrm{Tr}\ \log(-\not{k} - \overline{m}_t)\\ &= \frac{1}{2}\left[\mathrm{Tr}\ \log(\not{k} - \overline{m}_t) + \mathrm{Tr}\ \log(-\not{k} - \overline{m}_t)\right]\\ &= \frac{1}{2}\,\mathrm{Tr}\ \log\left(-k^2 + \overline{m}_t^2\right) = \frac{1}{2}4\times 3\ \log\left(-k^2 + \overline{m}_t^2\right),\end{aligned} \tag{4.67}$$

where the extra factors in the last step are due to the trace over Dirac and colour indices.

Including all the relevant degrees of freedom and working in dimensional regularization with $d = 4 - 2\epsilon$, the one-loop contribution to the effective potential equation (4.29) can be cast in the form

$$\begin{aligned}V_1(\phi,\epsilon) = -\frac{i}{2}\mu^{2\epsilon}\int\frac{\mathrm{d}^d k}{(2\pi)^d}&\left[-12\ \log\left(-k^2 + \overline{m}_t^2\right)\right.\\ &+ (d-1)\left(2\ \log\left(-k^2 + \overline{m}_W^2\right)\right.\\ &\left.+ \log\left(-k^2 + \overline{m}_Z^2\right)\right) + \log\left(k^2 - \overline{m}_h^2\right) + 2\ \log\left(k^2 - \overline{m}_{A+}^2\right)\\ &+ 2\ \log\left(k^2 - \overline{m}_{A-}^2\right)\\ &\left.+ \log\left(k^2 - \overline{m}_{B+}^2\right) + \log\left(k^2 - \overline{m}_{B-}^2\right)\right],\end{aligned} \tag{4.68}$$

where we have introduced a mass parameter μ in order to keep the correct dimension of the scalar potential.

The loop integral is easily evaluated after Wick rotation, yielding (see for instance reference [26])

$$-\frac{i}{2}\mu^{2\epsilon}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\ \log(-k^2 + m^2) = \frac{1}{4}\frac{m^4}{(4\pi)^2}\left(\log\,\frac{m^2}{\mu^2} - \frac{3}{2} - \Delta_\epsilon\right), \tag{4.69}$$

where we have introduced the modified minimal subtraction ($\overline{\mathrm{MS}}$) term [27]

$$\Delta_\epsilon = \frac{1}{\epsilon} - \gamma_{\mathrm{E}} + \log\,4\pi. \tag{4.70}$$

[8] Subtleties related to the proper definition of γ_5 in d dimensions only arise at higher orders in perturbation theory, while a different choice for the trace, e.g. $\mathrm{Tr}\,\mathbf{1}_{\mathrm{Dirac}} = 2^{d/2}$, would just lead to a different renormalization scheme [25].

After the ϵ-expansion the one-loop contribution to the effective potential is given by

$$\begin{aligned}V_1(\phi,\epsilon) = V_1^{\text{pole}}(\phi,\epsilon) + \frac{1}{4(4\pi)^2}\Bigg[&-12\overline{m}_t^4\left(\log\frac{\overline{m}_t^2}{\mu^2}-\frac{3}{2}\right)\\ &+6\overline{m}_W^4\left(\log\frac{\overline{m}_W^2}{\mu^2}-\frac{5}{6}\right)+3\overline{m}_Z^4\left(\log\frac{\overline{m}_Z^2}{\mu^2}-\frac{5}{6}\right)\\ &+\overline{m}_h^4\left(\log\frac{\overline{m}_h^2}{\mu^2}-\frac{3}{2}\right)+2\overline{m}_{A+}^4\left(\log\frac{\overline{m}_{A+}^2}{\mu^2}-\frac{3}{2}\right)\\ &+2\overline{m}_{A-}^4\left(\log\frac{\overline{m}_{A-}^2}{\mu^2}-\frac{3}{2}\right)+\overline{m}_{B+}^4\left(\log\frac{\overline{m}_{B+}^2}{\mu^2}-\frac{3}{2}\right)\\ &+\overline{m}_{B-}^4\left(\log\frac{\overline{m}_{B-}^2}{\mu^2}-\frac{3}{2}\right)\Bigg],\end{aligned} \tag{4.71}$$

where all divergent terms as $\epsilon\to 0$ are collected in $V_1^{\text{pole}}(\phi,\epsilon)$, which can be expressed as a function of the standard model couplings as

$$\begin{aligned}V_1^{\text{pole}}(\phi,\epsilon) = \frac{\Delta_\epsilon}{(4\pi)^2}\Bigg[&-m^4-\left(3\lambda-\frac{1}{8}\xi_B g'^2-\frac{3}{8}\xi_W g^2\right)m^2\phi^2\\ &+\left(-\frac{3}{64}g'^4-\frac{3}{32}g'^2g^2-\frac{9}{64}g^4+\frac{3}{4}y_t^4-3\lambda^2\right.\\ &\left.+\frac{1}{8}\xi_B g'^2\lambda+\frac{3}{8}\xi_W g^2\lambda\right)\phi^4\Bigg].\end{aligned} \tag{4.72}$$

While the m^4-dependent pole in equation (4.72) can be always subtracted by a constant shift in the effective potential, the remaining divergences are cancelled by the multiplicative renormalization of the bare field and couplings appearing in $V_{\text{eff}}^{(0)}$ (cf equation (4.38)):

$$\phi_0 = Z_\phi^{1/2}\phi,\qquad m_0^2 = Z_{m^2}m^2,\qquad \lambda_0 = Z_\lambda\lambda, \tag{4.73}$$

where the renormalization constants can be conveniently computed in the unbroken phase of the standard model. Their expressions at one loop in the $\overline{\text{MS}}$ scheme read (see e.g. [28, 29]):

$$Z_\phi^{1/2} = 1+\frac{\Delta_\epsilon}{(4\pi)^2}\left(\frac{3}{8}g'^2+\frac{9}{8}g^2-\frac{3}{2}y_t^2-\frac{1}{8}\xi_B g'^2-\frac{3}{8}\xi_W g^2\right) \tag{4.74}$$

$$Z_{m^2} = 1+\frac{\Delta_\epsilon}{(4\pi)^2}\left(-\frac{3}{4}g'^2-\frac{9}{4}g^2+3y_t^2+6\lambda\right) \tag{4.75}$$

$$\begin{aligned}Z_\lambda = 1+\frac{\Delta_\epsilon}{(4\pi)^2}\Bigg(&-\frac{3}{2}g'^2-\frac{9}{2}g^2+6y_t^2+12\lambda+\frac{3}{16}\frac{g'^4}{\lambda}\\ &+\frac{3}{8}\frac{g'^2g^2}{\lambda}+\frac{9}{16}\frac{g^4}{\lambda}-3\frac{y_t^4}{\lambda}\Bigg).\end{aligned} \tag{4.76}$$

It is a simple exercise to check that the renormalization of the tree-level potential, via the renormalization constants in equations (4.74) and (4.76), cancels the ϕ-dependent poles in equation (4.72). We point out that in the Fermi gauge the field ϕ gets only multiplicatively renormalized by the wavefunction of the Higgs field, but this feature does not necessarily hold in other gauges.

After renormalization, the one-loop contribution to the effective potential in the $\overline{\mathrm{MS}}$ scheme reads

$$V_1(\phi) = \frac{1}{4(4\pi)^2}\left[-12\overline{m}_t^4\left(\log\frac{\overline{m}_t^2}{\mu^2}-\frac{3}{2}\right)+6\overline{m}_W^4\left(\log\frac{\overline{m}_W^2}{\mu^2}-\frac{5}{6}\right)\right. \tag{4.77}$$

$$+3\overline{m}_Z^4\left(\log\frac{\overline{m}_Z^2}{\mu^2}-\frac{5}{6}\right)+\overline{m}_h^4\left(\log\frac{\overline{m}_h^2}{\mu^2}-\frac{3}{2}\right)$$

$$+2\overline{m}_{A+}^4\left(\log\frac{\overline{m}_{A+}^2}{\mu^2}-\frac{3}{2}\right)+2\overline{m}_{A-}^4\left(\log\frac{\overline{m}_{A-}^2}{\mu^2}-\frac{3}{2}\right)$$

$$\left.+\overline{m}_{B+}^4\left(\log\frac{\overline{m}_{B+}^2}{\mu^2}-\frac{3}{2}\right)+\overline{m}_{B-}^4\left(\log\frac{\overline{m}_{B-}^2}{\mu^2}-\frac{3}{2}\right)\right],$$

where the ϕ-dependent mass terms are defined in equations (4.43)–(4.47) and equations (4.64)–(4.66). Note that at the tree-level minimum, $m^2 = -\lambda\phi^2$, we have $\overline{m}_\chi = 0$ and $\overline{m}_{A^\pm} = \overline{m}_{B^\pm} = 0$, and the gauge dependence drops from $V_1(\phi)$. The meaning of the gauge dependence of the standard model effective potential will be discussed in section 4.5. In particular, for $\xi_W = \xi_B = 0$ one has $\overline{m}_{A+} = \overline{m}_{B+} = \overline{m}_\chi$ and $\overline{m}_{A-} = \overline{m}_{B-} = 0$, so that equation (4.77) reproduces the well-known one-loop result in the Landau gauge [22].

4.4. Large-field behaviour of the standard model effective potential

We shall see that the analysis of the stability of the electroweak ground state requires the knowledge of the effective potential $V_{\text{eff}}(\phi)$ at large values of the classical field ϕ. In this limit, the terms proportional to $\log\frac{\phi}{\mu}$ appearing in the one-loop correction to the effective potential grow large, and eventually spoil the reliability of perturbation theory. Such large logarithms need resummation in this case. The standard technique to resum such logarithms is the renormalization group, which we review here.

For fixed values of the bare parameters, the effective potential $V_{\text{eff}}(\phi)$, does not depend on the renormalization scale μ (see equation (4.14)). This is expressed by the renormalization group equation (RGE)

$$\left(\mu\frac{\partial}{\partial\mu}+\beta_i\frac{\partial}{\partial\lambda_i}-\gamma\phi\frac{\partial}{\partial\phi}\right)V_{\text{eff}}(\mu,\phi)=0, \tag{4.78}$$

where we have made explicit the dependence of V_{eff} on the scale μ and on the parameters of the theory (couplings and masses), collectively denoted by λ_i.

The formal solution of the RGE in equation (4.78) can be obtained by the method of the characteristics [30]: the effective potential, as a function of μ, is constant along the characteristic curves $\lambda_i = \lambda_i(\mu), \phi = \phi_i(\mu)$, solutions of the differential equations

$$\mu\frac{\mathrm{d}\lambda_i(\mu)}{\mathrm{d}\mu} = \beta_i; \quad \lambda_i(\mu_0) = \lambda_i \tag{4.79}$$

$$\mu\frac{\mathrm{d}\phi(\mu)}{\mathrm{d}\mu} = -\gamma\phi(\mu); \quad \phi(\mu_0) = \phi. \tag{4.80}$$

Table 1. The p-coefficients entering the expression of $\lambda_{\rm eff}$ in equation (4.83) for the Fermi gauge.

p	t	W	Z	h	$A^{\pm}$	$B^{\pm}$
N_p	-12	6	3	1	2	1
C_p	$3/2$	$5/6$	$5/6$	$3/2$	$3/2$	$3/2$
κ_p	$\frac{y_t^2}{2}$	$\frac{g^2}{4}$	$\frac{g^2+g'^2}{4}$	3λ	$\frac{1}{2}\left(\lambda \pm \sqrt{\lambda^2 - \lambda\xi_W g^2}\right)$	$\frac{1}{2}\left(\lambda \pm \sqrt{\lambda^2 - \lambda(\xi_W g^2 + \xi_B g'^2)}\right)$

The usefulness of the renormalization group is that μ can be chosen in such a way that the convergence of perturbation theory is improved. For instance, a standard choice in vacuum stability analyses is $\mu = \phi$. Without sticking, for the time being, to any specific choice of scale, the RG improved effective potential can be formally written as

$$V_{\rm eff}(\mu, \phi) = \Omega_{\rm eff}(\mu, \phi) + \frac{m^2_{\rm eff}(\mu, \phi)}{2}\phi^2 + \frac{\lambda_{\rm eff}(\mu, \phi)}{4}\phi^4. \tag{4.81}$$

In the case of the standard model the full expressions of $\Omega_{\rm eff}$, $m^2_{\rm eff}$ and $\lambda_{\rm eff}$ in the Fermi gauge can be read off equation (4.77). We do not report those expressions explicitly here; we just observe that $\Omega_{\rm eff}$ and $m^2_{\rm eff}$ are proportional to $m^4(\mu)$ and $m^2(\mu)$, respectively, where $m(\mu)$ is the $\overline{\rm MS}$ renormalized parameter of the Higgs potential, essentially an electroweak mass scale up to logarithmic running. Hence, in the $\phi \gg m$ limit equation (4.81) takes the simplified form

$$V_{\rm eff}(\phi) \simeq \frac{1}{4}\lambda_{\rm eff}(\mu, \phi)\phi^4, \tag{4.82}$$

where

$$\lambda_{\rm eff}(\mu, \phi) = e^{4\Gamma(\mu)}\left[\lambda(\mu) + \frac{1}{(4\pi)^2}\sum_p N_p \kappa_p^2(\mu)\left(\log\,\frac{\kappa_p(\mu)e^{2\Gamma(\mu)}\phi^2}{\mu^2} - C_p\right)\right] \tag{4.83}$$

and

$$\Gamma(\mu) = -\int_{\mu_0}^{\mu}\gamma\frac{{\rm d}\mu'}{\mu'}. \tag{4.84}$$

The coefficients N_p, C_p and $\kappa_p(\mu)$ appearing in equation (4.83) are listed in table 1 in the Fermi gauge (the corresponding Landau–gauge values are obtained by taking $\xi_W = \xi_B = 0$). Note that the gauge dependence of the RG improved effective potential is twofold: the gauge fixing parameters appear both in the couplings κ_p (cf table 1), and in the anomalous dimension of ϕ (cf equation (4.74)).

Equation (4.82) can be further simplified by choosing $\mu \sim \phi$, so that the logarithms in the last term of equation (4.83) are small, and neglecting the RGE evolution of the field ϕ, which amounts to taking $\Gamma(\mu) \sim 0$ (this is formally correct only at the LO). We hence obtain

$$V_{\rm eff}(\phi) \simeq \frac{1}{4}\lambda(\phi)\phi^4. \tag{4.85}$$

The RG evolution of λ in the $\overline{\rm MS}$ scheme is displayed in the left panel of figure 5 for central values of the standard model parameters, employing one-, two- and three-loop beta functions and two-loop matching at the M_t scale [31]. Similarly, in the right panel of figure 5 we show the extrapolation of the SM effective potential at high energies in the approximation of equation (4.85).

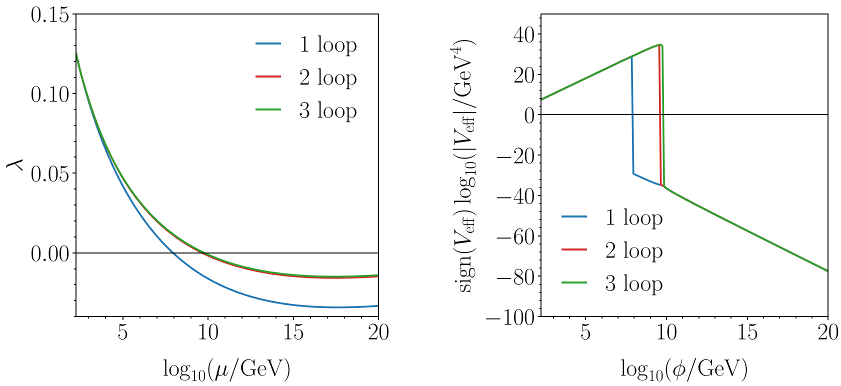

Figure 5. Left panel: RG evolution of λ in the standard model at one, two and three loops (respectively in blue, red and green). Right panel: standard model effective potential in the approximation $V_{\mathrm{eff}} \simeq \lambda(\phi)\phi^4/4$.

Note that within such an approximation the gauge dependence drops out, even though it is not perturbatively consistent to neglect the anomalous dimension and the fixed-order contribution of the effective potential in equation (4.83) when the running of λ is considered beyond one loop. This notwithstanding, already at the one-loop level the above approximation captures the most important physical feature: around $\mu = 10^8$ GeV the coupling λ becomes negative, thus triggering the instability of the Higgs potential. The electroweak minimum (whose fine structure is not visible in figure 5) is not anymore the absolute minimum of the Higgs potential after considering radiative corrections. Since λ keeps running negative at large field values, the effective potential must be ultraviolet completed in such a way that it is bounded from below and a second, deeper minimum arises. The latter might be due to Planck-scale physics or simply to the radiative self-completion of the SM, given the fact that λ crosses again the zero at trans-Planckian energies and grows positive due to the Landau pole of the hypercharge.

4.5. Gauge dependence

Although the effective potential turns out to be a very useful tool for the study of spontaneous symmetry breaking, there are some subtleties related to its physical interpretation which are worth to be discussed. In particular, we are going to review here the role of the gauge dependence of the effective potential in the vacuum stability analysis and, in the next section, the so-called convexity and reality issues.

The gauge dependence of the effective potential raises the question of which are the physical observables entering the vacuum stability analysis. To fix the ideas, let us assume that all the parameters of the standard model are exactly determined, but the Higgs boson mass. After choosing the renormalization scale μ, the RG improved effective potential, $V_{\mathrm{eff}}(\phi, M_h;\ \xi)$, is a function of ϕ, the Higgs pole mass M_h, and the gauge fixing parameters, collectively denoted by ξ. One can think of M_h as an order parameter, whose variation modifies the shape of the effective potential, as sketched in figure 6.

The absolute stability bound on the Higgs boson mass is obtained by defining a 'critical' mass, M_h^c, for which the value of the effective potential at the electroweak minimum, ϕ_1, and at

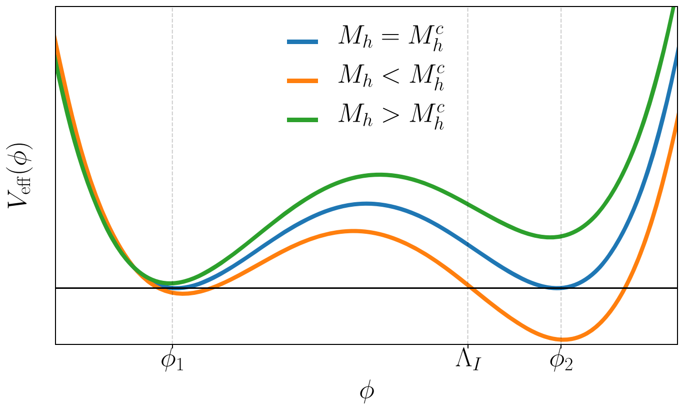

Figure 6. Schematic representation of the standard model effective potential for different values of the Higgs boson mass. For $M_h < M_h^c$, the electroweak vacuum is unstable.

a second minimum, $\phi_2 > \phi_1$, are the same. Analytically, this translates into the three conditions

$$V_{\rm eff}(\phi_1, M_h^c; \xi) - V_{\rm eff}(\phi_2, M_h^c; \xi) = 0, \tag{4.86}$$

$$\left.\frac{\partial V_{\rm eff}}{\partial \phi}\right|_{\phi_1, M_h^c} = \left.\frac{\partial V_{\rm eff}}{\partial \phi}\right|_{\phi_2, M_h^c} = 0. \tag{4.87}$$

We have seen in the previous section that, in the $\phi \gg \phi_1$ limit, the standard model effective potential is well approximated by equation (4.85),

$$V_{\rm eff}(\phi) \simeq \frac{1}{4}\lambda_{\rm eff}(\phi)\phi^4. \tag{4.88}$$

Hence, the absolute stability conditions in equations (4.86) and (4.87) can be equivalently rewritten in the following way:

$$\lambda_{\rm eff}(\phi_2, M_h^c; \xi) = 0, \tag{4.89}$$

$$\left.\frac{\partial \lambda_{\rm eff}}{\partial \phi}\right|_{\phi_2, M_h^c} = 0, \tag{4.90}$$

up to $(\phi_1/\phi_2)^2 \ll 1$ corrections.

On the other hand, due to the explicit presence of ξ in the vacuum stability condition, it is not obvious *a priori* which are the physical (gauge-independent) observables entering the vacuum stability analysis. The basic tool, in order to capture the gauge-invariant content of the effective potential is given by the Nielsen identity [32]

$$\left(\frac{\partial}{\partial \xi} + C(\phi, \xi)\frac{\partial}{\partial \phi}\right) V_{\rm eff}(\phi, \xi) = 0, \tag{4.91}$$

where $C(\phi, \xi)$ is a correlator involving the ghost fields and the gauge-fixing functional, whose explicit expression will not be needed for our argument. Equation (4.91) is valid for the class of linear gauges and can be derived from the BRST non-invariance of a composite operator involving the ghost field and the gauge fixing functional (see e.g. [33]).

The identity in equation (4.91) has the following physical interpretation: the effective potential is gauge independent where it is stationary and hence spontaneous symmetry breaking is a gauge-invariant statement. We can actually use the Nielsen identity, in combination with the vacuum stability condition in equations (4.86) and (4.87), to formally prove that the critical Higgs boson mass, M_h^c, is a gauge-independent quantity [34]. To this purpose, let us assume that simultaneously inverting equations (4.86) and (4.87) would yield gauge dependent field values and critical Higgs boson mass: $\phi_1 = \phi_1(\xi)$, $\phi_2 = \phi_2(\xi)$ and $M_h^c = M_h^c(\xi)$. The total differential of equation (4.86) with respect to ξ then reads

$$\left.\frac{\partial V_{\text{eff}}}{\partial \phi}\right|_{\phi_1, M_h^c} \frac{\partial \phi_1}{\partial \xi} + \left.\frac{\partial V_{\text{eff}}}{\partial M_h}\right|_{\phi_1, M_h^c} \frac{\partial M_h^c}{\partial \xi} + \left.\frac{\partial V_{\text{eff}}}{\partial \xi}\right|_{\phi_1, M_h^c}$$
$$= \left.\frac{\partial V_{\text{eff}}}{\partial \phi}\right|_{\phi_2, M_h^c} \frac{\partial \phi_2}{\partial \xi} + \left.\frac{\partial V_{\text{eff}}}{\partial M_h}\right|_{\phi_2, M_h^c} \frac{\partial M_h^c}{\partial \xi} + \left.\frac{\partial V_{\text{eff}}}{\partial \xi}\right|_{\phi_2, M_h^c}. \tag{4.92}$$

The first terms in both the lhs and the rhs of equation (4.92) vanish because of the stationary conditions in equation (4.87). The third terms in both the lhs and the rhs of equation (4.92) vanish for the same reason, after using the Nielsen identity. Hence, we are left with

$$\left(\left.\frac{\partial V_{\text{eff}}}{\partial M_h}\right|_{\phi_1, M_h^c} - \left.\frac{\partial V_{\text{eff}}}{\partial M_h}\right|_{\phi_2, M_h^c} \right) \frac{\partial M_h^c}{\partial \xi} = 0. \tag{4.93}$$

Since the expression in the bracket of equation (4.93) is in general different from zero, one concludes that

$$\frac{\partial M_h^c}{\partial \xi} = 0, \tag{4.94}$$

namely, the critical Higgs boson mass does not depend on the gauge-fixing parameters, and hence telling whether the electroweak minimum is absolutely stable or not is a physical statement. Note, however, that a non-consistent use of perturbation theory might still be responsible for a residual (spurious) gauge dependence in the determination of M_h^c [35, 36].

On the other hand, field values (as for instance the instability scale) are essentially gauge-dependent quantities. The standard model vacuum instability scale is operatively defined as the field value $\phi = \Lambda_I$, for which the effective potential has the same depth as the electroweak minimum (see figure 6). This is analytically expressed by

$$V_{\text{eff}}(\Lambda_I; \xi) = V_{\text{eff}}(\phi_1; \xi). \tag{4.95}$$

The rhs of equation (4.95) is a gauge-independent quantity, since ϕ_1 is by definition a minimum and we can apply the Nielsen identity. Hence, by solving equation (4.95), one has in general $\Lambda_I = \Lambda_I(\xi)$. In particular, by taking the total differential of equation (4.95) with respect to ξ, we get

$$\left.\frac{\partial V_{\text{eff}}}{\partial \phi}\right|_{\Lambda_I} \frac{\partial \Lambda_I}{\partial \xi} + \left.\frac{\partial V_{\text{eff}}}{\partial \xi}\right|_{\Lambda_I} = 0. \tag{4.96}$$

By using the Nielsen identity, we can substitute back the second term in equation (4.96), thus obtaining

$$\left(\frac{\partial \Lambda_I}{\partial \xi} - C(\Lambda_I, \xi) \right) \left.\frac{\partial V_{\text{eff}}}{\partial \phi}\right|_{\Lambda_I} = 0. \tag{4.97}$$

Since, in general, Λ_I is not an extremum of the effective potential, equation (4.97) yields

$$\frac{\partial \Lambda_I}{\partial \xi} = C(\Lambda_I, \xi). \tag{4.98}$$

It turns out that by varying ξ within its perturbative domain, the scale Λ_I in the standard model suffers from a gauge-fixing uncertainty of up to one order of magnitude in the Fermi gauge [34]. Although Λ_I is clearly unphysical, one can still identify some gauge-invariant scales related with the Higgs potential instability (see e.g. reference [37]).

4.6. Convexity and reality

The effective potential $V_{\text{eff}}(\phi_c)$, being defined as the Legendre transform of the generating functional of the connected Green's functions, must be a convex function of its argument[9]. Moreover, as we are going to show, it carries the following physical interpretation: it is the expectation value of the energy density $\langle \Psi | \mathcal{H} | \Psi \rangle$, for the lowest-energy state $|\Psi\rangle$, subject to the constraints $\langle \Psi | \Psi \rangle = 1$ and $\langle \Psi | \phi(x) | \Psi \rangle = \phi_c$. This also implies that V_{eff} must be real.

These two properties of the effective potential are in apparent contrast with what done so far, since any scalar potential relevant for spontaneous symmetry breaking has non-convex regions. Moreover, the convexity and reality problems are closely related. In fact, the one-loop correction to V_{eff} is proportional to $[V''(\phi_c)]^2 \, \log \, V''(\phi_c)$, so that non-convex regions ($V''(\phi_c) < 0$) lead to an imaginary part.

It is commonly understood that this paradox has nothing to do with the breakdown of perturbation theory, but it is rather an issue of terminology (see e.g. [38–40]). In the derivation of V_{eff} in section 4.1 we have implicitly assumed that $\phi_c[J]$ (defined via the relation $\phi_c = \delta W/\delta J$) is a single-valued function. It can be shown that whenever the tree-level potential has non-convex regions, the relation between ϕ_c and J becomes multi-valued and the standard one-loop derivation of what we improperly called the effective potential gets modified [38]. Specifically, the saddle point approximation employed in the background field method of section 4.2 features multiple stationary points with respect to which the effective action should be expanded, and the true effective potential is obtained by summing over all of them.

It is therefore more appropriate to distinguish between two objects: V_{eff}, defined in terms of the Legendre transform of the generating functional of connected Green's functions, and $V_{1\text{PI}}$, defined as the negative of the spatially independent part of the generating functional of the 1PI Green's functions (cf equation (4.14)). The two generating functionals are related by (see e.g. [41])

$$\exp\left[\frac{i}{\hbar}(W[J] + \mathcal{O}(\hbar))\right] = \mathcal{N} \int \mathcal{D}\phi_c \, \exp\left[\frac{i}{\hbar}\left(\Gamma_{1\text{PI}}[\phi_c] + \int \mathrm{d}^4 x \, J(x)\phi_c(x)\right)\right], \tag{4.99}$$

for some irrelevant normalization constant $\mathcal{N}$. In the semi-classical limit, $\hbar \to 0$, we can use the saddle point approximation to obtain

$$W[J] = \left[\Gamma_{1\text{PI}}[\phi_c] + \int \mathrm{d}^4 x \, J(x)\phi_c(x)\right]\Bigg|_{\frac{\delta\Gamma_{1\text{PI}}}{\delta\phi_c} = -J}. \tag{4.100}$$

[9] Some mathematical properties of Legendre transforms are reviewed in appendix A.4.

Hence, $W[J]$ is the Legendre transform of $\Gamma_{1\mathrm{PI}}[\phi_c]$. On the other hand, the effective action was defined as the Legendre transform of $W[J]$

$$\Gamma_{\mathrm{eff}}[\phi_c] = \left[W[J] - \int \mathrm{d}^4x\, J(x)\phi_c(x)\right]\Bigg|_{\frac{\delta W}{\delta J}=\phi_c}. \tag{4.101}$$

Therefore $\Gamma_{\mathrm{eff}}[\phi_c]$ is the double Legendre transform of $\Gamma_{1\mathrm{PI}}[\phi_c]$. From the properties of Legendre transforms it follows that $\Gamma_{\mathrm{eff}}[\phi_c]$ is the convex envelope of $\Gamma_{1\mathrm{PI}}[\phi_c]$. In turn, this implies that V_{eff} is the convex envelope of $V_{1\mathrm{PI}}$.

Note that the formalism behind V_{eff} is defined in the thermodynamical limit, where the system has relaxed on the ground state after an infinite amount of time. On the other hand, in most applications, we are rather interested in metastable regimes described by field configurations localized around a local minimum. The goal of this section is to show that $V_{1\mathrm{PI}}$ is actually the appropriate object in order to describe this physical situation.

Following Coleman [10], we start by considering the quantum mechanical problem of constructing a state $|a\rangle$ that is a stationary state of the quadratic form $\langle a|H|a\rangle$, subject to the constraints $\langle a|a\rangle = 1$ and $\langle a|A|a\rangle = A_c$, for some Hermitian operator A. This can be cast in the variational problem $\delta\langle a|H - E - JA|a\rangle = 0$, where E and J are Lagrangian multipliers. The variational equation leads to

$$[H - E - JA]|a\rangle = 0, \tag{4.102}$$

from which we see that $|a\rangle$ is an eigenstate of the perturbed Hamiltonian $H - JA$, with energy E. Equation (4.102) allows us to express $|a\rangle = |a(E, J)\rangle$ and, further using the normalization condition, we can express E in terms of J. Since we are eventually interested in the dependence on A_c, the relation between J and A_c can be obtained via standard perturbation theory[10]

$$A_c(J) = \langle a(J)|A|a(J)\rangle = -\frac{\mathrm{d}E}{\mathrm{d}J}. \tag{4.103}$$

Hence, we conclude that $|a\rangle$ is the stationary state of the quantity

$$\langle a(J)|H|a(J)\rangle = E(J) + JA_c = E(J) - J\,\frac{\mathrm{d}E}{\mathrm{d}J}. \tag{4.104}$$

The generalization to quantum field theory is obtained by replacing: $|a\rangle \to |\Psi\rangle$, $H \to \mathcal{H}$, $A \to \phi$, $A_c \to \phi_c$, $E \to W$, which corresponds to the definition of the effective action in equation (4.101). Hence, we conclude that

$$V_{\mathrm{eff}}(\phi_c) = \langle\Psi|\mathcal{H}|\Psi\rangle, \tag{4.105}$$

for a state $|\Psi\rangle$ such that $\delta\langle\Psi|\mathcal{H}|\Psi\rangle = 0$, under the constraints $\langle\Psi|\Psi\rangle = 1$ and $\langle\Psi|\phi(x)|\Psi\rangle = \phi_c$.

To investigate this further, let us consider for definiteness the case of the double-well potential $V(\phi) = -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4$ with minima in $\phi = \pm\sigma \equiv \sqrt{\mu^2/\lambda}$ (cf figure 7). The classical potential is clearly non-convex in the region $\phi < |\sigma|$. Following Weinberg and Wu [42], the main features of the effective potential can be understood via a classical argument. Let us consider the classical analog of $\langle\Psi|\phi(x)|\Psi\rangle$ to be the spatial average of ϕ, denoted by $\overline{\phi}$. Using the physical interpretation of the effective potential, the classical analog of the effective potential,

[10] This relation holds true in general (see e.g. [40]). However, it is sufficient for the sake of the argument to treat J as a perturbation.

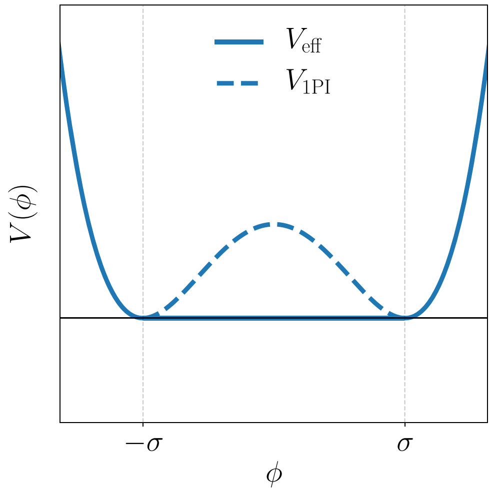

Figure 7. Example of a non-convex tree-level potential. The true effective potential ($V_{\rm eff}$) is the convex envelope of $V_{\rm 1PI}$.

$V_{\rm c}$, is the minimum value of the energy density among all the states with a given value of $\overline{\phi}$. For $\phi > |\sigma|$ the energy density

$$\mathcal{H} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi), \tag{4.106}$$

is minimized by a static homogeneous configuration with $\phi(x) = \overline{\phi}$ everywhere, and $V_{\rm c}$ is equal to $V(\phi)$. If $\phi < |\sigma|$, there is a non-trivial way in which the energy density can be minimized, namely via an *inhomogeneous* state

$$\phi(x) = \begin{cases} +\sigma & \text{for a fraction } f \text{ of space} \\ -\sigma & \text{otherwise,} \end{cases} \tag{4.107}$$

where $0 \leqslant f \leqslant 1$ is fixed by requiring that the spatial average of $\phi(x)$, $\sigma f - \sigma(1-f)$, is equal to $\overline{\phi}$, which yields $f = (\overline{\phi} + \sigma)/(2\sigma)$. Hence, in the $\phi < |\sigma|$ region the energy density is $V(\pm\sigma) < V(\phi)$. Note that the gradient term relevant at the boundaries of the two domains can be neglected in the infinite volume limit. $V_{\rm c}$ is thus flat in the region between the two minima, and corresponds to the convex envelop of the $V(\phi)$. In the quantum theory, the main difference is that the state which minimizes the energy density is not a spatially inhomogeneous mixture of two phases but rather a quantum superposition of two vacuum states.

If one is interested only in homogeneous states, i.e. states in which the field is concentrated about a specific value[11], then the effective potential is not the appropriate quantity to describe the region $\phi < |\sigma|$. In that case one would need a modified effective potential which might be defined as the expectation value of the energy density in a state $|\Psi\rangle$ which minimizes $\langle\Psi|\mathcal{H}|\Psi\rangle$ subject to the condition $\langle\Psi|\phi(x)|\Psi\rangle = \phi_c$ and to the further restriction that the wave functional for $|\Psi\rangle$ be concentrated on configurations with $\phi(x) \approx \phi_c$. Hence, to leading order

[11] This is clearly the case in the physical problem considered in the present review: we always start from a homogeneous field configuration localized around the electroweak minimum $\phi(x) \approx v$ and ask about its stability.

the minimum energy density is $V(\phi_c)$ which does not need to be convex. Zero point fluctuations are further taken into account by $V_{1\mathrm{PI}}(\phi_c)$, whose real part provides the relevant potential to address localized field configurations prior to a phase transition. Note that the restriction to localized configurations does not commute with the Hamiltonian (e.g. in the case if the double-well potential the energy eigenstates are even/odd linear combinations of the states localized in $\pm\sigma$), thus these configurations are unstable and can be interpreted as imaginary contributions to the energy [42]. This instability, related to the imaginary part of the perturbatively computed $V_{1\mathrm{PI}}$, should not to be confused with the non-perturbative decay of the false vacuum itself (for instance, one could consider a non-convex $V_{1\mathrm{PI}}$ with field configurations localized around the true vacuum).

Following a standard terminology, we will keep to refer to V_{eff} as the (localized) effective potential.

5. Decay of the standard model vacuum

The stability of the electroweak ground state is a time-honoured subject [40, 43–52]. Prior to the Higgs discovery in 2012 [53, 54] it served as a theoretical tool to set a lower bound on the Higgs boson mass [3, 55–60]. Recently, the analysis of the standard model vacuum stability has reached a higher level of complexity (see e.g. [31, 61–67]). In this section we describe the basic aspects of the calculation of the standard model vacuum decay rate and we introduce the so-called standard model phase diagram, namely the parameter space region in the standard model for which the electroweak vacuum is absolutely stable, metastable (i.e. with a sufficiently long lifetime compared to the age of the Universe) or worryingly unstable.

5.1. The origin of the standard model instability

The case of the standard model of electroweak interactions is different from the simple examples described in section 3. Indeed, the tree-level scalar potential of the standard model

$$V_0(\phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{5.1}$$

with $m^2 < 0$ and $\lambda > 0$ has a local maximum at $\phi = 0$, an absolute minimum at $\phi = \sqrt{-\frac{m^2}{\lambda}} = v$ (the electroweak minimum), with $v \sim 246$ GeV, and no other stationary points. Hence, at the classical level, there is no instability of the ground state.

The situation changes when loop corrections are included. As we have seen in the previous section, the effective potential behaves as $\lambda(\phi)\phi^4$ for large values of the classical field ϕ, and the running coupling $\lambda(\mu)$ becomes negative at $\mu = \Lambda_I \sim 10^{10}$ GeV if the top quark is sufficiently heavy. If this is the case, the electroweak minimum is not the absolute minimum of the effective potential: a second, deeper minimum, located somewhere between 10^{10} GeV and the position of the Landau singularity, must therefore exist, and quantum fluctuations may induce tunnelling from the electroweak minimum to the absolute minimum.

The study of the instability of the electroweak ground state must therefore be carried on consistently at the level of loop corrections.

Since such large energy scales and field values are involved, the calculation of the tunnelling probability is in general influenced by the ultraviolet completion of the standard model (if any), and by gravitational effects; it is however reasonable to take a conservative point of view, and perform a calculation of the false vacuum lifetime under the assumption that the standard model is a correct description of our Universe up to scales of the order of the Planck mass.

5.2. The standard model bounce

We have seen in section 3 that the calculation of the decay rate of an unstable ground state in quantum field theory is based on an $O(4)$-invariant bounce, $\phi(x) = \phi_b(r)$, solution of the field equation

$$\phi_b''(r) + \frac{3}{r}\phi_b'(r) = V'(\phi_b), \tag{5.2}$$

where $V(\phi)$ is the scalar potential expressed in terms of renormalized parameters, with boundary conditions

$$\lim_{r\to\infty} \phi_b(r) = v, \tag{5.3}$$

$$\phi_b'(0) = 0. \tag{5.4}$$

In the case of the standard model the value of ϕ (which denotes the Higgs background field, see equations (4.36) and (4.37)) at the true minimum of the scalar potential is typically many orders of magnitudes larger that the electroweak order parameter, $v \sim 246$ GeV. It seems therefore appropriate to simply neglect the finite value of v with respect to the typical energy scales involved in the problem of the quantum stability of the electroweak ground state. We will therefore adopt the boundary condition

$$\lim_{r\to\infty} \phi_b(r) = 0, \tag{5.5}$$

and we will approximate the scalar potential in the unstable region by

$$V(\phi) = \frac{1}{4}\lambda(\phi)\phi^4, \tag{5.6}$$

with the renormalized coupling $\lambda(\mu)$ computed at a scale $\mu \sim \phi$ to reduce the impact of large logarithms. Equation (5.2) takes the form

$$\phi_b'' + \frac{3}{r}\phi_b' = \phi_b^3\left[\lambda(\phi_b) + \frac{1}{4}\beta(\lambda(\phi_b))\right], \tag{5.7}$$

where

$$\beta(\lambda(\mu)) = \mu\,\frac{\mathrm{d}\lambda(\mu)}{\mathrm{d}\mu}. \tag{5.8}$$

(Although not explicitly indicated, the function $\beta(\lambda)$ actually depends on all coupling constants in the theory, and in particular on the top Yukawa coupling, which drives λ to negative values, as we have seen in section 4.) A bounce solution of equation (5.7), that is, a solution with $\phi_b'(0) = \phi_b(\infty) = 0$, certainly exists, because the scalar potential equation (5.6) has a local minimum in $\phi = 0$ and an absolute minimum at some very large value of ϕ, and the undershoot–overshoot argument of section 3.2 can be applied. We recall that, by the same argument, the value of the bounce at $r = 0$ is larger than the position of the first zero of $\lambda(\mu)$, $\mu \sim \Lambda_I = 10^{10}$ GeV. At such large values of the classical field, the use of the approximation (5.6) is largely justified; the mass term would induce a correction

$$\frac{\Delta\phi(0)}{\phi(0)} < \frac{v}{\Lambda_I} \sim 10^{-8}. \tag{5.9}$$

An analytic solution of equation (5.7) cannot be found. However, it can be checked that the dependence of the combination

$$\lambda(\phi) + \frac{1}{4}\beta(\lambda(\phi)) \tag{5.10}$$

on ϕ in the region where $\lambda(\phi) < 0$ is very weak: it is only logarithmic, and it is suppressed by powers of the coupling constants. An approximate solution can therefore be found by replacing the rhs of equation (5.7) by a negative constant λ. As we are going to show below, a good approximation is obtained by fixing $\lambda = \lambda(\overline{\mu})$ where $\overline{\mu}$ is such that $\beta(\lambda(\overline{\mu})) = 0$. Equation (5.7) then becomes

$$\phi_b''(r) + \frac{3}{r}\phi_b'(r) = \lambda\phi_b^3(r). \tag{5.11}$$

(Incidentally, we note that we are now precisely in the situation described in section 3.4, namely a scalar potential with no barrier.) As anticipated in section 3.5.1 and shown explicitly in appendix A.5, equation (5.11) with the boundary conditions (5.4) and (5.5) can be solved analytically. There is in fact an infinity of solutions:

$$\phi_b(r) = \frac{\phi_b(0)}{1 + \frac{|\lambda|}{8}\phi_b^2(0)r^2}, \tag{5.12}$$

parametrized by $\phi_b(0)$, which can take any positive value. Equation (5.12) is often rewritten in terms of a length scale R, rather than by its value at $r = 0$:

$$\phi_b(r) = \sqrt{\frac{8}{|\lambda|}}\frac{R}{r^2 + R^2}. \tag{5.13}$$

The parameter R can be physically interpreted as the approximate radius of the bubble, in three dimensional space, where the system has jumped from the false-vacuum configuration $\phi_b = 0$ to a non-zero value beyond the barrier. Indeed, $\phi_b(r)$ has its maximum in $r = 0$:

$$\phi_b(0) = \sqrt{\frac{8}{|\lambda|}}\frac{1}{R}, \tag{5.14}$$

and drops to zero for $r \gg R$. Note that $\phi_b(R) = \frac{\phi_b(0)}{2}$, which justifies our definition of R as the radius of the bubble.

As already noted in section 3.5.1, this degeneracy is a consequence of scale invariance. Indeed, the underlying action functional

$$S[\phi] = \int \mathrm{d}^4x \left[\frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{4}\lambda\phi^4\right] \tag{5.15}$$

is invariant upon scale transformations:

$$\phi(r) \to a\phi(ar), \tag{5.16}$$

since it does not contain any dimensionful parameter. The proof is straightforward:

$$S[a\phi(ar)] = \int \mathrm{d}^4x \left[\frac{a^2}{2}\left(\frac{\partial\phi(ar)}{\partial x_\mu}\right)^2 + \frac{a^4}{4}\lambda\phi^4(ar)\right] = S[\phi] \tag{5.17}$$

after the integration variable change $y = ax$. As a consequence, if $\phi(r)$ is a solution of equation (5.11), than $a\phi(ar)$ is also a solution, for any choice of the scale factor a:

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{3}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)[a\phi(ar)] = a^3\left[\phi''(ar) + \frac{3}{ar}\phi'(ar)\right] = \lambda[a\phi(ar)]^3. \tag{5.18}$$

The transformation $\phi(r) \to a\phi(ar)$ on equation (5.12) amounts to rescaling

$$\phi_b(0) \to a\phi_b(0). \tag{5.19}$$

However, the scale invariance of equation (5.11) is not a feature of our original problem, but rather a consequence of our approximations. The original field equation, equation (5.7), is *not* scale-invariant, even when mass terms are neglected, because of renormalization; therefore, its bounce solution has a definite value at $r = 0$. Unfortunately, the correct value of $\phi_b(0)$ cannot be determined analytically. A numerical study of the bounce solution to equation (5.7) was presented for instance in reference [20] based on the shooting method reviewed in appendix A.6. The results show that the commonly adopted approximation in equation (5.11) is rather accurate. However, its application in the calculation of the first quantum corrections requires some care, as discussed below in section 5.4.

We conclude this section by computing the action of the Fubini–Lipatov bounces equation (5.12), which enters the calculation of the decay rate of the metastable ground state. The action is easily computed using equation (3.12):

$$S[\phi_b] = -\int \mathrm{d}^4x\, V(\phi_b(x)) = -2\pi^2\int_0^{+\infty} \mathrm{d}r\, r^3 V(\phi_b(r)), \tag{5.20}$$

where in the last step we have used the $O(4)$ invariance of the bounce solution. We find

$$S[\phi_b] = \frac{8\pi^2}{3|\lambda|}, \tag{5.21}$$

independently of the value of $\phi_b(0)$, as expected.

5.3. A digression on scale transformations

In this section we collect some considerations about the role of scale invariance (or non-invariance) in the search for bounce solutions of the field equations. Let us consider a generic scalar field theory, with scalar potential $V(\phi)$, and let us assume that the action $S[\phi]$ is stationary around $\phi = \phi_b$. Then, in particular, it will be stationary upon infinitesimal scale transformations, that is we expect

$$\Delta B = \Delta(S[\phi_b] - S[\phi_{\mathrm{FV}}]) = 0 \tag{5.22}$$

upon a variation

$$\Delta\phi = (1 + x_\rho\partial_\rho)\phi_b \tag{5.23}$$

around a solution of the field equation. A straightforward calculation gives

$$\Delta B = \int \mathrm{d}^4x\left[\phi V'(\phi_b) - 4V(\phi_b) + 4V(\phi_{\mathrm{FV}})\right]. \tag{5.24}$$

The rhs of equation (5.24) is obviously zero for

$$V(\phi) = \frac{1}{4}\lambda\phi^4 + V(\phi_{\mathrm{FV}}), \tag{5.25}$$

because the integrand vanishes identically in this case. We already know that in this case bounce solutions of the field equation exist if $\lambda < 0$.

It is easy to find examples of scalar potentials that do not admit any bounce solution of the field equation. Consider for instance

$$V(\phi) = \frac{1}{4}\lambda\phi^4 + \frac{1}{2}m^2\phi^2 \tag{5.26}$$

with $\lambda < 0$ and constant m^2, either positive or negative. In this case,

$$\Delta B = \Delta S[\phi] = -m^2 \int \mathrm{d}^4x\, \phi^2(x), \tag{5.27}$$

which vanishes only for $\phi(x) = 0$.

There are however intermediate cases, that is, non-scale-invariant scalar potentials that admit a bounce solution of the field equation. For example

$$V(\phi) = g\left[\frac{1}{4}\phi^4 - \frac{1}{3}\phi^3(M + \phi_{\mathrm{TV}}) + \frac{1}{2}\phi^2 M\phi_{\mathrm{TV}}\right] \tag{5.28}$$

with $g > 0$ and

$$0 < M < \frac{\phi_{\mathrm{TV}}}{2} \tag{5.29}$$

has a local minimum for $\phi = 0$ and an absolute minimum for $\phi = \phi_{\mathrm{TV}}$, separated by a barrier with its maximum for $\phi = M$. A plot of $\frac{V(\phi)}{g\phi_{\mathrm{TV}}^4}$ as a function of $h = \frac{\phi}{\phi_{\mathrm{TV}}}$ for different values of $m = \frac{M}{\phi_{\mathrm{TV}}}$ is shown in figure 8. In this case the field equation

$$\phi'' + \frac{3}{r}\phi' = g\phi(\phi - M)(\phi - \phi_{\mathrm{TV}}) \tag{5.30}$$

cannot be solved analytically. We know, however, that it must have a bounce solution, as a consequence of the overshoot–undershoot argument illustrated in section 3.2. It is possible to solve the field equation numerically; the bounce $h(x)$ as a function of $x = \sqrt{g}\phi_{\mathrm{TV}}r$ is plotted in figure 9 for different values of m. Note that $h(0)$ gets close to 1, i.e. $\phi(0)$ close to ϕ_{TV}, as $m \to \frac{1}{2}$, where the two minima are degenerate. Also note that the solution is unique, as a consequence of the explicit breaking of scale invariance, contrary to what happens for the scale-invariant potential equation (5.25): the condition

$$\Delta B = g\int \mathrm{d}^4x \left[\frac{1}{3}\phi_b^3(M + \phi_{\mathrm{TV}}) - \phi_b^2 M\phi_{\mathrm{TV}}\right] = 0 \tag{5.31}$$

fixes the value of the bounce at $r = 0$. It can be shown that in this case the bounce is well approximated by a Gaussian

$$\phi_b(r) = \phi_0 e^{-\frac{r^2}{R^2}}, \tag{5.32}$$

in the sense that values of the parameters ϕ_0 and R can be found, such that the true bounce action is accurately reproduced.

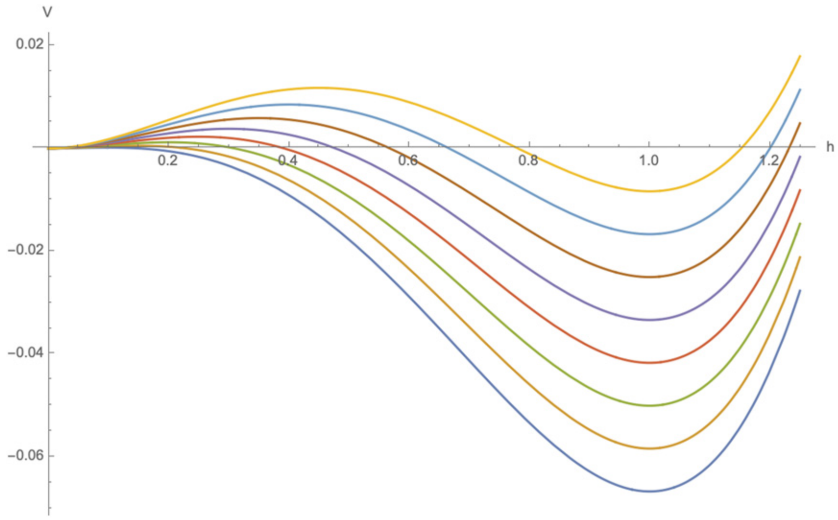

Figure 8. The scalar potential equation (5.28), normalized to $g\phi_{\mathrm{TV}}^4$, for different values of m. From top to bottom, $m = 0.45, 0.4, 0.35, 0.3, 0.25, 0.2, 0.15, 0.1$.

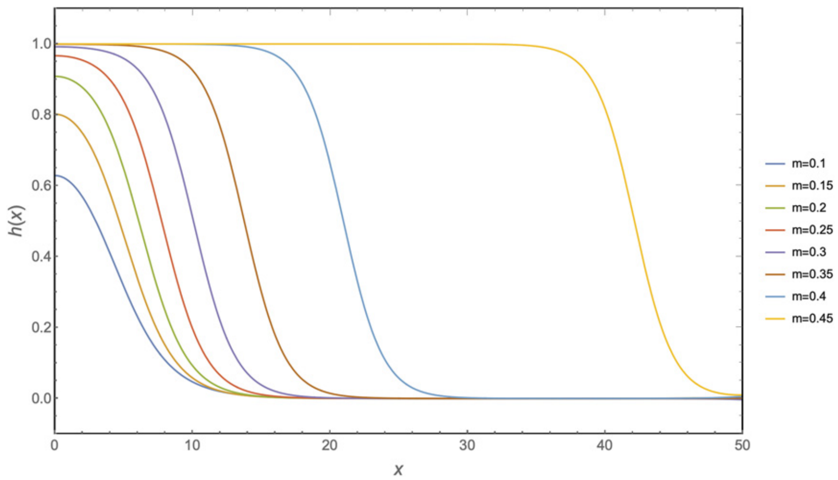

Figure 9. Bounce solutions relative to the scalar potential in equation (5.28) for different values of m.

The standard model scalar potential with the mass term neglected and the running coupling computed at $\mu \sim \phi$,

$$V(\phi) = \frac{1}{4}\lambda(\phi)\phi^4; \qquad \mu\frac{\mathrm{d}\lambda}{\mathrm{d}\mu} = \beta(\lambda) \tag{5.33}$$

also belongs to the same class: it has a local minimum at $\phi = 0$, increases with ϕ, then starts decreasing when $\lambda(\phi)$ becomes negative. We have argued in the previous section that in this case the bounce is well approximated by the Fubini–Lipatov bounce; the value of $\phi(0)$ is determined by the condition of stationarity of the action upon scale transformations,

$$\Delta B = \frac{1}{4}\int \mathrm{d}^4 x\, \beta(\lambda(\phi_b))\phi_b^4 = 0. \tag{5.34}$$

An analytical estimate of $\phi_b(0)$ in this case is presented in reference [68]: for values of ϕ not too far from the value of the scale $\overline{\mu}$ where $\lambda(\overline{\mu})$ has a minimum we have

$$\lambda(\phi) = \lambda(\overline{\mu}) + \frac{1}{2}\frac{\mathrm{d}\beta(\lambda(\phi))}{\mathrm{d}\,\log\,\phi}\bigg|_{\phi=\overline{\mu}} \log^2\frac{\phi}{\overline{\mu}} + O\left(\log^3\frac{\phi}{\overline{\mu}}\right) \tag{5.35}$$

and therefore

$$\beta(\lambda(\phi)) = \frac{\mathrm{d}\beta(\lambda(\phi))}{\mathrm{d}\,\log\,\phi}\bigg|_{\phi=\overline{\mu}} \log\frac{\phi}{\overline{\mu}} + O\left(\log^2\frac{\phi}{\overline{\mu}}\right). \tag{5.36}$$

Neglecting terms of order $\log^2\frac{\phi}{\overline{\mu}}$, the stationarity condition in equation (5.34) yields

$$\int \mathrm{d}^4\, x\, \phi_b^4(r)\log\,\phi_b(r) = \log\,\overline{\mu}\int \mathrm{d}^4 x\, \phi_b^4(r). \tag{5.37}$$

The two integrals are easily computed with $\phi(r)$ replaced by the Fubini–Lipatov bounce equation (5.12). We get

$$\begin{aligned}\log\,\frac{\phi_b(0)}{\overline{\mu}}&\int\frac{r^3\,\mathrm{d}r}{\left(1+\frac{|\lambda|}{8}\phi_b^2(0)r^2\right)^4}\\ &=\int\frac{r^3\,\mathrm{d}r}{\left(1+\frac{|\lambda|}{8}\phi_b^2(0)r^2\right)^4}\,\log\left(1+\frac{|\lambda|}{8}\phi_b^2(0)r^2\right)\end{aligned} \tag{5.38}$$

or

$$\phi_b(0) = \overline{\mu}e^{\frac{5}{6}}. \tag{5.39}$$

The result in equation (5.39) can be turned into an estimate of the bounce size R, defined by equation (5.14). We obtain

$$R = \sqrt{\frac{8}{|\lambda(\overline{\mu})|}}\frac{1}{\phi_b(0)} \sim \frac{10}{\overline{\mu}} \tag{5.40}$$

for $|\lambda(\overline{\mu})| \sim 0.013$. Because $\overline{\mu} \sim 10^{17}$ GeV, equation (5.39) is an *a posteriori* confirmation that the mass term in the potential can be safely neglected.

For earlier work discussing the breaking of scale invariance due to the running of λ in the context of the standard model vacuum tunnelling calculation see also references [51, 52].

5.4. First quantum corrections

The calculation of the first quantum corrections to the decay rate of the standard model vacuum is presented in reference [3]. The calculation is performed by the procedure illustrated in section 3.5, which produces the result

$$\frac{\gamma}{\mathcal{V}_3} = \frac{(S[\phi_b]-S[\phi_{\mathrm{FV}}])^2}{(2\pi)^2}\sqrt{\frac{\mathrm{Det}\; S''[\phi_{\mathrm{FV}}]}{|\mathrm{Det}'\, S''[\phi_b]|}}e^{-(S[\phi_b]-S[\phi_{\mathrm{FV}}])}e^{-(S_{\mathrm{ct}}[\phi_b]-S_{\mathrm{ct}}[\phi_{\mathrm{FV}}])}, \tag{5.41}$$

for the decay rate per unit volume of the metastable ground state.

In the case of the standard model, the calculation of the pre-exponential factor is complicated by the presence of fermion and vector degrees of freedom in the theory, which contribute fluctuations around the bounce configuration. We will not reproduce here the full calculation; we limit ourselves to pointing out a few important features.

The first point is the choice of a suitable basis for the space of eigenfunctions of $S''[\phi]$ for the evaluation of ratios of functional determinants

$$\frac{\text{Det } S''[\phi_b]}{\text{Det } S''[\phi_{\text{FV}}]}. \tag{5.42}$$

This requires solving eigenvalue equations of the type

$$S''[\phi]\, u = [-\partial^2 + W(\phi)]u = \lambda u, \tag{5.43}$$

where u generically denotes scalar, fermion or gauge fields, and the function $W(\phi)$ takes different forms for each contribution.

Because $W(\phi)$ in (5.43) depends only on the radial coordinate r, it is convenient to decompose the various fields in eigenstates of the four-dimensional angular momentum operator $L^2 = L_{\mu\nu}L_{\mu\nu}$, where

$$L_{\mu\nu} = \frac{i}{\sqrt{2}}\left(x_\mu\partial_\nu - x_\nu\partial_\mu\right). \tag{5.44}$$

The Laplace operator takes the form

$$\partial^2 = \partial_\mu\partial_\mu = \frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{3}{r}\frac{\mathrm{d}}{\mathrm{d}r} - \frac{L^2}{r^2}. \tag{5.45}$$

In the case of scalar fields, the eigenfunctions of L^2 are the four-dimensional spherical harmonics $Y_j(\theta)$ (where θ collectively denotes the three polar angles) with eigenvalues $4j(j+1)$, and degeneracy $(2j+1)^2$, where j takes integer and semi-integer values. In this basis, $S''[\phi_b]$ is block-diagonal, with blocks $\left(S''[\phi_b]\right)_j$ labelled by j; therefore

$$\log\ \text{Det } S''[\phi_b] = \sum_j \log\ \text{Det}\,(S''[\phi_b])_j. \tag{5.46}$$

The situation is slightly more complicated for fermion and vector fields; details can be found in reference [3].

As illustrated in section 2.2.2, in order to compute the ratio in equation (5.42) it is only necessary to solve equation (5.43) for $\lambda = 0$:

$$\rho_j \equiv \frac{\text{Det}\left[-\partial^2 + W(\phi_b)\right]}{\text{Det}\left[-\partial^2 + W(\phi_{\text{FV}})\right]} = \lim_{r\to\infty} \frac{\det u^j_{\phi_b}(r)}{\det u^j_{\phi_{\text{FV}}}(r)}, \tag{5.47}$$

where $u^j_\phi(r)$ are solutions of

$$[-\partial^2 + W(\phi)]\, u^j_\phi(r) = 0. \tag{5.48}$$

The symbol Det in equation (5.47) stands for the ordinary determinant over residual (spinorial, gauge group, etc) indices of these solutions.

The second important point is the subtraction of ultraviolet divergences. The ultraviolet behaviour of the functional determinant is encoded in the behaviour of the determinants for $j \to \infty$; as a consequence, the sum over j in equation (5.46) is not convergent, and the usual renormalization procedure is needed. The counterterm action $S_{\rm ct}$ removes the ultraviolet divergences that arise in the calculation of the functional determinants. Indeed,

$$\sqrt{\mathrm{Det}\ S''[\phi]}e^{S[\phi]+S_{\rm ct}[\phi]} = e^{S[\phi]+S_{\rm ct}[\phi]+\frac{1}{2}\log\mathrm{Det}\,S''[\phi]} = e^{\Gamma[\phi]}, \tag{5.49}$$

where $\Gamma[\phi]$ is the one-loop effective action, defined in section 4. For example, in the $\overline{\mathrm{MS}}$ scheme (dimensional regularization with $\overline{\mathrm{MS}}$ subtraction) the renormalized one-loop action can be written as

$$\Gamma[\phi] = S^{\overline{\mathrm{MS}}}[\phi] + \left[\Delta S[\phi] - (\Delta S[\phi])_{\rm pole}\right], \tag{5.50}$$

where $S^{\overline{\mathrm{MS}}}[\phi]$ is the lowest-order action expressed in terms of the renormalized $\overline{\mathrm{MS}}$ couplings, and $(\Delta S)_{\rm pole}$ is the divergent part of ΔS defined according to the $\overline{\mathrm{MS}}$ renormalization prescription. The calculation of the full determinant in $d = 4 - 2\epsilon$ dimensions can be avoided by observing that the ultraviolet divergences are all contained in $\Delta S^{[2]}$, defined as the expansion of ΔS up to second order in W:

$$\Delta S^{[2]}[\phi] = \frac{1}{2}\left[\log\ \mathrm{S\,Det}\left(-\partial^2 + W(\phi)\right)\right]_{O(W^2)}, \tag{5.51}$$

where $S\,\mathrm{Det}\,O = \mathrm{Det}\,O$ or $S\,\mathrm{Det}\,O = (\mathrm{Det}\,O)^{-2}$ depending on whether the operator O acts on boson or fermion fields. This is easily seen in the case of scalar degrees of freedom, when

$$W_{\rm scalar}(\phi) = V''(\phi). \tag{5.52}$$

In renormalizable theories, the second derivative of the scalar potential is a polynomial of degree 2 in the field, while counterterms are a polynomial of degree 4. Hence the difference $\Delta S - \Delta S^{[2]}$ is ultraviolet-finite. Equation (5.50) can be rewritten as

$$\Gamma[\phi] = S^{\overline{\mathrm{MS}}}[\phi] + \left[\Delta S[\phi] - \Delta S^{[2]}[\phi]\right] + \left[\Delta S^{[2]}[\phi] - (\Delta S[\phi])_{\rm pole}\right], \tag{5.53}$$

where the two terms in square brackets are separately finite. The advantage of this last expression is that $\Delta S^{[2]}[\phi]$ can be computed either as a (divergent) sum of terms corresponding to different values of the angular momentum, which gives a finite result when subtracted from $\Delta S[\phi]$, or by standard diagrammatic techniques in $4 - 2\epsilon$ dimensions.

The details of the calculation of one-loop contributions to $\Gamma[\phi_b]$ from field fluctuations in the different sectors of the theory described in reference [3] is performed in terms of the Fubini–Lipatov bounce, equation (5.12) or equation (5.13), with $\phi_b(0)$, or equivalently R, fixed as described at the end of section 5.3. In this approximation, the zero mode of scale invariance in the scalar sector is not normalizable, as shown section 3.5.1. This is however an artefact of our approximations: the bounce is a solution of equation (5.11), rather than equation (5.7). This difficulty is overcome by removing the zero eigenvalue of dilatations from the functional determinant, and replacing it by a slightly negative eigenvalue, induced by radiative corrections.

Additional zero modes of $S''[\phi_b]$ arise because of the gauge symmetry; the corresponding contributions to the path integral must be removed and replaced by an integration over suitable collective coordinates. This procedure is described in reference [17], and involves a jacobian factor proportional to

$$\int \mathrm{d}^4x\, \phi_b^2(r) \tag{5.54}$$

(see equation (3.62)), which is infrared divergent in the case of the Fubini–Lipatov bounce. Again, this is a consequence of our approximation of neglecting mass terms, and hence the nonvanishing vacuum expectation value of the Higgs field, in the scalar potential; in the real world, the Higgs mass acts as a regulator of the infrared divergence of equation (5.54). To see how this works we consider the field equation

$$\phi_b''(r) + \frac{3}{r}\phi_b'(r) = V'(\phi_b) \tag{5.55}$$

in the large-r limit. In this region, $\phi_b(r) \to \phi_{\mathrm{FV}}$, which is different from zero if the mass term is not neglected. In this limit, the field equation can be linearized by expanding the scalar potential in powers of $\phi_b - \phi_{\mathrm{FV}}$:

$$V(\phi_b) = V(\phi_{\mathrm{FV}}) + \frac{1}{2}V''(\phi_{\mathrm{FV}})(\phi_b - \phi_{\mathrm{FV}})^2 + \mathrm{O}\big((\phi_b - \phi_{\mathrm{FV}})^3\big), \tag{5.56}$$

where $V''(\phi_{\mathrm{FV}}) = m^2$ is a positive constant, related to the physical Higgs squared mass. Neglecting terms of order $(\phi_b - \phi_{\mathrm{FV}})^3$, equation (5.55) becomes

$$(\phi_b(r) - \phi_{\mathrm{FV}})'' + \frac{3}{r}(\phi_b(r) - \phi_{\mathrm{FV}})' = m^2(\phi_b(r) - \phi_{\mathrm{FV}}), \tag{5.57}$$

and can be solved analytically. In terms of the dimensionless quantities

$$x = rm; \qquad f(x) = r(\phi_b(r) - \phi_{\mathrm{FV}}). \tag{5.58}$$

Equation (5.57) reads

$$x^2 f''(x) + x f'(x) - (1 + x^2) f(x) = 0, \tag{5.59}$$

which is solved by a linear combination of modified Bessel functions

$$f(x) = \alpha K_1(x) + \beta I_1(x), \tag{5.60}$$

where

$$I_1(x) = \frac{x}{2}\sum_{n=0}^{\infty} \frac{1}{n!(n+1)!}\left(\frac{x^2}{4}\right)^n \tag{5.61}$$

$$K_1(x) = \left(\log\frac{x}{2} + \gamma\right) I_1(x) + \frac{1}{x} + \frac{x}{4}\sum_{n=0}^{\infty}\frac{1}{n!(n+1)!}\left(\sum_{k=1}^{n+1}\frac{1}{k} + \sum_{k=1}^{n}\frac{1}{k}\right)\left(\frac{x^2}{4}\right)^n. \tag{5.62}$$

The large-x behaviours of $I_1(x), K_1(x)$ are as follows:

$$I_1(x) \sim \frac{1}{\sqrt{2\pi x}} e^x \tag{5.63}$$

$$K_1(x) \sim \sqrt{\frac{\pi}{x}} e^{-x}. \tag{5.64}$$

The condition $\phi_b(r) - \phi_{\mathrm{FV}} \to 0$ for $r \to \infty$ gives $\beta = 0$, and therefore in this limit

$$\phi_b(r) - \phi_{\mathrm{FV}} \simeq \alpha m \frac{K_1(rm)}{rm} \sim \frac{1}{(rm)^{3/2}} e^{-rm}. \tag{5.65}$$

This shows that $\phi_b^2(r)$ is integrable, if the finite value of the Higgs mass is taken into account.

The final result for the tunnelling rate of the metastable ground state can now be written as

$$\frac{\gamma}{\mathcal{V}_3} = \frac{1}{\mathcal{V}_3 T} \frac{e^{-\Gamma[\phi_b]}}{e^{-\Gamma[\phi_{\mathrm{FV}}]}}; \tag{5.66}$$

the space–time volume factor $\mathcal{V}_3 T$ is cancelled by the exclusion of zero eigenvalues, as we have seen.

The final result is expressed in terms of the renormalized parameters λ, y_t and g_i, whose definitions depend on the renormalization scheme. Of course, the scheme dependence disappears once these couplings are re-expressed in terms of physical observables, such as Higgs and top pole masses. In practice, however, in the case of the gauge couplings it turns out to be more convenient to directly use the $\overline{\mathrm{MS}}$ definitions, since these parameters are accurately determined by fitting multiple observables.

An updated next-to-leading order analysis of the tunnelling rate of the standard model vacuum yields [66]

$$\tau_{\mathrm{SM}} = \left(\frac{\gamma_{\mathrm{SM}}}{\mathcal{V}_3}\right)^{-1/4} = 10^{161^{+160}_{-59}} \text{ year} \tag{5.67}$$

at 68% confidence level, where the uncertainty is roughly due in equal parts to the uncertainty on the value of the top quark mass and the strong-coupling constant and the theory uncertainty from threshold corrections at the top mass scale. Even taking into account these uncertainties, the lifetime of the standard model vacuum easily complies with the age of the Universe $\tau_{\mathrm{U}} \approx 13.8 \times 10^9$ year.

5.5. The standard model phase diagram

We now come to a discussion of the stability of the standard model ground state in the light of our present knowledge of the relevant parameters. The most relevant parameters in the calculation of the decay rate of the standard model vacuum are the value of the top quark mass M_t, which is responsible for the negative value of the beta function for $\lambda(\mu)$, the Higgs boson mass M_h, which determines the value of $\lambda(\mu)$ at the electroweak scale, and the strong coupling $\alpha_s(M_Z)$, which affects indirectly the calculation of the beta function. Traditionally, the result of the calculation if the vacuum decay rate are displayed in the (M_h, M_t) plane (cf Figure 10), often referred to as the standard model phase diagram[12].

Let us describe how the standard model phase diagram (which by definition assumes the standard model in isolation) is obtained and comment on possible interpretations of the fact that the experimental values of the standard model parameters seem to lie at the boundary between the stability and metastability region.

[12] Similar analyses have been performed also in references [36, 61–67].

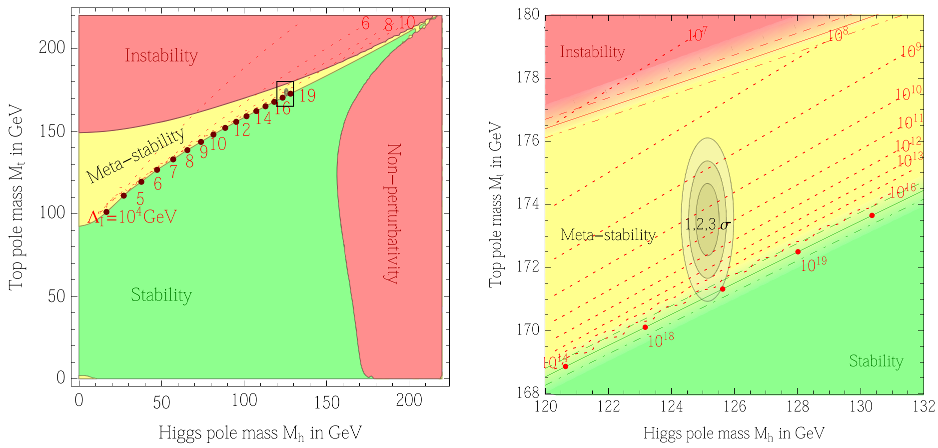

Figure 10. The standard model phase diagram in the (M_h, M_t) plane. Reproduced from [31]. CC BY 2.0.

5.5.1. Instability/metastability boundary. The probability of nucleating a bubble of true vacuum during the history of the Universe can be estimated as

$$p = \frac{\gamma}{\mathcal{V}_3}(\mathcal{V}_3 T)_{\text{light-cone}}, \tag{5.68}$$

with the space–time volume of our past light-cone given by [31]

$$(\mathcal{V}_3 T)_{\text{light-cone}} \approx \frac{0.15}{H_0^4}, \tag{5.69}$$

and $H_0 = 1.44 \times 10^{-42}$ GeV being the present Hubble constant. The boundary region between instability (red) and metastability (yellow) in figure 10 is obtained by setting $p = 1$ as a function of the top and Higgs pole masses.

5.5.2. Metastability/stability boundary. The absolute stability condition was already discussed in section 4.5 (cf equations (4.86), (4.87), (4.88) and (4.89)) and reads

$$\lambda_{\text{eff}}(\mu_c) = \beta_{\lambda_{\text{eff}}}(\mu_c) = 0, \tag{5.70}$$

where $\beta_{\lambda_{\text{eff}}} = \mu \, \frac{\mathrm{d}\lambda_{\text{eff}}}{\mathrm{d}\mu}$ and we have neglected small corrections of order $(v/\mu_c)^2 \approx 10^{-30}$ which corresponds to approximating $V_{\text{eff}}(\phi) = \frac{1}{4}\lambda_{\text{eff}}(\phi)\phi^4$. A more accurate stability criterium, which takes into account the consistent use of the $\hbar$ expansion of the standard model effective potential, was formulated in reference [36].

The critical line between metastability (yellow) and stability (green) in figure 10 is obtained by selecting the region in the (M_h, M_t) plane where equation (5.70) holds. The technical tools entering the high-energy extrapolation of the standard model effective potential have been developed dramatically after the Higgs discovery and involve the calculation of beta functions [28, 29, 69, 70], threshold corrections at the M_t mass scale [31, 61, 62, 71] and the fixed-order standard model effective potential [72–74]. With these ingredients at hand, one arrives to the remarkable conclusion that the electroweak vacuum is near-critical. While the central values of the standard model parameters point to the metastability region, with a lifetime way larger than the age of the Universe, absolute stability is excluded only between 2 and 3σ confidence level (cf right panel in figure 10). At the moment the leading uncertainty is of experimental nature and is due to the uncertainty on M_t and $\alpha_s(M_Z)$.

The main message of the standard model phase diagram is that the standard model can be consistently extrapolated up to the Planck scale. Pushing its interpretation even further, the apparent near-criticality of the standard model vacuum has triggered many speculations about the possible dynamics behind that (see e.g. [31, 75]).

5.6. Ultraviolet sensitivity and new physics

Physical degrees of freedom beyond those present in the standard model are expected, both on the basis of some experimental observations (massive neutrinos, dark matter, etc) and of theoretical considerations (e.g. the naturalness problem). The presence of extra degrees of freedom can in general affect the standard model effective potential, possibly modifying, or even invalidating, the conclusions of the previous sections about the stability of the electroweak ground state. In such cases, the study of the electroweak vacuum lifetime is still a useful tool in order to constrain extensions of the standard model: non-standard physics should not make the electroweak vacuum too unstable.

The lifetime of the electroweak vacuum is obviously sensitive to new physics if its mass scale is $M < \overline{\mu} \approx 10^{17}$ GeV, since the standard model bounce gets directly affected. On the other hand, also new physics with $M \gg \overline{\mu}$ can enhance the tunnelling rate, e.g. by opening new decay channels for the electroweak vacuum [4, 20, 76–79].[13]

Such non-decoupling behaviour of ultraviolet physics might appear counter-intuitive. In quantum mechanics the shape of the potential beyond the barrier does not affect the tunnelling rate through the barrier itself. The case of quantum field theory is different. As shown in section 3.4, the generalization of the potential barrier in quantum field theory involves a space–time integral of the scalar potential evaluated at the bounce configuration $\phi_b(r)$; as a consequence, all energy scales between 0 and $\phi_b(0)$ are relevant.

The impact of non-standard physics on the ground state lifetime can be assessed parametrizing new physics effects on the scalar potential by non-renormalizable operators. For example, in reference [4] it was proposed to modify the standard model scalar potential as

$$V(\phi) = \frac{1}{4}\lambda\phi^4 + \frac{\lambda_6}{6M^2}\phi^6 + \frac{\lambda_8}{8M^4}\phi^8, \tag{5.71}$$

where λ is the standard Higgs self-coupling, λ_6, λ_8 are assumed to be of order one, and M is the energy scale at which new physics effects become relevant. The potential in equation (5.71) is not meant to be a realistic ultraviolet completion of the standard model. From an effective field theory perspective, for instance, a more realistic picture would be obtained by considering a full tower of higher dimensional operators, including derivative operators as well. In the present discussion, we do not address the question of how specific models can affect the standard model vacuum lifetime, but just use the potential in equation (5.71) as a toy model in order to parametrize new physics effects.

The inclusion of two non-renormalizable terms in equation (5.71) does not allow us anymore to find an analytical solution of the bounce equation and one has to proceed numerically[14]. The sensitivity of the tunnelling rate to new physics at very large energy scales can be assessed by

[13] For the moment, we are neglecting Planck-scale physics in order to highlight some general features of tunnelling in flat space–time. The inclusion of gravitational effects will be discussed in section 6.

[14] Numerical techniques based on the shooting algorithm for the determination of the bounce are reviewed in appendix A.6.

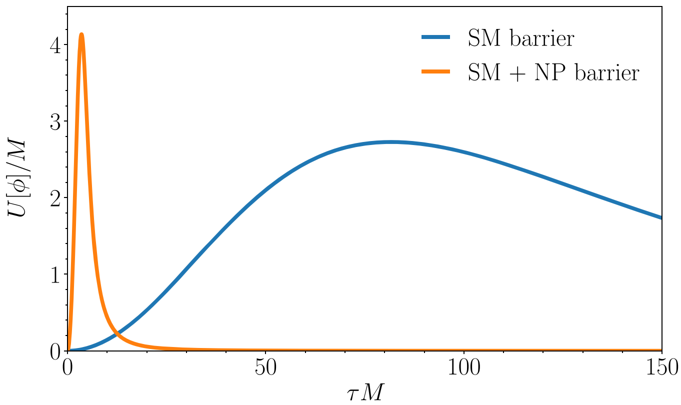

Figure 11. Potential barrier with and without the contribution of new physics. The values chosen for the new physics parameters are $\lambda_6 = -2$, $\lambda_8 = 2.1$.

comparing the potential energy barrier obtained with the potential in equation (5.71) with the one in the pure standard model. Following the definitions in equations (3.45)–(3.47), for the standard model case with $\lambda < 0$ one finds:

$$U_{\nabla}[\phi_b] = \frac{2\pi^2}{|\lambda|R}\left(\frac{1}{1+\frac{\tau^2}{R^2}}\right)^{\frac{3}{2}}, \qquad U_V[\phi_b] = -\frac{2\pi^2}{|\lambda|R}\left(\frac{1}{1+\frac{\tau^2}{R^2}}\right)^{\frac{5}{2}}, \tag{5.72}$$

which have opposite sign, while their sum

$$U[\phi_b] = U_{\nabla}[\phi_b] + U_V[\phi_b] = \frac{2\pi^2}{|\lambda|R}\left(\frac{1}{1+\frac{\tau^2}{R^2}}\right)^{\frac{3}{2}} \frac{\frac{\tau^2}{R^2}}{1+\frac{\tau^2}{R^2}}, \tag{5.73}$$

is positive, thus providing a true potential barrier.

The potential barrier in the presence of non-renormalizable terms in the scalar potential as in equation (5.71) can be computed by solving numerically the bounce equation. In figure 11 we compare the barrier $U[\phi_b]$ in the presence of new physics (orange) and in the pure standard model (blue). We have chosen $\lambda = \lambda(\overline{\mu}) \simeq -0.013\,45$ (corresponding to the minimum of $\lambda(\mu)$ in the standard model), $\lambda_6 = -2$ and $\lambda_8 = 2.1$, and we have set $M = M_{\rm Pl}$ for the scale of new physics. While the two barriers have comparable heights, the area under the curve which corresponds to the inclusion of new physics in this case is considerably smaller, which results in a much shorter lifetime of the metastable ground state for this particular choice of the parameters.

The above discussion can be generalized to the case of more non-renormalizable terms, including derivative operators, or even renormalizable ultraviolet completions of the standard model (see also [78, 80]). On the other hand, specific embeddings of the standard model into ultraviolet theories at the Planck scale can help in taming new instabilities of the Higgs potential [81].

Summarizing, while new physics at scales $M < \overline{\mu} \approx 10^{17}$ GeV can affect the decay rate of the standard model vacuum in any direction (possibly stabilizing the electroweak vacuum), modifications of the standard model potential at scales $M \gg \overline{\mu} \approx 10^{17}$ can only shorten the lifetime of the electroweak vacuum, whenever they lead to a new bounce solution which is related to a new, deeper minimum around the scale M.

6. Gravitational effects on vacuum decay

We have seen that very high energy scales are involved in the study of the metastability of the standard model ground state: the running coupling constant of the Higgs boson, $\lambda(\mu)$ becomes negative at $\mu \sim 10^{10}$ GeV, and reaches its maximum negative value (which determines the ground state lifetime) at $\mu = \overline{\mu} \approx 10^{17}$ GeV. It is therefore natural to ask whether gravitational effects become relevant in this regime. Indeed, the typical scale of gravitational interactions, the Planck mass $M_{\mathrm{Pl}} = 1.22 \times 10^{19}$ GeV, is only a couple of orders of magnitude larger than $\overline{\mu}$.

This issue was first addressed by Coleman and De Luccia in reference [5], where they considered the effects of gravity on the decay of a metastable vacuum in quantum field theory. In this section we will review their work, which allows us to introduce the relevant formalism. Next, following reference [6], we will apply the formalism of Coleman and De Luccia to the standard model. More recent developments on vacuum decay in the presence of gravity can be found e.g. in references [82, 83].

6.1. The tunnelling rate in the presence of gravity

The inclusion of gravity in the calculation of the decay rate of the standard model vacuum is in principle straightforward: one must find a configuration $\phi_b(x)$ of the Higgs field which is a stationary point of the Euclidean Einstein–Hilbert action

$$S_g = \int \mathrm{d}^4x \, \sqrt{g} \left[\frac{1}{2} \partial_\mu \phi_b \, \partial_\nu \phi_b \, g^{\mu\nu} + V(\phi_b) - \frac{\mathcal{R}}{2\kappa} \right], \tag{6.1}$$

where $\mathcal{R}$ is the Riemann curvature, and

$$\kappa = 8\pi G_{\mathrm{N}} = \frac{8\pi}{M_{\mathrm{Pl}}^2}, \tag{6.2}$$

where G_{N} is the Newton constant and $M_{\mathrm{Pl}} = 1.22 \times 10^{19}$ GeV is the Planck mass. We will limit ourselves to the case of minimal coupling to gravity, which formally corresponds to the leading order in the M_{Pl} expansion. The impact on the tunnelling rate of unknown non-minimal coupling terms in the Einstein–Hilbert–Higgs action was addressed e.g. in references [84, 85].

A difficulty immediately arises: contrary to the case of a flat spacetime, there is no guarantee that the bounce solution is $O(4)$ invariant. We will make this assumption, which simplifies considerably the problem, but it should be kept in mind that a bounce solution with a smaller value of the action and no $O(4)$ symmetry might in principle exist, although no example has been exhibited so far. We will therefore look for an $O(4)$ invariant metric

$$\mathrm{d}s^2 = \mathrm{d}r^2 + \rho^2(r)\mathrm{d}\Omega_3^2, \tag{6.3}$$

where $\mathrm{d}\Omega_3^2$ is the element of distance on a unit three-sphere,

$$\mathrm{d}\Omega_3^2 = \mathrm{d}\chi^2 + \sin^2\chi \, \mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2, \tag{6.4}$$

and an $O(4)$ invariant bounce $\phi_b(r)$, which obeys the field equations[15]

$$\phi_b'' + \frac{3\rho'}{\rho}\phi_b' = V'(\phi_b), \tag{6.5}$$

[15] Some details on the derivation of the field equations in the presence of gravity are provided in appendix A.7.

$$\rho'^2 = 1 + \frac{1}{3}\kappa\rho^2\left(\frac{1}{2}\phi_b'^2 - V(\phi_b)\right). \tag{6.6}$$

The boundary conditions for the field equations depend on the topology of the solution, which cannot be specified in advance. Since we are mainly interested in the problem of the stability of the standard model vacuum, we can neglect the small positive cosmological constant and assume a Minkowski-like background metric in the false vacuum. In such a case the bounce solution is non-compact, and $r \to \infty$. The boundary conditions leading to a finite bounce action are then

$$\phi_b'(0) = 0, \qquad \phi_b(\infty) = 0, \qquad \rho(0) = 0, \qquad \rho(\infty) = r, \tag{6.7}$$

where both the false vacuum and the value of the potential on the false vacuum are set to zero and the metric is asymptotically Minkowski-like. The condition $\rho(0) = 0$ is less obvious and comes from the fact that $\rho(r)$ must have at least one zero which can be set to $r = 0$ by translational invariance [7]. The Euclidean action equation (6.1) takes the form

$$S_g[\phi_b, g] = 2\pi^2\int_0^\infty \mathrm{d}r\left[\rho^3\left(\frac{\phi_b'^2}{2} + V(\phi_b)\right) + \frac{3}{\kappa}\left(\rho^2\rho'' + \rho\rho'^2 - \rho\right)\right]. \tag{6.8}$$

The bounce action can be rewritten in a simpler form by a generalization of the virial theorem, proven in section 3.1 in the flat-space case. Let us assume that the action is stationary upon variations of $\phi_b(x)$ and $g^{\mu\nu}(x)$; then in particular it should be stationary upon the transformations

$$\phi_b(x) \to \phi_b(bx), \tag{6.9}$$

$$g^{\mu\nu}(x) \to g^{\mu\nu}(bx), \tag{6.10}$$

which imply

$$\mathcal{R}(x) \to b^2\mathcal{R}(bx). \tag{6.11}$$

We find

$$S_g \to S_g(b), \tag{6.12}$$

where

$$\begin{aligned} S_g(b) &= \int \mathrm{d}^4\, x\, \sqrt{g(bx)}\left[\frac{1}{2}\frac{\partial\phi_b(bx)}{\partial x^\mu}\frac{\partial\phi_b(bx)}{\partial x^\nu}g^{\mu\nu}(bx) - \frac{b^2\mathcal{R}(bx)}{2\kappa} + V(\phi_b(bx))\right] \\ &= \int \mathrm{d}^4x\, \sqrt{g(x)}\left[\frac{1}{2b^2}\frac{\partial\phi_b(x)}{\partial x^\mu}\frac{\partial\phi_b(x)}{\partial x^\nu}g^{\mu\nu}(x) - \frac{\mathcal{R}(x)}{2b^2\kappa} + \frac{1}{b^4}V(\phi_b(x))\right]. \end{aligned} \tag{6.13}$$

The stationarity condition

$$\left.\frac{\mathrm{d}S_g(b)}{\mathrm{d}b}\right|_{b=1} = 0 \tag{6.14}$$

gives a relationship between the derivative terms and the potential term in the action:

$$\int \mathrm{d}^4x\,\sqrt{g(x)}\left[\frac{1}{2}\frac{\partial\phi_b(x)}{\partial x^\mu}\frac{\partial\phi_b(x)}{\partial x^\nu}g^{\mu\nu}(x)-\frac{\mathcal{R}(x)}{2\kappa}\right]$$
$$=-2\int \mathrm{d}^4x\,\sqrt{g(x)}V(\phi_b(x)). \tag{6.15}$$

Hence

$$S_g[\phi_b,g]=-\int \mathrm{d}^4x\,\sqrt{g(x)}V(\phi_b(x)). \tag{6.16}$$

In the flat-space limit this expression reproduces the standard virialized form of the action [2]. Assuming an $O(4)$ invariant metric, equation (6.16) becomes (see also [85])

$$S_g[\phi_b,g]=-2\pi^2\int_0^\infty \mathrm{d}r\,\rho^3(r)V(\phi_b(r)). \tag{6.17}$$

Equation (6.17) turns out to be much more useful than the equivalent form equation (6.8) in numerical estimates of the tunnelling rate, because the last two terms in equation (6.8) are both divergent for $r\to\infty$, while their difference is finite.

6.2. Perturbative solution of the field equations

The field equations (6.5) and (6.6) cannot be solved analytically. Since the effects of gravity are typically small, it is tempting to look for a solution as an expansion in powers of κ, as suggested in reference [6], and to neglect systematically all terms of order κ^2 and higher. To this purpose, one must expand ϕ_b to order κ and ρ to order κ^2, because of the κ^{-1} coefficient in the Einstein–Hilbert action:

$$\phi_b=\phi_{b0}+\kappa\phi_{b1}+O(\kappa^2), \tag{6.18}$$
$$\rho=\rho_0+\kappa\rho_1+\kappa^2\rho_2+O(\kappa^3), \tag{6.19}$$

where ϕ_{b0} is the bounce in the absence of gravity, and

$$\rho_0(r)=r, \tag{6.20}$$

as appropriate for a flat metric. The two equations for the order-κ corrections:

$$\phi_{b1}''+\frac{3}{r}\phi_{b1}'+\frac{3}{r}\left(\rho_1'-\frac{\rho_1}{r}\right)\phi_{b0}'=V''(\phi_{b0})\phi_{b1} \tag{6.21}$$
$$\rho_1'=\frac{1}{12}r^2\left({\phi_{b0}'}^2-2V(\phi_{b0})\right) \tag{6.22}$$

are no longer coupled to each other.

It is interesting to note that, to order κ, the correction terms ϕ_{b1} and ρ_2 do not affect the Euclidean action. Indeed, the first term in equation (6.8) becomes

$$\begin{aligned}\rho^3\left(\frac{{\phi_b'}^2}{2}+V(\phi_b)\right)\\ &= r^3\left(\frac{{\phi_{b0}'}^2}{2}+V(\phi_{b0})\right)\\ &\quad+\kappa\left[3r^2\rho_1\left(\frac{{\phi_{b0}'}^2}{2}+V(\phi_{b0})\right)+r^3\left(\phi_{b1}'\phi_{b0}'+\phi_{b1}V'(\phi_{b0})\right)\right]+O(\kappa^2).\end{aligned}\tag{6.23}$$

We now observe that

$$r^3\left(\phi_{b1}'\phi_{b0}'+\phi_{b1}V'(\phi_{b0})\right)=\frac{\mathrm{d}}{\mathrm{d}r}\left(r^3\phi_{b1}\phi_{b0}'\right),\tag{6.24}$$

where we have used the flat-space field equation for ϕ_{b0}

$$\phi_{b0}''+\frac{3}{r}\phi_{b0}'=V'(\phi_{b0})\tag{6.25}$$

to eliminate $V'(\phi_{b0})$. Hence, the last term in equation (6.23) gives no contribution to the action, provided

$$\lim_{r\to 0}\left(r^3\phi_{b1}\phi_{b0}'\right)=\lim_{r\to+\infty}\left(r^3\phi_{b1}\phi_{b0}'\right)=0.\tag{6.26}$$

Furthermore, expanding ρ to order κ^2 we find

$$\rho^2\rho''+\rho\rho'^2-\rho=\kappa\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\rho_1'\right)+\kappa^2\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\rho_2'\right)+\kappa^2\left(2\rho_1\rho_1'+2r\rho_1\rho_1''+r{\rho_1'}^2\right).\tag{6.27}$$

The only ρ_2-dependent term vanishes upon integration. Therefore, the bounce action depends only on $\phi_{b0}, \rho_0, \rho_1$, up to corrections of order κ^2:

$$\begin{aligned}S_g[\phi_b,g]=2\pi^2\int_0^\infty\mathrm{d}r\left[r^3\left(\frac{{\phi_{b0}'}^2}{2}+V(\phi_{b0})\right)+3\kappa r^2\rho_1\left(\frac{{\phi_{b0}'}^2}{2}+V(\phi_{b0})\right)\right.\\ \left.+3\kappa\left(2\rho_1\rho_1'+2r\rho_1\rho_1''+r{\rho_1'}^2\right)\right]+O(\kappa^2).\end{aligned}\tag{6.28}$$

6.3. Decay of the standard model vacuum in the presence of gravity

The formalism outlined in the previous sections can be applied to the standard model [6]. As in the case of a flat space–time, we approximate the standard model scalar potential by

$$V(\phi)=\frac{1}{4}\lambda\phi^4,\tag{6.29}$$

with $\lambda = \lambda(\overline{\mu}) < 0$. The linearized field equations, equations (6.21) and (6.22), take the form

$$\phi_{b1}'' + \frac{3}{r}\phi_{b1}' + \frac{3}{r}\left(\rho_1' - \frac{\rho_1}{r}\right)\phi_{b0}' = 3\lambda\phi_{b0}^2\phi_{b1} \tag{6.30}$$

$$\rho_1' = \frac{1}{12}r^2\left({\phi_{b0}'}^2 - \frac{\lambda}{2}\phi_{b0}^4\right), \tag{6.31}$$

where ϕ_{b0} is the bounce in the absence of gravity, the Fubini–Lipatov bounce equation (5.12) or equation (5.13). We have shown in the previous section that the first order correction to the bounce, $\phi_{b1}(r)$, does not affect the value of the bounce action. On the other hand, equation (6.22) is immediately integrated, to give

$$\rho_1(r) = \frac{1}{3|\lambda|R}\frac{(1+x^2)^2\ \arctan\, x + x(x^2-1)}{(1+x^2)^2}; \quad x = \frac{r}{R}. \tag{6.32}$$

This is sufficient to make an estimate of the impact of gravitational interactions on the decay rate of the standard model vacuum, to first order in κ. One finds that the value of the effective bounce action is shifted by

$$\Delta\Gamma[\phi_b] = \frac{256\pi^3}{45(RM_{\mathrm{Pl}\lambda})^2}. \tag{6.33}$$

There is however a difficulty that arises in the computation of higher-order corrections, as pointed out in reference [86], which is again related to the transformation properties of the action upon scale transformations. Adopting the approximate form of the scalar potential equation (6.29) a scale transformation

$$\phi_b(x) \to a\phi_b(ax), \tag{6.34}$$

$$g^{\mu\nu}(x) \to g^{\mu\nu}(ax), \tag{6.35}$$

$$\mathcal{R}(x) \to a^2\mathcal{R}(ax) \tag{6.36}$$

transforms the action as

$$S_g \to S_g(a), \tag{6.37}$$

where

$$\begin{aligned} S_g(a) &= \int \mathrm{d}^4x\,\sqrt{g(ax)}\left[\frac{a^2}{2}\frac{\partial\phi_b(ax)}{\partial x^\mu}\frac{\partial\phi_b(ax)}{\partial x^\nu}g^{\mu\nu}(ax)\right. \\ &\qquad \left. + \frac{a^4\lambda}{4}\phi_b^4(ax) - \frac{a^2\mathcal{R}(ax)}{2\kappa}\right] \\ &= \int \mathrm{d}^4x\,\sqrt{g(x)}\left[\frac{1}{2}\frac{\partial\phi_b(x)}{\partial x^\mu}\frac{\partial\phi_b(x)}{\partial x^\nu}g^{\mu\nu}(x) + \frac{\lambda}{4}\phi_b^4(x) - \frac{\mathcal{R}(x)}{2a^2\kappa}\right]. \end{aligned} \tag{6.38}$$

The stationarity condition

$$a\frac{\mathrm{d}S_g(a)}{\mathrm{d}a}\bigg|_{a=1} = \int \mathrm{d}^4x\,\sqrt{g(x)}\frac{\mathcal{R}(x)}{\kappa} = 0 \tag{6.39}$$

can only be satisfied in the flat limit $\mathcal{R}(y) = 0$, or equivalently $\rho(r) = r$. But in this case equation (6.6) gives

$$\frac{1}{2}\phi_b'^2 - V(\phi_b) = \frac{1}{2}\phi_b'^2 + \frac{|\lambda|}{4}\phi_b^4 = 0, \tag{6.40}$$

which is only possible if $\phi_b(r) = 0$, similarly to what happens in the flat-space–time case with the introduction of a mass term in the scalar potential equation (6.29): the breaking of scale invariance leads to the disappearance of the bounce. This is confirmed by an explicit calculation of the order-κ correction to the bounce. Equation (6.30) can be solved analytically [84], with the following result:

$$\begin{aligned}\phi_{b1}(r) = \frac{1}{3R^3|\lambda|^{3/2}}\frac{4\sqrt{2}}{(1+x^2)^2}\Bigg[&c_1(x^2-1)\\ &+c_2\frac{1-17x^2-x^2(1-x^2)(x^2+12\,\log\,x)}{2x^2}\\ &+\frac{1}{15x^2}\frac{1+x^2(7-9x^2)}{1+x^2} - x\,\arctan\,x + \frac{2}{5}(1-x^2)\log\,\frac{1+x^2}{x^2}\Bigg],\end{aligned} \tag{6.41}$$

where c_1 and c_2 are arbitrary constants. Since $\phi_{b0}(r)$ obeys the required asymptotic conditions for $\phi_b(r)$, namely $\phi_b'(0) = \phi_b(\infty) = 0$, the same conditions must also apply to $\phi_{b1}(r)$. We find

$$\lim_{r\to\infty}\phi_{b1}(r) = \frac{2\sqrt{2}}{3R^3|\lambda|^{3/2}}c_2 \tag{6.42}$$

and

$$\phi_{b1}(r) \sim \frac{1}{R^3|\lambda|^{3/2}}\frac{4\sqrt{2}}{3}\left(\frac{c_2}{2}+\frac{1}{15}\right)\left(\frac{1}{x^2} - 12\,\log\,x\right) + \text{regular terms} \tag{6.43}$$

for r close to zero. Hence, the two boundary conditions cannot be satisfied simultaneously.

On the other hand, when the scale dependence of the renormalized coupling is taken into account, the stationarity condition upon scale transformations becomes

$$a\frac{\mathrm{d}S_g(a)}{\mathrm{d}a}\bigg|_{a=1} = \int \mathrm{d}^4x\,\sqrt{g(x)}\left[\frac{\mathcal{R}(x)}{\kappa} + \frac{1}{4}\beta(\lambda(\phi_b))\phi_b^4\right] = 0, \tag{6.44}$$

and a non-vanishing bounce solution of the field equation can be found. We have checked that this is indeed the case by a numerical calculation (numerical methods to find the bounce in the presence of gravity are also reviewed in appendix A.6). Comparing the action computed taking gravity into account, S_g, with the action computed in a flat space–time, S, we notice that gravity stabilizes the ground state, thus making its lifetime slightly longer.

Since one of the most important parameters describing the running of the coupling constant λ is the mass of the top quark M_t, it is interesting to study how the lifetime of the ground state is modified by varying M_t within its experimental uncertainty. In the left panel of figure 12 we represent the value of the action in the presence of gravity as a function of the top quark mass. As in the flat space–time case, a lower value for the top quark stabilizes the potential, making the ground state lifetime longer.

Furthermore, from the fractional difference between the action including or neglecting gravity in the right panel of figure 12 one can see that (especially for lower values of the top quark

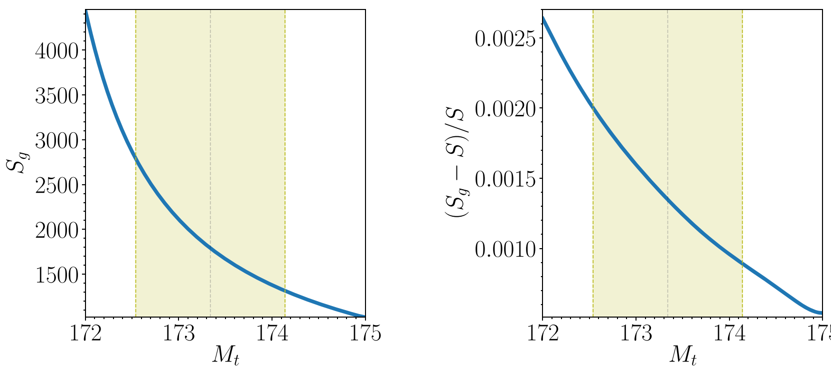

Figure 12. Left panel: value of the action as a function of the top quark mass, including gravitational effects. Right panel: fractional difference between the action including or neglecting gravity. Displayed are also the central value of the top mass and its 1σ uncertainty, $M_t = 173.34 \pm 0.8$.

mass) the bounce action becomes larger in the presence of gravity and hence the vacuum lifetime is longer. The reason of this phenomenon can be better understood by thinking at the running of λ. A lower mass for the top quark means a larger energy scale involved in the tunnelling process since the zero of β_λ is shifted at higher values, and hence gravitational effects are more important.

Finally, we remark that in the presence of new Planck-scale suppressed effective operators (cf section 5.6) the tunnelling rate might still be strongly enhanced, even after taking into account gravitational effects [87].

7. Conclusions

In this work, we have reviewed the phenomenon of quantum tunnelling both in ordinary quantum mechanics and quantum field theory. In particular, we have applied the formalism to the relevant case of the stability of the electroweak ground state in the standard model, which presents several subtle aspects related to the radiative instability of the Higgs potential and the approximate scale invariance of the electroweak theory at high-energy scales.

The present situation regarding the phase diagram of the standard model vacuum is represented in figure 10 (from reference [31]). Given the present value of the top quark mass, the standard model vacuum is metastable (with a lifetime much larger than the age of the Universe), although remarkably close to the absolute stability boundary which can be reached at the 2–3σ level, depending on the uncertainty on the top quark mass. Future improvements on the measurement of the top quark mass, which may be decisively achieved with a dedicated top-threshold lepton collider [88], will be crucial in order to establish whether the electroweak ground state is metastable or absolutely stable.

The fate of the standard model vacuum is also deeply intertwined with the cosmological evolution of the Universe, a fascinating subject that we have completely omitted in this introductory review. For instance, the electroweak vacuum instability might have played an important role for the evolution of the Higgs field during inflation (see e.g. [89]). Moreover, finite temperature effects might enhance the decay rate due to thermal fluctuations which should be taken into account for the vacuum stability analysis in the early Universe (see e.g. [90]).

Although the vacuum decay rate in quantum field theory is inherently ultraviolet sensitive (new physics in the deep ultraviolet can always enhance the decay rate), the apparent near-criticality of the standard model vacuum (under the conservative hypothesis of no new physics between the electroweak and the Planck scale) hints to a deeper structure, that is possibly related to the hidden meaning of the standard model itself.

Acknowledgments

We thank Gino Isidori, Alessandro Strumia and Lorenzo Ubaldi for discussions. The work of FD is partially supported by ERC Starting Grant 804394 hipQCD. The work of SD is supported by the Italian Ministry of Research (MIUR) under Grant PRIN 201719AVICI-01. The work of GR is partially supported by the Italian Ministry of Research (MIUR) under Grant PRIN 20172LNEEZ.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A

A.1. The semi-classical approximation in quantum mechanics

We consider a particle of mass m moving in one space dimension: x is the space coordinate, p its conjugate momentum, and $U(x)$ the potential energy. The Hamiltonian is

$$H(p, x) = \frac{p^2}{2m} + U(x). \tag{A.1}$$

In classical physics, the particle's motion is only allowed in the region defined by $E \geqslant U(x)$, where E is the total particle energy. The values of x for which $U(x) = E$ are usually called turning points, for obvious reasons.

In ordinary quantum mechanics one looks for stationary states $\psi(x)$, solutions of the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\psi''(x) + U(x)\psi(x) = E\psi(x) \tag{A.2}$$

(primes denote differentiation with respect to the argument). Apart from a few simple cases, the stationary state wave functions can be only be determined in the context of some approximation framework. In the case of barrier penetration, the so-called semi-classical approximation proves especially useful [8, 9, 91].

The classical limit of quantum mechanics, $\hbar \to 0$, is a singular limit of equation (A.2): for $\hbar = 0$ the derivative term disappears, and the equation has a different space of solutions. The correct way to take the classical limit is the following: the wave function is written

$$\psi(x) = \exp\left[i\frac{\sigma(x)}{\hbar}\right], \tag{A.3}$$

where $\sigma(x)$ is a complex function. In this way, all terms in equation (A.2) are of the same order as $\hbar \to 0$, because

$$\psi''(x) = -\frac{1}{\hbar^2}\left[\sigma'^2(x) - i\hbar\sigma''(x)\right]\psi(x). \tag{A.4}$$

Equation (A.2) takes the form

$$\sigma'^2(x) - i\hbar\sigma''(x) = 2m[E - U(x)]. \tag{A.5}$$

The complex function $\sigma(x)$ is assumed to have a power expansion in $\hbar$:

$$\sigma(x) = \sigma_0(x) + \frac{\hbar}{i}\sigma_1(x) + \left(\frac{\hbar}{i}\right)^2\sigma_2(x) + O(\hbar^3). \tag{A.6}$$

Replacing equation (A.6) in equation (A.5) we obtain a tower of differential equations for the coefficients $\sigma_n(x)$. To lowest order we get

$$\sigma_0'^2(x) = 2m[E - U(x)], \tag{A.7}$$

which is immediately integrated:

$$\sigma_0(x) = \pm\int \mathrm{d}x\, p(x), \tag{A.8}$$

where

$$p(x) = +\sqrt{2m[E - U(x)]} \tag{A.9}$$

is the classical particle momentum. Note that the term proportional to $\hbar$ in the lhs of equation (A.5) can be neglected provided

$$|\sigma'^2(x)| \gg \hbar|\sigma''(x)|, \tag{A.10}$$

or

$$\left|\hbar\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{\sigma'(x)}\right| \simeq \left|\frac{\mathrm{d}}{\mathrm{d}x}\frac{\hbar}{p(x)}\right| \ll 1, \tag{A.11}$$

which means that the approximation is only valid where the particle wavelength $\lambda = \frac{h}{p}$ does not vary too much over distances of its same order. This is clearly not the case in the vicinity of turning points, where $p(x) = 0$ and $\lambda \to \infty$.

Including one more term in the $\hbar$ expansion we get

$$2\sigma_1'(x)\sigma_0'(x) + \sigma_0''(x) = 0, \tag{A.12}$$

which gives

$$\sigma_1'(x) = -\frac{1}{2}\frac{\sigma_0''(x)}{\sigma_0'(x)} = -\frac{p'(x)}{2p(x)}; \qquad \sigma_1(x) = \log\frac{1}{\sqrt{|p(x)|}} \tag{A.13}$$

up to an arbitrary integration constant. Hence, the wave function in the semi-classical approximation, including corrections up to order $\hbar$ in $\sigma(x)$, is given by

$$\psi(x) = \frac{C_1}{\sqrt{p(x)}}e^{\frac{i}{\hbar}\int_a^x p(x')\mathrm{d}x'} + \frac{C_2}{\sqrt{p(x)}}e^{-\frac{i}{\hbar}\int_a^x p(x')\mathrm{d}x'} \tag{A.14}$$

in the classically allowed regions, where $p(x)$ is real and positive, and

$$\psi(x) = \frac{D_1}{\sqrt{|p(x)|}} e^{-\frac{1}{\hbar}\int_a^x |p(x')|\mathrm{d}x'} + \frac{D_2}{\sqrt{|p(x)|}} e^{\frac{1}{\hbar}\int_a^x |p(x')|\mathrm{d}x'} \tag{A.15}$$

in the classically forbidden ones, where $p(x)$ is purely imaginary. In the latter case, when $x \to \pm\infty$ only one of the two terms must be retained, to ensure normalizability of the wave function. Otherwise, one of the two terms is exponentially suppressed with respect to the other, and must be consistently neglected. In equations (A.14) and (A.15), a is an arbitrary reference point on the x axis. A different choice of a can be absorbed in a redefinition of the constants C_i, D_i. Higher-order terms in the expansion of $\sigma(x)$ in powers of $\hbar$ are very rarely needed.

The constants C_1, C_2, D_1, D_2 must be fixed on the basis of normalization and continuity of the wave function and its first derivative at turning points. This is not a trivial task, because the semi-classical approximation does not hold in the vicinity of a turning point. A way to circumvent this difficulty is the following. To be definite, let us assume that $x = a$ is a turning point, with $E < U(x)$ for $x > a$. Then

$$\psi(x) = \begin{cases} \dfrac{C_1}{\sqrt{p(x)}} e^{\frac{i}{\hbar}\int_a^x p(x')\mathrm{d}x'} + \dfrac{C_2}{\sqrt{p(x)}} e^{-\frac{i}{\hbar}\int_a^x p(x')\mathrm{d}x'} & x \ll a \\ \dfrac{D}{\sqrt{|p(x)|}} e^{-\frac{1}{\hbar}\int_a^x |p(x')|\mathrm{d}x'} & x \gg a \end{cases}. \tag{A.16}$$

Our task is relating D to C_1 and C_2 by requiring the continuity of the wave function and its first derivative at $x = a$. These conditions cannot be imposed directly on $\psi(x)$ as given in equation (A.16), because the semi-classical approximation fails in the vicinity of a turning point. We will instead continue analytically $\psi(x)$ in the complex plane x from the real axis to the right of the turning point, $\mathrm{Re}\, x > a$, to the real axis to the left $\mathrm{Re}\, x < a$, following a path which is never too close to the turning point. For example, we may choose a half-circle with centre a and radius r, in the upper or in the lower complex half-plane, provided r is large enough for the semi-classical approximation to hold. We further assume that we can choose r so that the potential energy $U(x)$ can be reliably expanded in Taylor series to first order in $x - a$ for all complex x with $|x - a| = r$:

$$U(x) = U(a) + U'(a)(x-a) + O(r^2) = E - F(x-a) + O(r^2), \tag{A.17}$$

with $F = -U'(a) < 0$ by assumption. Neglecting terms of order r^2 and higher in $U(x)$ the exponent of $\psi(x)$ for $x > a$ reads

$$\frac{i}{\hbar}\sigma_0(x) = -\frac{1}{\hbar}\int_a^x |p(x')|\mathrm{d}x' = -\frac{2}{3\hbar}\sqrt{2m|F|}(x-a)^{\frac{3}{2}}. \tag{A.18}$$

We now set

$$x = a + r\mathrm{e}^{\mathrm{i}\theta} \tag{A.19}$$

and continue analytically the wave function following the half-circle in the upper complex plane. We get

$$\frac{i}{\hbar}\sigma_0(x) = -\frac{2}{3\hbar}\sqrt{2m|F|}r^{\frac{3}{2}}e^{\frac{3}{2}i\theta}; \quad 0 \leqslant \theta \leqslant \pi. \tag{A.20}$$

For $\theta = \pi$

$$\frac{i}{\hbar}\sigma_0(x) = \frac{2i}{3\hbar}\sqrt{2m|F|}r^{\frac{3}{2}} = -\frac{i}{\hbar}\int_a^x p(x')\mathrm{d}x', \tag{A.21}$$

which is the exponent of the term proportional to C_2 in equation (A.16). On the other hand,

$$\frac{1}{\sqrt{|p(x)|}} = \frac{1}{(2m|F|)^{\frac{1}{4}}}(x-a)^{-\frac{1}{4}} \tag{A.22}$$

which, by the same argument, is continued to

$$\frac{1}{(2m|F|)^{\frac{1}{4}}}r^{-\frac{1}{4}}e^{-\frac{1}{4}i\pi} = \frac{1}{\sqrt{p(x)}}e^{-\frac{1}{4}i\pi}. \tag{A.23}$$

We therefore conclude that

$$C_2 = De^{-\frac{1}{4}i\pi}. \tag{A.24}$$

Following the half-circle in the lower half-plane we obtain, by an analogous argument,

$$C_1 = De^{\frac{1}{4}i\pi}. \tag{A.25}$$

Finally,

$$\psi(x) = \begin{cases} \dfrac{2D}{\sqrt{p(x)}}\cos\left[\dfrac{1}{\hbar}\displaystyle\int_a^x p(x')\mathrm{d}x' + \dfrac{\pi}{4}\right] & x \ll a \\ \dfrac{D}{\sqrt{|p(x)|}}e^{-\frac{1}{\hbar}\int_a^x |p(x')|\mathrm{d}x'} & x \gg a \end{cases}. \tag{A.26}$$

This formula was obtained in the case when the region $x > a$ is classically forbidden; it is easy to show that in the general case

$$\psi(x) = \begin{cases} \dfrac{2D}{\sqrt{p(x)}}\cos\left[\dfrac{1}{\hbar}\left|\displaystyle\int_a^x p(x')\mathrm{d}x'\right| - \dfrac{\pi}{4}\right] & U(x) < E \\ \dfrac{D}{\sqrt{|p(x)|}}e^{-\frac{1}{\hbar}\left|\int_a^x p(x')\mathrm{d}x'\right|} & U(x) > E \end{cases}. \tag{A.27}$$

A.2. Double-well potential in the semi-classical approximation

In section 2.2 we have discussed how the double-well potential can be studied in the path integral formalism. In this appendix, we review the standard formulation of the same problem in the semi-classical approximation. We will follow the discussion presented in reference [8].

We consider again the Hamiltonian of a particle moving in one dimension under the effect of a potential with two degenerate minima:

$$H = \frac{1}{2}m\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + U(x), \tag{A.28}$$

where $U(x)$ is an even function, with two minima at $x = \pm a$ and a local maximum at $x = 0$. We further assume $U(\pm a) = 0$, and define $U''(\pm a) = m\omega^2$. An example of a potential energy with these features is

$$U(x) = \frac{m\omega^2}{8a^2}(x^2 - a^2)^2, \tag{A.29}$$

but the discussion can be kept at a general level, and reference to a specific form of the potential is not needed.

If the potential energy barrier separating the two minima were infinitely high and broad, there would be two identical towers of energy eigenvalues and eigenstates, corresponding to the two potential wells at $x = \pm a$. The wave functions $\psi(\pm x)$ of the fundamental states, with energy E_0, are strongly peaked around $x = \pm a$, and decrease exponentially away from the potential minima. The function $\psi(x)$ is a solution of

$$-\frac{\hbar^2}{2m}\psi''(x) + U(x)\psi(x) = E_0\psi(x). \tag{A.30}$$

We now take quantum tunnelling into account, under the assumption that the semi-classical approximation is reliable. Because the potential is parity-even, the wave functions of the two lowest-energy states are

$$\psi_1(x) = \frac{1}{\sqrt{2}}[\psi(x) + \psi(-x)] \tag{A.31}$$

$$\psi_2(x) = \frac{1}{\sqrt{2}}[\psi(x) - \psi(-x)] \tag{A.32}$$

with energies E_1 and E_2 respectively:

$$-\frac{\hbar^2}{2m}\psi_1''(x) + U(x)\psi_1(x) = E_1\psi_1(x) \tag{A.33}$$

$$-\frac{\hbar^2}{2m}\psi_2''(x) + U(x)\psi_2(x) = E_2\psi_2(x), \tag{A.34}$$

and $E_1 < E_2$ because $\psi_2(x)$ has a node in $x = 0$. Furthermore, the two wave functions are approximately normalized, because $\psi(x)$ has almost no overlap with $\psi(-x)$. Equations (A.30), (A.33) and (A.34) can be combined and integrated, to obtain

$$-\frac{\hbar^2}{2m}\int_0^{+\infty} \mathrm{d}x\,\left[\psi_1(x)\psi''(x) - \psi_1''(x)\psi(x)\right] = (E_0 - E_1)\int_0^{+\infty} \mathrm{d}x\,\psi(x)\psi_1(x) \tag{A.35}$$

$$-\frac{\hbar^2}{2m}\int_0^{+\infty} \mathrm{d}x\,\left[\psi_2(x)\psi''(x) - \psi_2''(x)\psi(x)\right] = (E_0 - E_2)\int_0^{+\infty} \mathrm{d}x\,\psi(x)\psi_2(x). \tag{A.36}$$

The integrals are readily performed by means of equations (A.31) and (A.32), which give $\psi_1'(0) = \psi_2(0) = 0$. We obtain

$$E_1 = E_0 - \frac{\hbar^2}{m}\psi(0)\psi'(0) \tag{A.37}$$

$$E_2 = E_0 + \frac{\hbar^2}{m}\psi'(0)\psi(0) \tag{A.38}$$

and therefore

$$E_2 - E_1 = \frac{2\hbar^2}{m}\psi(0)\psi'(0). \tag{A.39}$$

It is at this point that the semi-classical approximation comes into play. The value of the wave function at the origin can be computed using the approximate expression equation (A.27) for the wave function; we find

$$\psi(0) = \frac{D}{\sqrt{|p(0)|}} e^{-\frac{1}{\hbar}\int_0^{x_-} \mathrm{d}x\,|p(x)|}, \qquad \psi'(0) = \frac{D\sqrt{|p(0)|}}{\hbar} e^{-\frac{1}{\hbar}\int_0^{x_-} \mathrm{d}x\,|p(x)|}, \quad \text{(A.40)}$$

where $x_\pm$ are the classical turning points corresponding to the energy E_0 for the potential well at $x = a$:

$$U(x_\pm) = E_0; \quad x_\pm > 0. \quad \text{(A.41)}$$

Therefore

$$E_2 - E_1 = \frac{2\hbar}{m} D^2 e^{-\frac{1}{\hbar}\int_{-x_-}^{x_-} \mathrm{d}x\,|p(x)|}, \quad \text{(A.42)}$$

where we have used $U(x) = U(-x)$.

The constant D can be obtained from the normalization condition of the wave function. Neglecting the exponentially suppressed contributions from the classically forbidden regions, we get

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 \mathrm{d}x = 4D^2 \int_{x_-}^{x_+} \mathrm{d}x\, \frac{1}{p(x)} \cos^2\left[\frac{1}{\hbar}\left|\int_{x_-}^{x} p(x')\mathrm{d}x'\right| - \frac{\pi}{4}\right] = 1. \quad \text{(A.43)}$$

We may replace the cosine squared in the integrand by its average over a period of the classical motion, $1/2$, to obtain

$$2D^2 \int_{x_-}^{x_+} \frac{\mathrm{d}x}{p(x)} = 1. \quad \text{(A.44)}$$

Furthermore,

$$\int_{x_-}^{x_+} \frac{\mathrm{d}x}{p(x)} = \int_{t_1}^{t_2} \frac{\mathrm{d}x}{\mathrm{d}t} \frac{\mathrm{d}t}{p(x)} = \frac{1}{m}(t_2 - t_1) = \frac{\pi}{m\omega}, \quad \text{(A.45)}$$

which gives

$$D^2 = \frac{m\omega}{2\pi} \quad \text{(A.46)}$$

and

$$E_2 - E_1 = \frac{\hbar\omega}{\pi} e^{-\frac{1}{\hbar}\int_{-x_-}^{x_-} \mathrm{d}x\,|p(x)|}; \quad |p(x)| = \sqrt{2m(U(x) - E_0)}; \quad U(x_-) = E_0. \quad \text{(A.47)}$$

This is how the result is usually presented in textbooks (see e.g. [8]).

Since

$$U(x) = \frac{1}{2} m\omega^2 (x-a)^2 + O((x-a)^3) \quad \text{(A.48)}$$

in the vicinity of $x = a$, we have

$$E_0 \simeq \frac{\hbar\omega}{2}; \qquad x_\pm \simeq a \pm \sqrt{\frac{\hbar}{m\omega}}, \quad \text{(A.49)}$$

and the exponent of equation (A.47) reads

$$F(\hbar) = -\frac{2}{\hbar}\int_0^{a-\sqrt{\frac{\hbar}{m\omega}}} \mathrm{d}x\, \sqrt{2mU(x) - \hbar m\omega}. \tag{A.50}$$

The exponent $F(\hbar)$ is singular as $\log\hbar$ as $\hbar \to 0$, as can be seen by observing that, for small $\hbar$, the integration variable x is close to a at the upper integration bound, and $U(x) \sim \frac{1}{2}m\omega^2(a-x)^2$ in this region. In order take this singularity into account, we split the integration range in equation (A.50):

$$F(\hbar) = -\frac{2}{\hbar}\int_0^{a-\epsilon} \mathrm{d}x\, \sqrt{2mU(x) - \hbar m\omega} - \frac{2}{\hbar}\int_{a-\epsilon}^{a-\sqrt{\frac{\hbar}{m\omega}}} \mathrm{d}x\, \sqrt{2mU(x) - \hbar m\omega}, \tag{A.51}$$

where ϵ is an arbitrary constant in the range $\sqrt{\frac{\hbar}{m\omega}} < \epsilon < a$. The integration range in the first integral now does not contain the singular point, and we may expand the integrand in powers of $\hbar$ to first order, while in the second integral we may approximate $U(x)$ as in equation (A.48) if ϵ is not too large. We get

$$\begin{aligned} F(\hbar) = &-\frac{2}{\hbar}\int_0^{a-\epsilon} \mathrm{d}x\, \sqrt{2mU(x)} + m\omega \int_0^{a-\epsilon} \frac{\mathrm{d}x}{\sqrt{2mU(x)}} \\ &- \frac{2m\omega}{\hbar}\int_{a-\epsilon}^{a-\sqrt{\frac{\hbar}{m\omega}}} \mathrm{d}x\, \sqrt{(a-x)^2 - \frac{\hbar}{m\omega}} + O(\hbar). \end{aligned} \tag{A.52}$$

The last integral can be computed analytically. Neglecting $O(\hbar)$ corrections, and using equation (A.48) again whenever allowed, we get

$$\begin{aligned} F(\hbar) = &-\frac{2}{\hbar}\int_0^{a} \mathrm{d}x\, \sqrt{2mU(x)} - \frac{2m\omega}{\hbar}\int_a^{a-\epsilon} \mathrm{d}x\,(a-x) \\ &+ \int_0^{a-\epsilon} \mathrm{d}x \left[\frac{m\omega}{\sqrt{2mU(x)}} - \frac{1}{a-x}\right] + \int_0^{a-\epsilon} \mathrm{d}x\, \frac{1}{a-x} \\ &- \log\left[\sqrt{\frac{m\omega}{\hbar}}\left(\epsilon - \sqrt{\epsilon^2 - \frac{\hbar}{m\omega}}\right)\right] - \frac{m\omega}{\hbar}\epsilon\sqrt{\epsilon^2 - \frac{\hbar}{m\omega}} \\ = &-\frac{2}{\hbar}\int_0^{a} \mathrm{d}x\, \sqrt{2mU(x)} + \int_0^{a-\epsilon} \mathrm{d}x \left[\frac{m\omega}{\sqrt{2mU(x)}} - \frac{1}{a-x}\right] \\ &- \log\left[\frac{\epsilon^2}{a}\sqrt{\frac{m\omega}{\hbar}}\left(1 - \sqrt{1 - \frac{\hbar}{m\omega\epsilon^2}}\right)\right] + \frac{m\omega\epsilon^2}{\hbar}\left(1 - \sqrt{1 - \frac{\hbar}{m\omega\epsilon^2}}\right). \end{aligned} \tag{A.53}$$

In the semi-classical limit $\hbar \ll \epsilon^2 m\omega$ we have

$$1 - \sqrt{1 - \frac{\hbar}{m\omega\epsilon^2}} = \frac{\hbar}{2m\omega\epsilon^2} + O(\hbar^2), \tag{A.54}$$

so that

$$F(\hbar) = -\frac{2}{\hbar}\int_0^a \mathrm{d}x\,\sqrt{2mU(x)} + \int_0^{a-\epsilon}\mathrm{d}x\left[\frac{m\omega}{\sqrt{2mU(x)}} - \frac{1}{a-x}\right] + \log\left[2a\sqrt{\frac{m\omega}{\hbar}}\right] + \frac{1}{2}. \tag{A.55}$$

The only residual dependence on the arbitrary parameter ϵ is in the upper integration bound of the second integral, where it has a negligible effect because the integrand is essentially zero in the range $a - \epsilon < x < a$. We have therefore

$$E_2 - E_1 = \frac{\hbar\omega}{\pi}\,\exp F(\hbar) = \sqrt{\frac{e}{\pi}}\hbar\omega\left[2\hat{a}\sqrt{\frac{m\omega}{\pi\hbar}}\right]\exp\left[-\frac{1}{\hbar}\int_{-a}^{a}\mathrm{d}x\,\sqrt{2mU(x)}\right], \tag{A.56}$$

where

$$\hat{a} = a\,\exp\int_0^a \mathrm{d}x\left[\frac{m\omega}{\sqrt{2mU(x)}} - \frac{1}{a-x}\right]. \tag{A.57}$$

The result equations (A.56) and (A.57) must be compared with the result equations (2.123) and (2.124) with $U_0 = 0$, obtained with the path integral method. We see that the two expressions differ by a factor

$$\sqrt{\frac{e}{\pi}} \sim 0.93. \tag{A.58}$$

This small discrepancy can be removed by a more accurate treatment of the matching of the wave function at classical turning points in the semi-classical approximation, see for example [9].

A.3. The Maupertuis principle

The Maupertuis principle is a modified version of Hamilton's principle of classical mechanics, which allows one to determine the trajectory of the actual motion (with no reference to the time dependence of the coordinates), when energy is conserved.

We consider the action of a system with N degrees of freedom as a function of the generalized coordinates at the final time t, $\mathbf{x}(t)$ and of t itself:

$$S(\mathbf{x}(t),t) = \int_{t_0}^{t}\mathrm{d}t'\,L(\mathbf{x}(t'),\dot{\mathbf{x}}(t')) = \int_{t_0}^{t}\mathrm{d}t'\left[\frac{1}{2}\dot{x}_i\dot{x}_i - U(\mathbf{x})\right], \tag{A.59}$$

where a sum over the repeated index i is understood. Here $\mathbf{x}(t')$ is a solution of the equation of motion, with $\mathbf{x}(t_0) = \mathbf{a}$. We find

$$\begin{aligned}\frac{\partial}{\partial x_i(t)}S(\mathbf{x}(t),t) &= \int_{t_0}^{t}\mathrm{d}t'\left[\frac{\partial L}{\partial x_j(t')}\frac{\partial x_j(t')}{\partial x_i(t)} + \frac{\partial L}{\partial \dot{x}_j(t')}\frac{\mathrm{d}}{\mathrm{d}t'}\frac{\partial x_j(t')}{\partial x_i(t)}\right] \\ &= \int_{t_0}^{t}\mathrm{d}t'\left[\frac{\partial L}{\partial x_j(t')} - \frac{\mathrm{d}}{\mathrm{d}t'}\frac{\partial L}{\partial \dot{x}_j(t')}\right]\frac{\partial x_j(t')}{\partial x_i(t)} \\ &\quad + \left.\frac{\partial L}{\partial \dot{x}_j(t')}\frac{\partial x_j(t')}{\partial x_i(t)}\right|_{t_0}^{t}.\end{aligned} \tag{A.60}$$

Since $\mathbf{x}(t')$ is assumed to be a solution of the Euler–Lagrange equations, the integrand in the first term vanishes identically, and we are left with the result

$$\frac{\partial}{\partial x_i(t)} S(\mathbf{x}(t), t) = p_i(t), \tag{A.61}$$

where

$$p_i(t) = \frac{\partial L}{\partial \dot{x}_i(t)} = \dot{x}_i(t) \tag{A.62}$$

is the conjugate momentum of x_i at time t. Also,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} S(\mathbf{x}(t), t) &= \frac{\partial}{\partial t} S(\mathbf{x}(t), t) + \frac{\partial S(\mathbf{x}(t), t)}{\partial x_i(t)} \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} \\ &= \frac{\partial}{\partial t} S(\mathbf{x}(t), t) + p_i \dot{x}_i(t). \end{aligned} \tag{A.63}$$

On the other hand, by definition,

$$\frac{\mathrm{d}}{\mathrm{d}t} S(\mathbf{x}(t), t) = L(\mathbf{x}(t), \dot{\mathbf{x}}(t)). \tag{A.64}$$

By comparison,

$$\frac{\partial}{\partial t} S(\mathbf{x}(t), t) = L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - p_i(t)\dot{x}_i(t) = -H(\mathbf{p}(t), \mathbf{x}(t)). \tag{A.65}$$

The action can therefore be written

$$S(\mathbf{x}(t), t) = S_0(\mathbf{x}(t)) - \int_{t_0}^{t} \mathrm{d}t'\, H(\mathbf{p}(t'), \mathbf{x}(t')), \tag{A.66}$$

where

$$S_0(\mathbf{x}(t)) = \sum_i \int_{a_i}^{x_i(t)} p_i \, \mathrm{dx}_i \tag{A.67}$$

is sometimes called the reduced action, and $\mathbf{a}$ is a reference point in configuration space. When H is constant in time,

$$H(p_i(t), x_i(t)) = \frac{1}{2} \sum_i \frac{\mathrm{d}x_i}{\mathrm{d}t} \frac{\mathrm{d}x_i}{\mathrm{d}t} + U(\mathbf{x}) = E, \tag{A.68}$$

we get

$$S(\mathbf{x}(t), t) = S_0(\mathbf{x}(t)) - E(t - t_0). \tag{A.69}$$

We know that the action S is stationary in correspondence of the solutions of the equations of motion, provided the variation is performed with $\mathbf{x}(t_0)$, $\mathbf{x}(t)$ and $t - t_0$ fixed. If we now allow the time interval to vary as well, we obtain, from the above argument,

$$\delta S = -E\delta t. \tag{A.70}$$

On the other hand, from equation (A.69) we get

$$\delta S = \delta S_0 - E\delta t, \tag{A.71}$$

and hence, comparing equations (A.70) and (A.71),

$$\delta S_0 = 0. \tag{A.72}$$

Thus we have shown that the reduced action is stationary in correspondence of the solutions of the equations of motion if energy is conserved.

We may now choose a parametrization of the trajectory, $\mathbf{x} = \mathbf{x}(s)$, with $0 \leqslant s \leqslant s_f$, $\mathbf{x}(0) = \mathbf{x}_0, \mathbf{x}(s_f) = \mathbf{X}$, such that

$$\left|\frac{\mathrm{d}\mathbf{x}(s)}{\mathrm{d}s}\right|^2 = \sum_i \frac{\mathrm{d}x_i(s)}{\mathrm{d}s}\frac{\mathrm{d}x_i(s)}{\mathrm{d}s} = 1 \tag{A.73}$$

or equivalently

$$\sum_i \frac{\mathrm{d}x_i}{\mathrm{d}t}\frac{\mathrm{d}x_i}{\mathrm{d}t} = \sum_i \frac{\mathrm{d}x_i}{\mathrm{d}s}\frac{\mathrm{d}x_i}{\mathrm{d}s}\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2. \tag{A.74}$$

Using equation (A.68),

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \sqrt{2[E - U(\mathbf{x})]}. \tag{A.75}$$

Then

$$\begin{aligned} S_0(\mathbf{x}(t), t) &= \sum_i \int_{a_i}^{x_i(t)} p_i \, \mathrm{d}x_i = \int_0^{s_f} \left(\sum_i \frac{\mathrm{d}x_i}{\mathrm{d}s}\frac{\mathrm{d}x_i}{\mathrm{d}s}\right)\frac{\mathrm{d}s}{\mathrm{d}t}\mathrm{d}s \\ &= \int_0^{s_f} \sqrt{2[E - U(\mathbf{x}(s))]}\mathrm{d}s, \end{aligned} \tag{A.76}$$

where we have used equations (A.75) and (A.73).

A.4. Legendre transforms

Given a function f of a single variable x, its Legendre transform, denoted by $g(p)$, is defined as

$$g(p) = \max_x [px - f(x)]. \tag{A.77}$$

Graphically this corresponds to drawing a line through the origin with slope p and taking a point $x = x_0(p)$ such that the vertical distance between $f(x)$ and the line px is maximized. The maximization of the function $px - f(x)$ implies:

$$\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=x_0} = p \quad \text{and} \quad \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x=x_0} \geqslant 0, \tag{A.78}$$

so that we can equivalently rewrite

$$g(p) = x_0(p)p - f(x_0(p)), \tag{A.79}$$

whose form resembles that of the effective action, after the proper generalization to a functional space (cf equation (4.5)). The first equation in (A.78) implies that the line tangent to f at the point x_0 must have slope p, i.e. it must be parallel to the line px. Note that if $f(x)$ is convex

($f''(x) \geqslant 0$ everywhere) there will be only one value of x_0, otherwise there might be several and one has to pick up that which maximizes $px - f(x)$.

The Legendre transform has the following properties (which we report without proof):

- g is always a convex function (even if f is non convex).
- The double Legendre transform of a convex function f is f itself. If f is non convex, the double Legendre transform of f is its convex envelope.

A.5. The Fubini–Lipatov bounce

Solutions of the second-order differential equation

$$\phi''(r) + \frac{3}{r}\phi'(r) = \lambda\phi^3(r) \tag{A.80}$$

with $\phi(0) > 0, \phi'(0) = 0, \phi(\infty) = 0$ can be found in the form of a Taylor expansion around $r = 0$:

$$\phi(r) = \sum_{k=0}^{\infty} A_k r^k \tag{A.81}$$

with $A_0 > 0$. Equation (A.80) takes the form of a recurrence relation for the coefficients A_k:

$$\frac{3A_1}{r} + \sum_{k=0}^{\infty}(k+2)(k+4)A_{k+2}r^k = \lambda\sum_{k=0}^{\infty} r^k \sum_{i=0}^{k}\sum_{j=0}^{k-i} A_iA_jA_{k-i-j}. \tag{A.82}$$

It follows that

$$A_1 = 0. \tag{A.83}$$

The first derivative of the solution at $r = 0$ vanishes as a consequence of the assumption equation (A.81). The remaining coefficients are given by

$$A_{k+2} = \frac{\lambda}{(k+2)(k+4)}\sum_{i=0}^{k}\sum_{j=0}^{k-i} A_iA_jA_{k-i-j}. \tag{A.84}$$

The coefficients A_k with k odd are zero: indeed, when k is odd, at least one of the three summation indices $i, j, k - i - j$ is also odd. Hence $A_1 = 0$ implies $A_3 = 0$, and so on. Thus, we may rewrite equation (A.81) as

$$\phi(r) = \sum_{k=0}^{\infty} a_k r^{2k}; \quad a_k = A_{2k}, \tag{A.85}$$

and the recurrence relation for the coefficients becomes

$$a_{k+1} = \frac{\lambda}{8}\frac{2}{(k+1)(k+2)}\sum_{i=0}^{k}\sum_{j=0}^{k-i} a_ia_ja_{k-i-j}. \tag{A.86}$$

The coefficients a_k are determined by the single number a_0, the value of $\phi(r)$ at $r = 0$. We now show that

$$a_j = \left(\frac{\lambda}{8}\right)^j a_0^{2j+1}. \tag{A.87}$$

The proof is by induction. For $k = 0$ equation (A.86) gives

$$a_1 = \frac{\lambda}{8} a_0^3. \tag{A.88}$$

We now assume that equation (A.87) holds for $0 \leqslant j \leqslant k$. Then

$$\begin{aligned} a_{k+1} &= \frac{\lambda}{8} \frac{2}{(k+1)(k+2)} \sum_{i=0}^{k} \sum_{j=0}^{k-i} a_i a_j a_{k-i-j} \\ &= \frac{\lambda}{8} \frac{2}{(k+1)(k+2)} \sum_{i=0}^{k} \sum_{j=0}^{k-i} \left(\frac{\lambda}{8}\right)^k a_0^{2k+3} = \left(\frac{\lambda}{8}\right)^{k+1} a_0^{2k+3} \end{aligned} \tag{A.89}$$

which is what we set out to prove. The Taylor expansion in equation (A.85) can now be summed. We find

$$\phi(r) = a_0 \sum_{k=0}^{\infty} \left(\frac{\lambda}{8}\right)^k a_0^{2k} r^{2k} = \frac{a_0}{1 - \frac{\lambda}{8} a_0^2 r^2}. \tag{A.90}$$

The series has convergence radius

$$R = \sqrt{\frac{8}{|\lambda|}} \frac{1}{a_0}. \tag{A.91}$$

If $\lambda < 0$ the sum can be analytically continued to the whole positive real axis, and vanishes as $r \to \infty$. Finally, we note that for $\lambda < 0$ the bounce in equation (A.90) coincides with the solution given in equation (5.13), with R as in equation (A.91). The above construction shows that, given the value of the bounce at the origin and the requirement of regularity in the range $0 < r < \infty$, the solution is unique. An alternative derivation of the result equation (A.90) is given in reference [16].

A.6. Numerical methods for the bounce equation

Some of the results presented in this review have been obtained by numerical methods. This appendix contains an illustration of the techniques adopted for the numerical determination of the bounce solutions for a generic potential, both in the flat and curved space–time cases.

A.6.1. Flat space. It is convenient to introduce dimensionless variables:

$$x = Mr\,; \qquad \tilde{\phi}(x) = \frac{\phi(r)}{M}\,; \qquad \tilde{V}(\tilde{\phi}) = \frac{V(\phi)}{M^4}, \tag{A.92}$$

where M is an arbitrary mass scale. It will prove convenient to choose M of the same order of magnitude as the value of the bounce in its centre, $M \simeq \phi(0)$, in order to keep the value of $\tilde{\phi}$ of order one.

In terms of dimensionless variables, the field equation reads

$$\tilde{\phi}''(x) + \frac{3}{x}\tilde{\phi}'(x) = \tilde{V}'(\tilde{\phi}), \tag{A.93}$$

while the boundary conditions become

$$\lim_{x\to\infty}\tilde{\phi}(x)=\frac{\phi_{\mathrm{FV}}}{M} \tag{A.94}$$

$$\left.\frac{\mathrm{d}\tilde{\phi}(x)}{\mathrm{d}x}\right|_{x=0}=0, \tag{A.95}$$

ϕ_{FV} being the value of the field in the false vacuum configuration.

Finding a numerical solution of equation (A.93) is not straightforward since the boundary conditions equations (A.94) and (A.95) do not define a Cauchy problem, but rather a boundary value problem. The solution can be found via a *shooting* method. We define a Cauchy problem choosing as boundary conditions equation (A.95) and

$$\tilde{\phi}(0)=a. \tag{A.96}$$

This problem has a unique solution, which can be found by means of available numerical routines, as a function of the parameter a. Next, the value of a is adjusted so that the boundary condition equation (A.94) is fulfiled.

Note that the origin $x=0$ is a singular point for equation (A.93): thus, the numerical solution must be looked for in a range $x_{\min}\leqslant x\leqslant x_{\max}$ with $x_{\min}>0$. As a consequence, we need the initial conditions at $x=x_{\min}$ which correspond to those at $x=0$. To this purpose, we relate $\phi(x_{\min})$ to $\phi(0)$ by a Taylor expansion:

$$\begin{aligned}\tilde{\phi}(x_{\min})&=a+\tilde{\phi}'(0)x_{\min}+\frac{1}{2}\tilde{\phi}''(0)x_{\min}^2+O(x_{\min}^3)\\&=a+\frac{1}{2}\tilde{\phi}''(0)x_{\min}^2+O(x_{\min}^3),\end{aligned} \tag{A.97}$$

where we have used equation (A.95). The field equation equation (A.93) gives

$$\tilde{\phi}''(0)=\tilde{V}(\tilde{\phi}(0))-\left.\frac{3}{x}\tilde{\phi}'(x)\right|_{x=0}=\tilde{V}(a)-3\tilde{\phi}''(0), \tag{A.98}$$

which gives

$$\tilde{\phi}''(0)=\frac{1}{4}\tilde{V}'(a). \tag{A.99}$$

So finally

$$\tilde{\phi}(x_{\min})=a+\frac{1}{8}\tilde{V}'(a)x_{\min}^2+O(x_{\min}^3) \tag{A.100}$$

$$\tilde{\phi}'(x_{\min})=\frac{1}{4}\tilde{V}'(a)x_{\min}+O(x_{\min}^3). \tag{A.101}$$

The Cauchy problem which determines the bounce can now be solved numerically.

From the numerical point of view, the calculation of the Euclidean action is most conveniently performed by means of equation (3.12). Exploiting the $O(4)$ symmetry of the bounce, we find

$$S_E=-2\pi^2\int_{x_{\min}}^{x_{\max}}\mathrm{d}x\,x^3\tilde{V}(\tilde{\phi}(x)). \tag{A.102}$$

As a final step, one iterates the procedure decreasing x_{min} and increasing x_{max} until the result is numerically stable.

A.6.2. Curved space. In the presence of gravitational effects, the algorithm closely follows the one described in the case of flat space–time. We have now to solve the coupled equations (6.5) and (6.6). A complication immediately arises, due to the square root ambiguity in equation (6.6). To get rid of it we differentiate equation (6.6) with respect to r [84]. We get

$$\rho'' = -\frac{\kappa}{3}\rho(\phi_b'^2 + V(\phi_b)). \tag{A.103}$$

Following the same strategy as in the flat case, we introduce the dimensionless variables

$$x = Mr\,; \qquad \tilde{\phi}(x) = \frac{\phi_b(r)}{M}\,; \qquad \tilde{V}(\tilde{\phi}) = \frac{V(\phi_b)}{M^4}; \qquad \tilde{\rho}(x) = \rho(r), \tag{A.104}$$

where M is an arbitrary mass scale. In terms of the new variables, the system of differential equations becomes

$$\tilde{\phi}'' = -3\frac{\tilde{\rho}'}{\tilde{\rho}}\tilde{\phi}' + \frac{\mathrm{d}\tilde{V}}{\mathrm{d}\tilde{\phi}}, \tag{A.105}$$

$$\tilde{\rho}'' = -\frac{8\pi}{3}\epsilon^2\tilde{\rho}\left(\tilde{\phi}'^2 + \tilde{V}\right), \tag{A.106}$$

where $\kappa = 8\pi/M_{\mathrm{Pl}}^2$, $\epsilon = M/M_{\mathrm{Pl}}$.

The boundary conditions for ϕ are the same as in flat space. Thus it is not a Cauchy problem and we have to apply a shooting algorithm similarly to the case in absence of gravity.

On the other hand the boundary condition for the function ρ are

$$\rho'(0) = 1, \tag{A.107}$$

$$\rho(0) = 0 \tag{A.108}$$

from equations (6.6) and (6.7).

The boundary conditions in terms of dimensionless variables read

$$\lim_{x\to\infty}\tilde{\phi}(x) = \frac{\phi_{\mathrm{FV}}}{M}, \tag{A.109}$$

$$\left.\frac{\mathrm{d}\tilde{\phi}(x)}{\mathrm{d}x}\right|_{x=0} = 0, \tag{A.110}$$

$$\tilde{\rho}(0) = 0, \tag{A.111}$$

$$\tilde{\rho}'(0) = 1. \tag{A.112}$$

Proceeding as in the previous case, we solve numerically the two differential equations equations (A.105) and (A.106) in the range $(x_{\rm min}, x_{\rm max})$ with initial conditions

$$\tilde{\phi}(x_{\rm min}) = a + \frac{1}{8}\frac{\mathrm{d}\tilde{V}}{\mathrm{d}\tilde{\phi}}\bigg|_{\tilde{\phi}=a} x_{\rm min}^2 + \mathcal{O}(x_{\rm min}^4), \tag{A.113}$$

$$\tilde{h}'(x_{\rm min}) = \frac{1}{4}\frac{\mathrm{d}\tilde{V}}{\mathrm{d}\tilde{\phi}}\bigg|_{\tilde{h}=a} x_{\rm min} + \mathcal{O}(x_{\rm min}^3), \tag{A.114}$$

$$\tilde{\rho}(x_{\rm min}) = x_{\rm min}, \tag{A.115}$$

$$\tilde{\rho}'(x_{\rm min}) = 1 \tag{A.116}$$

and we adjust the value of a so that equation (A.109) is fulfiled. The algorithm is iterated lowering $x_{\rm min}$ and increasing $x_{\rm max}$, until the desired level of accuracy is achieved.

A.7. Gravitational field equations

In this appendix we provide some details about the field equations in the case of a scalar field minimally coupled to gravity. Following reference [5], we add the Euclidean version of the Einstein–Hilbert action to the action of the scalar field[16]:

$$S[\phi, g] = S_g[g] + S_\phi[\phi, g], \tag{A.117}$$

with

$$S_g[g] = \int \mathrm{d}^4x\,\sqrt{g}\left[-\frac{1}{2\kappa}\mathcal{R}\right], \tag{A.118}$$

$$S_\phi[\phi, g] = \int \mathrm{d}^4x\,\sqrt{g}\left[\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + V(\phi)\right], \tag{A.119}$$

where $\kappa = 8\pi G_{\rm N} = 8\pi/M_{\rm Pl}^2$, $\mathcal{R}$ is the Riemann curvature, and the metric tensor $g_{\mu\nu}$ is understood to have a positive-definite Euclidean signature. Note that, for a scalar field, there is no difference between ordinary and covariant derivative: $\nabla_\mu\phi = \partial_\mu\phi$.

There is an important difference with respect to the flat space–time case: without gravity, an additive constant to the scalar potential has no effect on the field equations. When gravity is taken into account, instead, a constant term in the scalar potential contributes with a term proportional $\sqrt{g}$ to the action integrand, which results in a contribution to the field equations via a cosmological constant. Remarkably, vacuum decay affects gravitation itself: when tunnelling takes place, the value of the potential at the minimum (and hence the cosmological constant) gets modified.

We first derive the Euclidean field equations in the presence of gravity. By taking the variation of $S[\phi, g]$ with respect to the metric we obtain the Einstein equations

$$\mathcal{G}_{\mu\nu} = -\kappa\mathcal{T}_{\mu\nu}, \tag{A.120}$$

where $\mathcal{G}_{\mu\nu}$ is the Einstein tensor

$$\mathcal{G}_{\mu\nu} = \mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}\,g_{\mu\nu}, \tag{A.121}$$

[16] We neglect a non-minimal coupling of gravity to the Higgs of the form $\xi\mathcal{R}|H|^2$. The effect of this operator on the tunnelling rate of the electroweak vacuum has been considered e.g. in [85, 92].

and $\mathcal{T}_{\mu\nu}$ the stress–energy tensor

$$\mathcal{T}_{\mu\nu} = -\frac{2}{\sqrt{g}}\frac{\delta S_\phi}{\delta g^{\mu\nu}} = -\partial_\mu\phi\partial_\nu\phi + g_{\mu\nu}\left[\frac{1}{2}g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi + V(\phi)\right]. \tag{A.122}$$

The minus sign in equation (A.120) is due to the fact that the metric is Euclidean.

The field equation for the scalar field is obtained by taking the variation with respect to ϕ:

$$\begin{aligned}\delta_\phi S[\phi, g] &= \int \mathrm{d}^4x\,\sqrt{g}\left[\frac{\mathrm{d}V}{\mathrm{d}\phi}\delta\phi + \frac{1}{2}g^{\mu\nu}\frac{\partial(\partial_\mu\phi\partial_\nu\phi)}{\partial\partial_\rho\phi}\delta(\partial_\rho\phi)\right]\\ &= \int \mathrm{d}^4x\,\sqrt{g}\left[\frac{\mathrm{d}V}{\mathrm{d}\phi} + g^{\mu\nu}(\partial_\mu\phi)\partial_\nu\right]\delta\phi\\ &= \int \mathrm{d}^4x\,\left[\sqrt{g}\,\frac{\mathrm{d}V}{\mathrm{d}\phi} - \partial_\nu(\sqrt{g}g^{\mu\nu}(\partial_\mu\phi))\right]\delta\phi + \text{surface term}\\ &= \int \mathrm{d}^4x\,\sqrt{g}\left[\frac{\mathrm{d}V}{\mathrm{d}\phi} - \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}(\partial^\mu\phi))\right]\delta\phi + \text{surface term},\end{aligned} \tag{A.123}$$

which yields

$$\frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}\partial^\mu\phi) = \frac{\mathrm{d}V}{\mathrm{d}\phi}. \tag{A.124}$$

To find the bounce we thus have to solve the two coupled equations equations (A.120) and (A.124) for the field ϕ and the metric g. The solution, once replaced in equation (A.117), gives the value of the bounce action, relevant for the calculation of the tunnelling rate.

ORCID iDs

Federica Devoto https://orcid.org/0000-0001-6397-920X
Simone Devoto https://orcid.org/0000-0001-9146-190X
Luca Di Luzio https://orcid.org/0000-0002-3740-7846
Giovanni Ridolfi https://orcid.org/0000-0003-4887-4359

References

[1] Coleman S 1977 Fate of the false vacuum: semiclassical theory *Phys. Rev.* D **15** 2929–36
[2] Callan C G, Coleman S and Coleman S R 1977 Fate of the false vacuum: II. First quantum corrections *Phys. Rev.* D **16** 1762–8
[3] Isidori G, Ridolfi G and Strumia A 2001 On the metastability of the standard model vacuum *Nucl. Phys.* B **609** 387–409
[4] Branchina V and Messina E 2013 Stability, Higgs boson mass, and new physics *Phys. Rev. Lett.* **111** 241801
[5] Coleman S and De Luccia F 1980 Gravitational effects on and of vacuum decay *Phys. Rev.* D **21** 3305
[6] Isidori G, Rychkov V S, Strumia A and Tetradis N 2008 Gravitational corrections to standard model vacuum decay *Phys. Rev.* D **77** 025034
[7] Weinberg E J 2012 *Classical Solutions in Quantum Field Theory: Solitons and Instantons in High Energy Physics* (*Cambridge Monographs on Mathematical Physics* vol 9) (Cambridge: Cambridge University Press)

[8] Landau L and Lifshitz E 2013 *Quantum Mechanics: Non-relativistic Theory* (Amsterdam: Elsevier) Teoreticheskaia Fizika
[9] Konishi K and Paffuti G 2009 *Quantum Mechanics* (Oxford: Oxford University Press)
[10] Coleman S 1985 *Aspects of Symmetry* (Cambridge: Cambridge University Press)
[11] Weinberg S 2013 *The Quantum Theory of Fields* (*Modern Applications* vol 2) (Cambridge: Cambridge University Press)
[12] Banks T, Bender C M and Wu T T 1973 Coupled anharmonic oscillators: I. Equal-mass case *Phys. Rev.* D **8** 3346–66
[13] Banks T and Bender C M 1973 Coupled anharmonic oscillators: II. Unequal-mass case *Phys. Rev.* D **8** 3366–78
[14] Coleman S, Glaser V and Martin A 1978 Action minima among solutions to a class of Euclidean scalar field equations *Commun. Math. Phys.* **58** 211
[15] Ivanov A, Matteini M, Nemevšek M and Ubaldi L 2022 Analytic thin wall false vacuum decay rate *J. High Energy Phys.* JHEP03(2022)209
[16] Lee K and Weinberg E J 1986 Tunneling without barriers *Nucl. Phys.* B **267** 181
[17] Kusenko A, Lee K and Weinberg E J 1997 Vacuum decay and internal symmetries *Phys. Rev.* D **55** 4903–9
[18] Fubini S 1976 A new approach to conformal invariant field theories *Nuovo Cimento* A **34** 521
[19] Lipatov L N 1977 Divergence of the perturbation theory series and the quasiclassical theory *Sov. Phys. JETP* **45** 216–23
Lipatov L N 1977 *Zh. Eksp. Teor. Fiz.* **72** 411
[20] Di Luzio L, Isidori G and Ridolfi G 2016 Stability of the electroweak ground state in the standard model and its extensions *Phys. Lett.* B **753** 150–60
[21] Jona-Lasinio G 1964 Relativistic field theories with symmetry-breaking solutions *Nuovo Cimento* **34** 1790–5
[22] Coleman S and Weinberg E 1973 Radiative corrections as the origin of spontaneous symmetry breaking *Phys. Rev.* D **7** 1888–910
[23] Jackiw R 1974 Functional evaluation of the effective potential *Phys. Rev.* D **9** 1686
[24] Martin S P 2017 Effective potential at three loops *Phys. Rev.* D **96** 096005
[25] Collins J C 1986 *Renormalization* (*Cambridge Monographs on Mathematical Physics* vol 26) (Cambridge: Cambridge University Press)
[26] Peskin M E and Schroeder D V 1995 *An Introduction to Quantum Field Theory* (Reading, MA: Addison-Wesley Developers Press)
[27] Bardeen W A, Buras A J, Duke D W and Muta T 1978 Deep-inelastic scattering beyond the leading order in asymptotically free gauge theories *Phys. Rev.* D **18** 3998
[28] Chetyrkin K G and Zoller M F 2012 Three-loop β-functions for top-Yukawa and the Higgs self-interaction in the standard model *J. High Energy Phys.* JHEP06(2012)033
[29] Mihaila L N, Salomon J and Steinhauser M 2012 Renormalization constants and beta functions for the gauge couplings of the standard model to three-loop order *Phys. Rev.* D **86** 096008
[30] Ford C, Jones D R T, Stephenson P W and Einhorn M B 1993 The effective potential and the renormalisation group *Nucl. Phys.* B **395** 17–34
[31] Buttazzo D, Degrassi G, Giardino P P, Giudice G F, Sala F, Salvio A and Strumia A 2013 Investigating the near-criticality of the Higgs boson *J. High Energy Phys.* JHEP12(2013)089
[32] Nielsen N K 1975 On the gauge dependence of spontaneous symmetry breaking in gauge theories *Nucl. Phys.* B **101** 173–88
[33] Metaxas D and Weinberg E J 1996 Gauge independence of the bubble nucleation rate in theories with radiative symmetry breaking *Phys. Rev.* D **53** 836–43
[34] Di Luzio L and Mihaila L 2014 On the gauge dependence of the standard model vacuum instability scale *J. High Energy Phys.* JHEP06(2014)079
[35] Andreassen A, Frost W and Schwartz M D 2015 Consistent use of effective potentials *Phys. Rev.* D **91** 016009
[36] Andreassen A, Frost W and Schwartz M D 2014 Consistent use of the standard model effective potential *Phys. Rev. Lett.* **113** 241801
[37] Espinosa J R, Garny M, Konstandin T and Riotto A 2017 Gauge-independent scales related to the standard model vacuum instability *Phys. Rev.* D **95** 056004
[38] Fujimoto Y, O'Raifeartaigh L and Parravicini G 1983 Effective potential for non-convex potentials *Nucl. Phys.* B **212** 268–300
[39] Dannenberg A 1988 Dysfunctional methods and the effective potential *Phys. Lett.* B **202** 110–6

[40] Sher M 1989 Electroweak Higgs potential and vacuum stability *Phys. Rep.* **179** 273–418
[41] Brandenberger R H 1985 Quantum field theory methods and inflationary Universe models *Rev. Mod. Phys.* **57** 1
[42] Weinberg E J and Wu A 1987 Understanding complex perturbative effective potentials *Phys. Rev.* D **36** 2474
[43] Krive I V and Linde A D 1976 On the vacuum stability in the σ model *Nucl. Phys.* B **117** 265–8
[44] Krasnikov N V 1978 Restriction of the fermion mass in gauge theories of weak and electromagnetic interactions *Yad. Fiz.* **28** 549–51
[45] Maiani L, Parisi G and Petronzio R 1978 Bounds on the number and masses of quarks and leptons *Nucl. Phys.* B **136** 115–24
[46] Politzer H D and Wolfram S 1979 Bounds on particle masses in the Weinberg–Salam model *Phys. Lett.* B **82** 242–6
Politzer H D and Wolfram S 1979 *Phys. Lett.* B **83** 421 (erratum)
[47] Hung P Q 1979 Vacuum instability and new constraints on fermion masses *Phys. Rev. Lett.* **42** 873
[48] Cabibbo N, Maiani L, Parisi G and Petronzio R 1979 Bounds on the fermions and Higgs boson masses in grand unified theories *Nucl. Phys.* B **158** 295–305
[49] Linde A D 1980 Vacuum instability, cosmology and constraints on particle masses in the Weinberg–Salam model *Phys. Lett.* B **92** 119–21
[50] Lindner M, Sher M and Zaglauer H W 1989 Probing vacuum stability bounds at the fermilab collider *Phys. Lett.* B **228** 139–43
[51] Arnold P B 1989 Can the electroweak vacuum be unstable? *Phys. Rev.* D **40** 613
[52] Arnold P and Vokos S 1991 Instability of hot electroweak theory: bounds on $m(H)$ and $M(t)$ *Phys. Rev.* D **44** 3620–7
[53] Aad G *et al* (ATLAS Collaboration) 2012 Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC *Phys. Lett.* B **716** 1–29
[54] Chatrchyan S *et al* (CMS Collaboration) 2012 Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC *Phys. Lett.* B **716** 30–61
[55] Sher M 1993 Precise vacuum stability bound in the standard model *Phys. Lett.* B **317** 159–63
Sher M 1994 *Phys. Lett.* B **331** 448 (addendum)
[56] Altarelli G and Isidori G 1994 Lower limit on the Higgs mass in the standard model: an update *Phys. Lett.* B **337** 141–4
[57] Casas J A, Espinosa J R and Quirós M 1995 Improved Higgs mass stability bound in the standard model and implications for supersymmetry *Phys. Lett.* B **342** 171–9
[58] Espinosa J R and Quirós M 1995 Improved metastability bounds on the standard model Higgs mass *Phys. Lett.* B **353** 257–66
[59] Espinosa J R, Giudice G F and Riotto A 2008 Cosmological implications of the Higgs mass measurement *J. Cosmol. Astropart. Phys.* JCAP05(2008)002
[60] Ellis J, Espinosa J R, Giudice G F, Hoecker A and Riotto A 2009 The probable fate of the standard model *Phys. Lett.* B **679** 369–75
[61] Bezrukov F, Kalmykov M Y, Kniehl B A and Shaposhnikov M 2012 Higgs boson mass and new physics *J. High Energy Phys.* JHEP10(2012)140
[62] Degrassi G, Di Vita S, Elias-Miro J, Espinosa J R, Giudice G F, Isidori G and Strumia A 2012 Higgs mass and vacuum stability in the standard model at NNLO *J. High Energy Phys.* JHEP08(2012)098
[63] Masina I 2013 Higgs boson and top quark masses as tests of electroweak vacuum stability *Phys. Rev.* D **87** 053001
[64] Bednyakov A V, Kniehl B A, Pikelner A F and Veretin O L 2015 Stability of the electroweak vacuum: gauge independence and advanced precision *Phys. Rev. Lett.* **115** 201802
[65] Iacobellis G and Masina I 2016 Stationary configurations of the standard model Higgs potential: electroweak stability and rising inflection point *Phys. Rev.* D **94** 073005
[66] Andreassen A, Frost W and Schwartz M D 2018 Scale invariant instantons and the complete lifetime of the standard model *Phys. Rev.* D **97** 056006
[67] Chigusa S, Moroi T and Shoji Y 2017 State-of-the-art calculation of the decay rate of electroweak vacuum in the standard model *Phys. Rev. Lett.* **119** 211801
[68] Espinosa J R 2020 Vacuum decay in the standard model: analytical results with running and gravity *J. Cosmol. Astropart. Phys.* JCAP06(2020)052
[69] Mihaila L N, Salomon J and Steinhauser M 2012 Gauge coupling beta functions in the standard model to three loops *Phys. Rev. Lett.* **108** 151602

[70] Chetyrkin K G and Zoller M F 2013 β-function for the Higgs self-interaction in the standard model at three-loop level *J. High Energy Phys.* JHEP04(2013)091
Chetyrkin K G and Zoller M F 2013 *J. High Energy Phys.* JHEP09(2013)155
[71] Kniehl B A, Pikelner A F and Veretin O L 2015 Two-loop electroweak threshold corrections in the standard model *Nucl. Phys.* B **896** 19–51
[72] Martin S P 2014 Three-loop standard model effective potential at leading order in strong and top Yukawa couplings *Phys. Rev.* D **89** 013003
[73] Martin S P 2015 Four-loop standard model effective potential at leading order in QCD *Phys. Rev.* D **92** 054029
[74] Martin S P and Patel H H 2018 Two-loop effective potential for generalized gauge fixing *Phys. Rev.* D **98** 076008
[75] Cline J M and Espinosa J R 2018 Axionic landscape for Higgs coupling near-criticality *Phys. Rev.* D **97** 035025
[76] Branchina V, Messina E and Platania A 2014 Top mass determination, Higgs inflation, and vacuum stability *J. High Energy Phys.* JHEP09(2014)182
[77] Branchina V, Messina E and Sher M 2015 Lifetime of the electroweak vacuum and sensitivity to Planck scale physics *Phys. Rev.* D **91** 013003
[78] Branchina V and Messina E 2017 Stability and UV completion of the standard model *Europhys. Lett.* **117** 61002
[79] Andreassen A, Farhi D, Frost W and Schwartz M D 2017 Precision decay rate calculations in quantum field theory *Phys. Rev.* D **95** 085011
[80] Patel H H and Radovčić B 2017 On the decoupling theorem for vacuum metastability *Phys. Lett.* B **773** 527–33
[81] Branchina V, Contino F and Pilaftsis A 2018 Protecting the stability of the electroweak vacuum from Planck-scale gravitational effects *Phys. Rev.* D **98** 075001
[82] Ghosh J K, Kiritsis E, Nitti F and Witkowski L T 2021 Revisiting Coleman–de Luccia transitions in the AdS regime using holography *J. High Energy Phys.* JHEP09(2021)065
[83] Espinosa J R, Fortin J F and Huertas J 2021 Exactly solvable vacuum decays with gravity *Phys. Rev.* D **104** 065007
[84] Rajantie A and Stopyra S 2017 Standard model vacuum decay with gravity *Phys. Rev.* D **95** 025008
[85] Salvio A, Strumia A, Tetradis N and Urbano A 2016 On gravitational and thermal corrections to vacuum decay *J. High Energy Phys.* JHEP09(2016)054
[86] Branchina V, Messina E and Zappalà D 2016 Impact of gravity on vacuum stability *Europhys. Lett.* **116** 21001
[87] Bentivegna E, Branchina V, Contino F and Zappalà D 2017 Impact of new physics on the EW vacuum stability in a curved spacetime background *J. High Energy Phys.* JHEP12(2017)100
[88] Franceschini R, Strumia A and Wulzer A 2022 The collider landscape: which collider for establishing the SM instability? (arXiv:2203.17197 [hep-ph])
[89] Espinosa J R, Giudice G F, Morgante E, Riotto A, Senatore L, Strumia A and Tetradis N 2015 The cosmological Higgstory of the vacuum instability *J. High Energy Phys.* JHEP09(2015)174
[90] Delle Rose L, Marzo C and Urbano A 2016 On the fate of the standard model at finite temperature *J. High Energy Phys.* JHEP05(2016)050
[91] Griffiths D and Schroeter D 2018 *Introduction to Quantum Mechanics* (Cambridge: Cambridge University Press)
[92] Branchina V, Bentivegna E, Contino F and Zappalà D 2019 Direct Higgs–gravity interaction and stability of our Universe *Phys. Rev.* D **99** 096029