12

Vacuum decay

Consider a scalar field theory with a Lagrangian density of the form

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - V(\phi)\,, \tag{12.1}$$

where the scalar field potential has two unequal minima, as in Fig. 12.1. The lower of the two, at $\phi = \phi_{\rm tv}$, corresponds to the vacuum state of the theory and is called the true vacuum. The higher minimum, at $\phi = \phi_{\rm fv}$, is called the false vacuum. The values of the potential at these minima will be denoted $V_{\rm tv}$ and $V_{\rm fv}$, respectively. The false vacuum would be stable classically, but quantum mechanically it is only metastable and can decay via tunneling processes that we will study in this chapter.

These processes can be of importance in a cosmological context. For example, if the universe underwent any first-order phase transitions, below the critical temperature the metastable high-temperature phase would play the role of the false vacuum. If there is a string theory landscape, vacuum tunneling plays a crucial role in the evolution of the multiverse. Finally, we cannot exclude the possibility that our universe today is not in its ground state, but instead only in a very long-lived metastable state.

Although thermal and gravitational effects can often be significant, I will start with the case of vacuum decay at zero temperature in flat spacetime, and then later consider the effects of nonzero temperature and of gravity.

12.1 Bounces in a scalar field theory

Although Fig. 12.1 looks very much like Fig. 9.2, there is a crucial difference. The latter figure is a plot of potential energy, and the barrier is a finite one. The former is a plot of $V(\phi)$, which is an energy density. The actual potential energy is

$$U[\phi(\mathbf{x})] = \int d^3x \, \left[\frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi)\right] . \tag{12.2}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

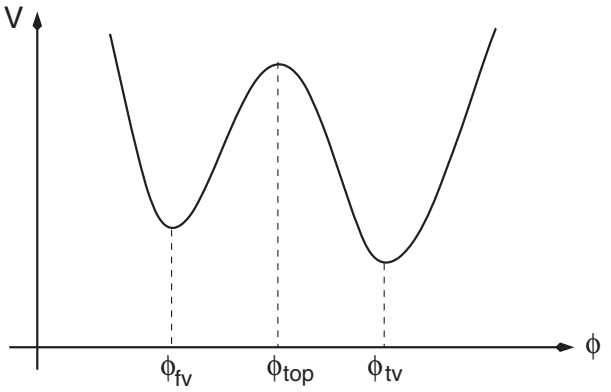


Fig. 12.1. A typical potential with a false vacuum.

To go from the false vacuum through a series of spatially homogeneous configurations would require traversing an infinite potential energy barrier. The tunneling amplitude for this vanishes. Instead, the false vacuum decays by a tunneling process that takes a spatially homogeneous state to one with a region of approximate true vacuum—a bubble—embedded in a false vacuum background. Because the bubble can be nucleated anywhere, the decay rate is proportional to the volume of space, and thus formally infinite. The finite physically measurable quantity that we need is the bubble nucleation rate per unit volume, $\Gamma/\mathcal{V}$.

One can envision many paths through the space of field configurations that connect the pure false vacuum to a configuration with a bubble. Two of these are illustrated in Fig. 12.2. Each path specifies a series of field configurations that define a slice through the potential energy barrier. A plot of $U[\phi(\mathbf{x})]$ along the path would be similar to the one-dimensional potential energy barrier shown in Fig. 9.3. The end point of the path, corresponding to the field configuration at the time that the bubble nucleates, has the same potential energy as the initial, pure false vacuum, configuration; quantum tunneling conserves energy.

As described in Chap. 9, the tunneling amplitude is dominated by the path that minimizes the barrier penetration integral B. This path can be found by finding the bounce solution to the Euclidean equation of motion [226], which in the present case is the field equation

$$0 = \frac{d^2\phi}{d\tau^2} + \boldsymbol{\nabla}^2\phi - \frac{dV}{d\phi} \tag{12.3}$$

that follows from the Euclidean action[1]

[1] Because almost all actions in this chapter will be Euclidean, I will generally omit an explicit subscript E on the action.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

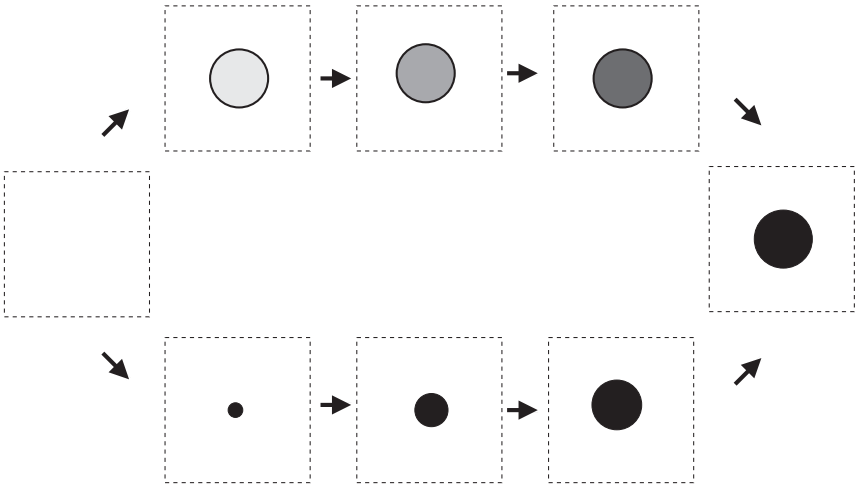


Fig. 12.2. Two of the infinite number of paths through the potential energy barrier that connect the pure false vacuum, on the left, with a configuration of the same energy, on the right, that contains a true vacuum bubble in a false vacuum background. In the upper path the bubbles are all of the same size, with the field in the interior progressing through the barrier in $V(\phi)$. In the lower path the bubble interior is always in the true vacuum, while the bubble radius increases from zero to that of the nucleated bubble.

$$S = \int d\tau\, d^3x \left[\frac{1}{2} \left(\frac{\partial \phi}{\partial \tau} \right)^2 + \frac{1}{2} (\boldsymbol{\nabla} \phi)^2 + V(\phi) \right] . \tag{12.4}$$

I will denote the bounce solution by $\phi_b(\mathbf{x}, t)$. The nucleation rate is then of the form

$$\frac{\Gamma}{\mathcal{V}} = A e^{-B} \tag{12.5}$$

where

$$B = S(\phi_b) - S(\phi_{\rm fv}) \, . \tag{12.6}$$

Here $S(\phi_b)$ is the Euclidean action of the bounce solution and

$$S(\phi_{\rm fv}) = \int d\tau\, d^3x\, V_{\rm fv} \tag{12.7}$$

that of the homogeneous false vacuum. Although both are infinite, their difference is finite.

Solving Eq. (12.3) is equivalent to finding a static solution in four spatial dimensions. It might appear that this is forbidden by Derrick's theorem. However, the proof of Derrick's theorem assumes that ϕ approaches the absolute minimum of V at spatial infinity, which is not the case for the bounce. For a theory in D spatial dimensions we can define

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

$$I_K = \frac{1}{2}\int d\tau\, d^D x\, (\partial_a \phi)^2 \tag{12.8}$$

and

$$I_V = \int d\tau\, d^D x\, [V(\phi) - V_{\rm fv}]\,. \tag{12.9}$$

The manipulations that led to Derrick's theorem then give

$$I_V = \frac{1-D}{1+D}\, I_K\,. \tag{12.10}$$

This tells us that I_V must be negative if $D > 1$. In contrast with the case of the soliton, this is not a problem because $V(\phi)$ is less than $V_{\rm fv}$ near the center of the bounce. Similarly, we can have $I_V = 0$, as required for $D = 1$, and yet have a nontrivial scalar field configuration. It follows from this virial identity that

$$B = S(\phi_b) - S(\phi_{\rm fv}) = I_K + I_V = \left(\frac{2}{1+D}\right) I_K\,. \tag{12.11}$$

The boundary conditions for Eq. (12.3) follow from the fact that the bounce represents a path through configuration space that begins with the initial configuration (the homogeneous false vacuum) at $\tau_{\rm init}$, reaches a turning point on the opposite side of the potential energy barrier (a configuration with a bubble) at $\bar\tau$ and then returns[2] in a τ-reversed fashion to the initial configuration at $\tau_{\rm fin}$. This is illustrated by Fig. 12.3. Note that $\partial\phi/\partial\tau = 0$ everywhere on the hypersurfaces $\tau = \tau_{\rm init}$ and $\tau = \bar\tau$ because these are turning points at which the kinetic energy vanishes.

Examination of the field equations reveals that $\tau_{\rm init}$ and $\tau_{\rm fin}$ must be $-\infty$ and ∞, respectively. We thus have

$$\phi(\mathbf{x}, \pm\infty) = \phi_{\rm fv}\,. \tag{12.12}$$

We must also require that the configurations along the tunneling path all have finite potential energy, measured relative to the false vacuum. This gives the requirement that

$$\phi(|\mathbf{x}| = \infty, \tau) = \phi_{\rm fv}\,. \tag{12.13}$$

Both the field equation and the boundary conditions display an O(4) invariance, despite the fact that τ and $\mathbf{x}$ have rather different physical meanings. This suggests setting $\bar\tau = 0$ and using an O(4)-invariant ansatz when seeking the bounce. Indeed, one might guess that the bounce with the lowest action (and therefore the dominant one) would have such a symmetry; this has been rigorously demonstrated, subject to mild assumptions about the form of $V(\phi)$, for the case of a theory with a single scalar field [271]. Thus, let us define $s = \sqrt{\tau^2 + \mathbf{x}^2}$

[2] Recall, from the discussion in Sec. 9.2, that the return part of the path just gives the factor of two needed to go from the tunneling amplitude to the tunneling probability.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

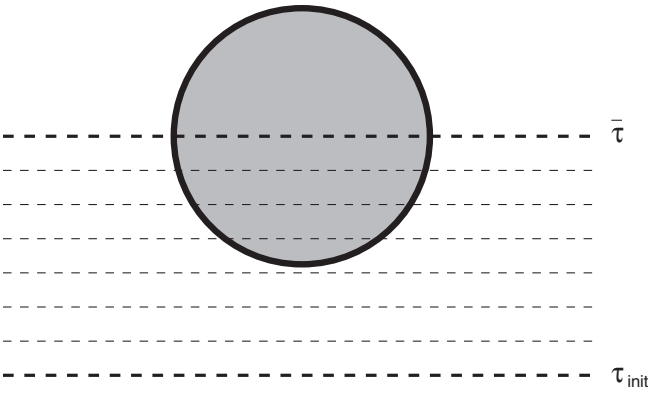


Fig. 12.3. Schematic view of a bounce solution. The field is approximately at its true vacuum value in the shaded region. The heavy circle surrounding this region is a wall region where the field crosses the barrier in $V(\phi)$. The dashed lines represent a sequence of spatial hypersurfaces of constant τ that describe a path through the configuration space. The lower heavy dashed line corresponds to the initial configuration, at $\tau_{\rm init}$, and the upper one, through the center of the true vacuum region, the configuration after emergence from the potential energy barrier. Although not indicated in the figure, there is an exponentially small deviation from the false vacuum that slowly increases as the configurations progress from $\tau_{\rm init}$ to the first appearance of the wall region. The upper half of the bounce is a mirror image of the lower half.

and seek a bounce solution that depends only on s. The Euclidean field equation then becomes

$$\phi'' + \frac{3}{s}\phi' = \frac{dV}{d\phi}\,, \tag{12.14}$$

where primes denote differentiation with respect to s. The boundary conditions on the bounce reduce to

$$\phi(\infty) = \phi_{\rm fv}\,, \tag{12.15}$$

while requiring that the solution be nonsingular at the origin gives

$$\phi'(0) = 0\,. \tag{12.16}$$

The existence of a solution to these equations can be demonstrated by an "overshoot–undershoot" argument due to Coleman [226]. The idea, illustrated in Fig. 12.4, is to reinterpret Eq. (12.14) as the equation for a particle with position ϕ moving in an "upside-down" potential energy $-V(\phi)$, with s denoting the time. The ϕ' term is interpreted as a damping force that decreases with time. The boundary conditions require that the particle start at rest at some initial point ϕ_0 on the true vacuum side of the potential well, and that it come to rest at $\phi_{\rm fv}$ as $s \to \infty$. What we need to do is to show that a suitable ϕ_0 exists.

First, suppose that we take ϕ_0 to be a point on the true vacuum side such that $-V(\phi_0) \le -V_{\rm fv}$. The particle's energy decreases with time, because of the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

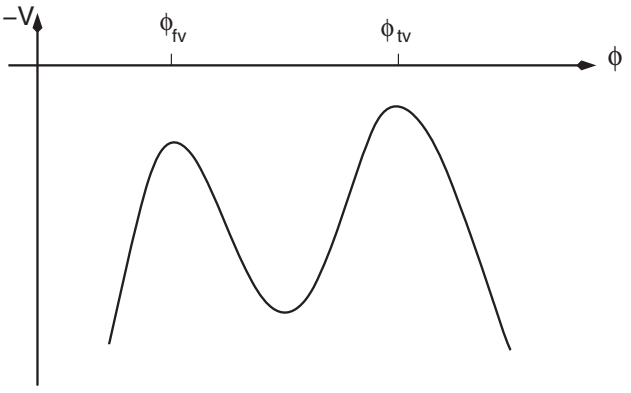


Fig. 12.4. The upside-down potential used in the overshoot–undershoot argument. The field begins, at $s = 0$, near the true vacuum peak and must finally come to rest at $\phi_{\rm fv}$.

damping term, and so the particle will never have sufficient energy to reach $\phi_{\rm fv}$; i.e., it undershoots. Now suppose instead that ϕ_0 is taken to differ only infinitesimally from $\phi_{\rm tv}$. It will then start to move very slowly, so that by the time it has moved appreciably from its initial position the damping force will have almost died away and the particle's energy will be essentially conserved. It will then have nonzero kinetic energy when it reaches $\phi_{\rm fv}$ and so will continue on, never to return; it overshoots. By continuity, there must be a range of ϕ_0 near the top of the potential that lead to overshoots and another range, lower down, giving undershoots. The boundary between these two ranges gives the desired ϕ_0 that determines the bounce. (While this argument demonstrates the existence of a bounce, it does not prove the uniqueness of the bounce. It is not hard to find potentials that have several ranges of overshoot and undershoot, and thus several distinct bounce solutions.)

Assuming O(4) symmetry, the tunneling exponent from the bounce can be written as

$$B = 2\pi^2 \int_0^\infty ds\, s^3 \left[\frac{1}{2}(\phi')^2 + V(\phi) - V_{\rm fv} \right] , \tag{12.17}$$

where the factor of $2\pi^2$ is from the four-dimensional angular integration. We can estimate the magnitude of B by scaling arguments. Let us assume that the theory has a single characteristic mass scale m and that the scalar field potential can be written as

$$V(\phi) = g^2 v^4 \tilde{V}(\phi/v) + V_{\rm fv} , \tag{12.18}$$

where $v = m/g$ and the dimensionless parameters in $\tilde{V}$ are all of order unity. (In particular, this implies that $|\phi_{\rm tv} - \phi_{\rm fv}|$ is of order v.) We can then define rescaled variables $u = ms$ and $f = \phi/v$ and write

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

$$B = \frac{2\pi^2}{g^2} \int_0^\infty du\, u^3 \left[\frac{1}{2} \left(\frac{df}{du} \right)^2 + \tilde{V}(f) \right] . \tag{12.19}$$

For generic potentials the characteristic length scale of the bounce is of order m^{-1} (an exception will be discussed in the next section), so the integral is dominated by the region with u of order unity. Because the integral contains no small parameters, we expect it to be a number C of order unity. Hence,

$$B \approx \frac{2\pi^2 C}{g^2} . \tag{12.20}$$

In order for our semiclassical approximation, with its neglect of higher-order corrections, to be reliable the couplings must be small. It immediately follows that $e^{-B} \ll 1$, which guarantees the validity of the dilute-gas approximation that was used in the derivations of Chap. 9

As a concrete example, let us consider tunneling with a quartic field potential $V(\phi)$ that has two unequal minima. If an overall additive constant is chosen so that V vanishes at the false vacuum, then the potential depends on four constants. Two of these can be fixed by rescaling and shifting the field so that the false and true vacua are at $\phi = -a$ and $\phi = a$, respectively. An overall dimensionless constant can be factored out, as described above. The actual shape of the potential is determined by the remaining constant. If we take this to be a quantity b lying in the range $0 < b < 1$, and write

$$V(\phi) = \frac{g^2}{4} \left[(\phi^2 - a^2)^2 + \frac{4b}{3} (a\phi^3 - 3a^3\phi - 2a^4) \right] , \tag{12.21}$$

then the top of the barrier separating the false and true vacua is at $\phi = -ba$. The shape of this potential for several values of b is shown in Fig. 12.5.

Numerically solving the bounce equation for several values of b gives the solutions shown in Fig. 12.6. Notice that for $b \gtrsim 0.6$ the field at the center of the bounce, $s = 0$, is visibly different from its true vacuum value. The actions for these bounces are shown in Fig. 12.7, where $g^2 B$ is plotted as a function of b. It can be seen that the quantity C in Eq. (12.20), although formally of order unity, can be quite large.

An expression for the pre-exponential factor in the decay rate for a single degree of freedom was obtained in Sec. 9.4. This is easily generalized to the field theory case [230]. Because the bounce solution is not spatially homogeneous, there are three zero modes, corresponding to spatial translation of the bounce, in addition to the zero mode corresponding to translation in Euclidean time.[3]

[3] Further zero modes can appear if the theory has an internal symmetry that is broken by the bounce [272]. However, as discussed in Sec. 9.5, there must be only a single negative mode because a solution with multiple negative modes gives only a saddle point, rather than a minimum, of the tunneling exponent B.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

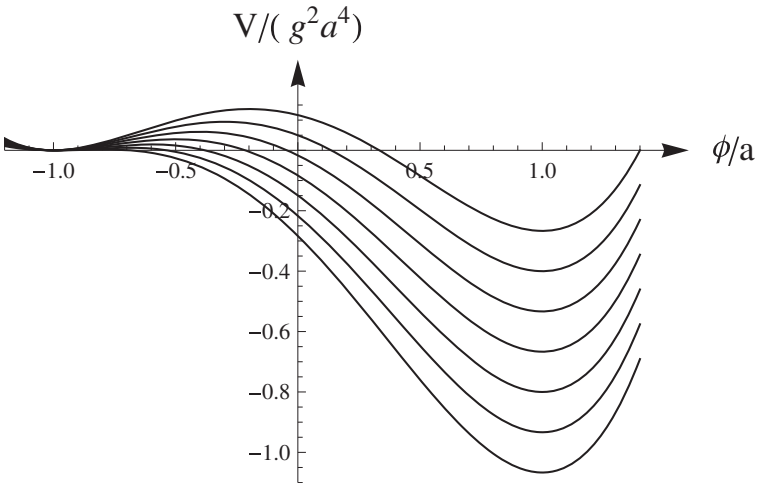


Fig. 12.5. The potential of Eq. (12.21) for various values of b. From top to bottom, the curves correspond to $b = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$.

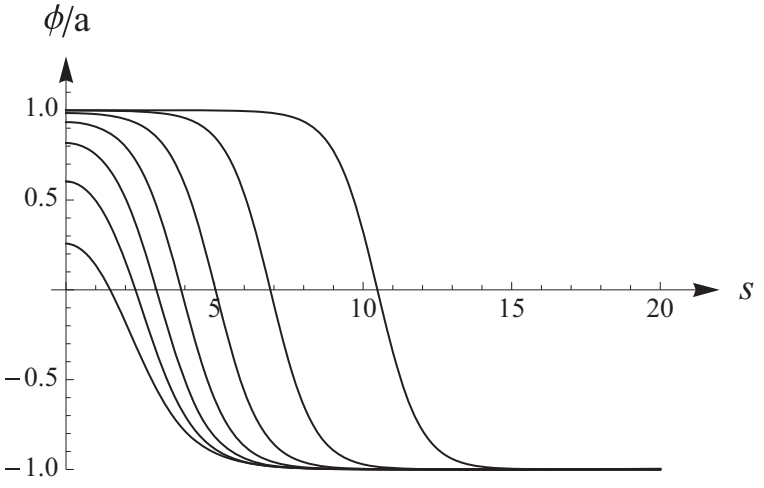


Fig. 12.6. The bounce solutions for the potential of Eq. (12.21). Reading from left to right, these correspond to $b = 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2$.

These new zero modes require the introduction of three collective coordinates, which can be chosen to be the location of the center of the bounce.

A second factor to be taken into account is that in a field theory the determinant factors have divergences that remain even after their ratio is taken. These are completely analogous to the divergence that we encountered in the course of calculating the first quantum corrections to the energy of a soliton in Sec. 2.2.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

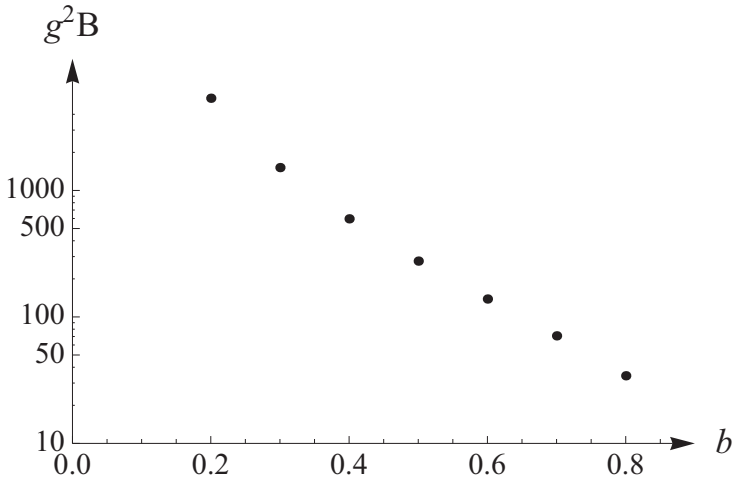


Fig. 12.7. The quantity g^2B as a function of b for the potential of Eq. (12.21).

The remedy is similar—we simply add the leading counterterm corrections to the calculation of our Euclidean actions.

Making the appropriate modifications to Eq. (9.57), we obtain

$$\Gamma = \mathcal{V}\left[\prod_{a=1}^{4}\left(\frac{N_a}{2\pi}\right)^{1/2}\right]\left|\frac{\det{}' S''(\phi_b)}{\det S''(\phi_{\rm fv})}\right|^{-1/2} e^{-B} e^{-[S_{\rm ct}(\phi_b)-S_{\rm ct}(\phi_{\rm fv})]}\,. \tag{12.22}$$

Here N_a is the normalization factor for the zero mode corresponding to translation in the x_a direction (with $\tau \equiv x_4$). The factor of $\mathcal{V}$, the volume of space, comes from integrating over the location of the bounce and must be divided out to give the physically measurable bubble nucleation rate per unit volume.[4]

For an O(4)-symmetric bounce the N_a are all equal, and are given by

$$N_a = \int d^4x (\partial_a \phi)^2 = \frac{1}{2} I_K = B\,, \tag{12.23}$$

so

$$\frac{\Gamma}{\mathcal{V}} = \left(\frac{B}{2\pi}\right)^2 \left|\frac{\det{}' S''(\phi_b)}{\det S''(\phi_{\rm fv})}\right|^{-1/2} e^{-[S_{\rm ct}(\phi_b)-S_{\rm ct}(\phi_{\rm fv})]} e^{-B}\,. \tag{12.24}$$

Evaluation of the determinants in the pre-exponential factor is a nontrivial matter. A rough estimate of their ratio can be obtained by dimensional arguments. Because of the omission of the zero modes from the determinant about

[4] The generalization of this to a theory with many fields is fairly straightforward. However, extra care is required in situations where the vacuum structure is not evident at tree level, but only after one-loop corrections to the effective potential are taken into account [125]. In such cases one must be careful to avoid double-counting of one-loop terms in both the effective action and in the determinant factors. For a discussion of this, see Ref. [273].

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

the bounce, the ratio of determinant factors has the dimensions of (mass)^4, just as needed to give the correct dimensions for $\Gamma/\mathcal{V}$. For a generic potential the only relevant mass is m. A rescaling such as that described above Eq. (12.19) shows that the determinants give no additional powers of g. Hence, suppressing numerical factors in the prefactor, we have

$$\frac{\Gamma}{\mathcal{V}} \approx \left(\frac{m}{g}\right)^4 e^{-B} \approx v^4 e^{-B}\,, \tag{12.25}$$

with the factor of $1/g^4$ coming from the B^2 in the prefactor.

12.2 The thin-wall approximation

An interesting and instructive case to consider is the one where

$$\epsilon \equiv V_{\rm fv} - V_{\rm tv} \tag{12.26}$$

is small compared to the height of the barrier in $V(\phi)$ [226]. If we analyze this in terms of the analogous problem of a particle moving in the upside-down potential, we see that the particle, starting near $\phi_{\rm tv}$, can only lose a very small amount of energy to damping forces if it is to avoid an undershoot. This can only happen if the particle stays very close to $\phi_{\rm tv}$ until some large time $s \approx R$ when the damping force has become negligible. It can then move quickly down the valley and back up the other side to come to rest at $\phi_{\rm fv}$.

A potential of this form can be constructed by starting with a potential $V_1(\phi)$ that has two degenerate minima, at $\phi = \phi_{\rm fv}$ and $\phi = \phi_{\rm tv}$, and then defining

$$V(\phi) = V_1(\phi) + \epsilon\left(\frac{\phi_{\rm tv} - \phi}{\phi_{\rm tv} - \phi_{\rm fv}}\right). \tag{12.27}$$

(Note that the actual minima of V deviate from $\phi_{\rm fv}$ and $\phi_{\rm tv}$ by amounts of order ϵ, which I will ignore.) To a first approximation, near $s = R$ we can neglect both the $3\phi'/s$ term in the field equation and the ϵ-dependent term in $V(\phi)$. We then have

$$\phi'' = \frac{dV_1}{d\phi}\,. \tag{12.28}$$

This is just the equation for a one-dimensional soliton. Applying the methods of Sec. 2.1 [see Eq. (2.21)] gives a solution $\phi_1(x)$ defined by

$$x = \int_{\phi_{\rm top}}^{\phi_1} \frac{d\phi}{\sqrt{2[V_1(\phi) - V_1(\phi_{\rm fv})]}}\,. \tag{12.29}$$

Away from $x = 0$, ϕ_1 approaches the vacuum values exponentially fast. Ignoring terms of order ϵ, we find that the soliton's one-dimensional energy is

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

$$\sigma = \int dx \left\{ \frac{1}{2}(\phi_1')^2 + [V(\phi) - V_{\rm fv}] \right\}$$
$$= \int_{\phi_{\rm fv}}^{\phi_{\rm tv}} d\phi \sqrt{2[V(\phi) - V_{\rm fv}]}\,. \tag{12.30}$$

Our approximate solution for the bounce is then

$$\phi(s) = \begin{cases} \phi_{\rm tv}\,, & s \ll R\,, \\ \phi_1(s-R)\,, & s \approx R\,, \\ \phi_{\rm fv}\,, & s \gg R, \end{cases} \tag{12.31}$$

if $\phi_{\rm tv} > \phi_{\rm fv}$; if instead $\phi_{\rm tv} < \phi_{\rm fv}$, then the field in the wall region is $\phi_1(R-s)$. This is essentially a four-dimensional bubble of true vacuum surrounded by a thin wall. Its Euclidean action, relative to that of the false vacuum, is

$$S[R] = 2\pi^2 R^3 \sigma - \frac{1}{2}\pi^2 R^4 \epsilon\,. \tag{12.32}$$

The first term is the contribution from the wall and shows that σ can be viewed as a surface tension, while the second is from the true vacuum interior.

The bounce radius R is determined by requiring that this be a stationary point of the Euclidean action, which in turn requires that

$$0 = \frac{dS}{dR} = 6\pi^2 R^2 \sigma - 2\pi^2 R^3 \epsilon\,. \tag{12.33}$$

This is solved by

$$R_b = \frac{3\sigma}{\epsilon}\,, \tag{12.34}$$

which gives

$$B = S = \frac{27\pi^2 \sigma^4}{2\epsilon^3}\,. \tag{12.35}$$

Differentiating Eq. (12.33) yields

$$\left.\frac{d^2 S}{dR^2}\right|_{R_b} = -\frac{18\pi^2\sigma^2}{\epsilon} < 0\,. \tag{12.36}$$

This shows that R_b maximizes the action among the family of configurations given by Eq. (12.31), and that the expected negative mode about the bounce corresponds to variation of R. More explicitly, differentiating Eq. (12.14) with respect to s gives

$$\phi''' + \frac{3}{s}\phi'' - \frac{3}{s^2}\phi' = \frac{d^2 V}{d\phi^2}\,\phi' \tag{12.37}$$

or, equivalently,

$$S''(\phi_b)\phi' = \left[-\frac{d^2}{ds^2} - \frac{3}{s}\frac{d}{ds} + \frac{d^2 V}{d\phi^2} \right] \phi'$$
$$= -\frac{3}{s^2}\,\phi'. \tag{12.38}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

In the thin-wall approximation, where ϕ' is non-negligible only for $s \approx R$, we can replace s by R in the last line. Hence, in this approximation ϕ' is an eigenfunction of S'' with a negative eigenvalue $-3/R^2$.

Despite the attractive simplicity of the formulas obtained by the thin-wall approximation, one must keep in mind that they are not generally applicable. In the regime of nearly degenerate minima where the approximation is valid the bounce action is generally quite large, as can be seen, for example, from the data shown in Fig. 12.7.

12.3 Evolution of the bubble after nucleation

The bounce plays a dual role. Not only does its action determine the bubble nucleation rate, but a spatial slice through its center gives the most probable emergence point from the potential energy barrier; i.e., the most likely initial configuration for the classical evolution of the bubble after nucleation. In fact, the bounce solution also gives us much of the future development of ϕ directly, without the need for further calculation.

Let us denote the Euclidean space bounce solution by $\phi_E(\mathbf{x}, \tau)$ and the Minkowski spacetime field by $\phi_M(\mathbf{x}, t)$. These satisfy

$$0 = \frac{d^2\phi_E}{d\tau^2} + \boldsymbol{\nabla}^2\phi_E - \frac{dV}{d\phi} \tag{12.39}$$

and

$$0 = -\frac{d^2\phi_M}{dt^2} + \boldsymbol{\nabla}^2\phi_M - \frac{dV}{d\phi}\,, \tag{12.40}$$

as well as the initial value conditions

$$\begin{aligned} 0 &= \left.\frac{d\phi_E(\mathbf{x},\tau)}{d\tau}\right|_{\tau=0}\,, \\ 0 &= \left.\frac{d\phi_M(\mathbf{x},\tau)}{dt}\right|_{t=0}\,. \end{aligned} \tag{12.41}$$

Examining these, we see that starting with a Euclidean solution and making the replacement $\tau \to it$ gives a solution of the Minkowskian equations. If the bounce has the O(4)-symmetric form $\phi_E = f(\sqrt{\mathbf{x}^2 + \tau^2})$, then the Minkowskian solution is given by the O(3,1) symmetric $\phi_M = f(\sqrt{\mathbf{x}^2 - t^2})$ with the same function f. In the region outside the light cone of the origin, $r = \sqrt{\mathbf{x}^2} > t$, this gives ϕ_M directly. The region $r < t$ inside the light cone has no Euclidean counterpart. In principle, ϕ_M could be obtained here by analytic continuation, but in practice one writes $\phi_M = \phi_M(u)$, with $u = \sqrt{t^2 - r^2}$, and solves

$$\phi_M''(u) + \frac{3}{u}\phi_M' = -\frac{dV}{d\phi}\,. \tag{12.42}$$

In the thin-wall approximation the bubble at nucleation has a thin wall of radius $R_0 = 3\sigma/\epsilon$ that separates an exactly true vacuum interior from an exactly

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

false vacuum exterior.[5] Because the Minkowskian solution depends only on r^2-t^2, the bubble wall at future times must be at

$$R(t) = \sqrt{R_0^2 + t^2}\,, \tag{12.43}$$

which means that it is expanding at a velocity

$$v = \frac{dR}{dt} = \frac{t}{R} = \frac{\sqrt{R^2 - R_0^2}}{R} \tag{12.44}$$

that approaches the speed of light.

The energy density in this expanding bubble wall can be obtained by Lorentz-boosting the surface tension of the wall at rest, giving

$$E_{\rm wall} = 4\pi R^2 \frac{\sigma}{\sqrt{1-v^2}} = 4\pi\sigma \frac{R^3}{R_0} = \frac{4\pi}{3}\, R^3 \epsilon\,. \tag{12.45}$$

This is precisely equal to the latent heat released by the conversion of false vacuum to true. All of the released energy has gone into the kinetic energy of the walls, with none of it exciting the field in the bubble interior.

This is somewhat surprising, and certainly contrasts with our experience with nonrelativistic bubbles expanding in a medium. One might wonder whether this result is an artifact of the thin-wall approximation. To see that it is not, let us consider the solutions of Eq. (12.42). At $u = 0$ we have $\phi_M(0) = \phi_E(0)$, the value at the center of the bounce. As u increases, ϕ_M approaches $\phi_{\rm tv}$ and then oscillates about the true vacuum with an amplitude that falls as $u^{3/2}$. Let us define u_i as the value where ϕ_M approximates the true vacuum to some predetermined accuracy, thus defining $r_i = \sqrt{t^2 - u_i^2}$ to be the inner surface of the bubble wall. Similarly, we can define a value s_o where the field approximates the false vacuum, so that $r_o = \sqrt{s_o^2 + t^2}$ is the location of the outer surface of the bubble wall. The thickness of the bubble wall at large t is then

$$r_o - r_i \approx \frac{s_o^2 + u_i^2}{2t}\,. \tag{12.46}$$

Thus, even if the bubble does not initially have a thin wall, it develops one as time goes on. Although the interior is not precisely a true vacuum, as it is in the thin-wall approximation, the energy density from oscillations about the true vacuum decreases, eventually falling as t^{-3}. Since the walls are expanding outward at almost the speed of light, the total energy in the interior of the bubble remains essentially constant.

From an exterior point of view, it seems natural to use the coordinates r and t to describe the bubble, and to view it as a finite, but expanding, region

[5] The energy in the walls of this bubble is exactly canceled by the negative volume energy from the replacement of false vacuum by true in the bubble interior. Thus, energy is conserved, as it should be.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

of approximate true vacuum. However, this is not the most natural description from the point of view of an interior observer. Inside the light cone of the origin we can define

$$\tilde{t} = \sqrt{t^2 - r^2}\,,$$
$$\tilde{r} = \frac{r}{\sqrt{t^2 - r^2}}\,. \tag{12.47}$$

Because ϕ_M depends only on $\tilde{t}$, these coordinates give a description in which the bubble interior (or at least the part within the light cone of the origin) is a time-dependent but spatially homogeneous region of infinite spatial extent. There is no privileged point that can be uniquely defined to be the "center" of the bubble. In terms of the new coordinates, the Minkowskian metric takes the form

$$\begin{aligned} ds^2 &= d\tilde{t}^{\,2} - \tilde{t}^{\,2}\left[\frac{d\tilde{r}^2}{1+\tilde{r}^2} + \tilde{r}^2(d\theta^2 + \sin^2\theta\, d\varphi^2)\right] \\ &\equiv d\tilde{t}^{\,2} - \tilde{t}^{\,2}\, d\Omega_T^2\,. \end{aligned} \tag{12.48}$$

This describes an open Friedmann–Robertson–Walker universe with infinite negatively curved spacelike surfaces of constant time. These surfaces are spacelike hyperboloids (i.e., hyperboloids with timelike normals) and $d\Omega_T^2$ is the metric on the unit spacelike hyperboloid.

This picture of an almost empty bubble interior is only valid to the extent that collisions with other bubbles can be ignored. One might expect that the continued nucleation and growth of bubbles would eventually cause them to coalesce, so that the bubble walls would disappear and the latent heat would be released. This is indeed what happens in cosmological first-order phase transitions that proceed rapidly with little supercooling. However, as will be discussed in Sec. 12.7, bubble nucleation in a quasi-de Sitter background can lead to a very different type of scenario [126].

12.4 Tunneling at finite temperature

Let us now consider the effects of a nonzero temperature $T = 1/\beta$. Specifically, let us assume that our system is initially in a quasi-equilibrium, with a thermal distribution of the states built upon the false vacuum, but with states built upon the true vacuum being unpopulated.

The picture of bubble nucleation outlined in the previous sections must be modified in several ways [274, 275]. First, it is clear that in finding the bounce we must use the finite temperature effective potential $V_{\rm eff}(\phi, T)$, which was discussed in Sec. 7.2, instead of the zero-temperature scalar field potential. Otherwise, we would find a nonzero rate for nucleating bubbles of the low-temperature phase even above the critical temperature of a first-order transition. To simplify the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

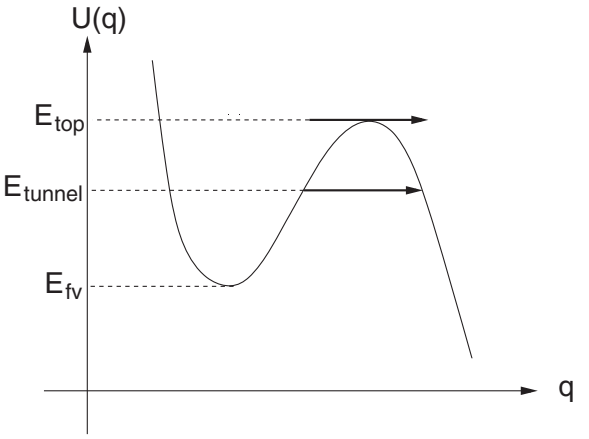


Fig. 12.8. The two modes for transitions at finite temperature. One is a fluctuation of the system to $E_{\rm top}$, allowing it to pass over the potential energy barrier, the other thermally assisted tunneling at an energy $E_{\rm tunnel}$.

notation, I will continue to simply write $V(\phi)$, but temperature dependence should be understood.

Second, there are new mechanisms available for the transition. The first of these might be termed thermally assisted quantum tunneling. Instead of tunneling directly from the false vacuum at the bottom of the potential energy well, the system can tunnel from a thermally excited higher energy state for which the barrier penetration integral is smaller. This is illustrated in Fig. 12.8 for a system with only a single degree of freedom. Taking a thermal average of the tunneling rate for this system and keeping track only of exponential factors gives

$$\begin{aligned}\Gamma_{\rm tunn} &\sim \int_{E_{\rm fv}}^{E_{\rm top}} dE\, e^{-\beta(E-E_{\rm fv})} e^{-B(E)} \\ &\sim e^{-\beta(E_*-E_{\rm fv})} e^{-B(E_*)}\,, \end{aligned} \tag{12.49}$$

where E_* is the value for which $\beta(E-E_{\rm fv})+B(E)$ is a minimum.

With many degrees of freedom and a Lagrangian with a quadratic kinetic energy, as in Eq. (9.7), we must minimize

$$\beta(E-E_{\rm fv}) + 2\int_{s_0}^{s_f} ds\,\sqrt{2[V(q(s))-E]} \tag{12.50}$$

with respect to path, end points, and energy. We know that the minimization with respect to the first two of these is achieved by finding a solution of the Euclidean equations of motion. Minimization with respect to energy requires that

$$\beta = -2\frac{d}{dE}\int_{s_0}^{s_f} ds\,\sqrt{2[V(q(s))-E]}\,. \tag{12.51}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

The integral has implicit dependences on E through the path and the end points. However, since we have already minimized with respect to path and end points, these implicit dependences do not contribute here, so we have

$$\begin{aligned}\beta &= 2\int_{s_0}^{s_f} ds\, \frac{1}{\sqrt{2[V(q(s))-E]}} \\ &= 2\int_{s_0}^{s_f} ds\, \left[\sum_j \left(\frac{dq_j}{d\tau}\right)^2\right]^{-1/2} \\ &= 2\int_{s_0}^{s_f} ds\, \frac{d\tau}{ds} \\ &= 2(\tau_2-\tau_1)\,. \end{aligned} \tag{12.52}$$

Thus, the passage through the barrier must take a Euclidean time $\tau_2 - \tau_1 = \beta/2$; adding the τ-reversed passage doubles the time and tells us that we must look for a bounce solution that is periodic[6] in τ with period β. Repeating the manipulations of the zero-temperature case then gives

$$\Gamma_{\rm tunn} \sim e^{-[S(\phi_b)-S(\phi_{\rm fv})]}\,, \tag{12.53}$$

where in both actions the τ integration is understood to be over a single period β.

The second new possibility is for the system to be thermally excited all the way to the top of the barrier, with no quantum tunneling required at all, as also illustrated in Fig. 12.8. Again ignoring pre-exponential factors, we have

$$\Gamma_{\rm therm} \sim e^{-\beta(E_{\rm top}-E_{\rm fv})}\,. \tag{12.54}$$

When there is more than one degree of freedom, what appears as the top of the barrier in Fig. 12.8 is actually a saddle point, with the potential energy increasing in all of the remaining directions; i.e., it is the highest point on the lowest pass over the potential energy barrier. Because it is a stationary point of the potential energy, it is a static solution of the field equations.

In the field theory context this solution is a three-dimensional critical bubble, one where the outward pressure from the approximately true vacuum interior is precisely balanced by the surface tension; the density of such configurations in the thermal ensemble determines the bubble nucleation rate. In the thin-wall approximation the energy, relative to the false vacuum background, of a bubble of radius R is

$$E = 4\pi R^2\sigma - \frac{4\pi}{3}R^3\epsilon\,. \tag{12.55}$$

[6] The fact that the bounce interpolates between a pair of classical turning points on opposite sides of the potential energy barrier means that, in addition to being periodic, the bounce must have two hypersurfaces of constant τ, separated by half a period, on which $\partial\phi/\partial\tau$ is identically zero.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

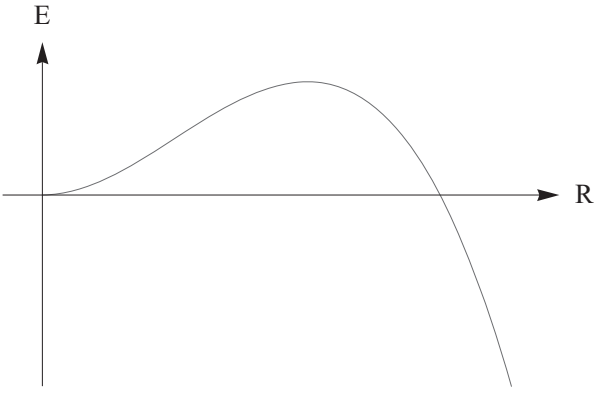


Fig. 12.9. Plot of energy as a function of bubble radius for thin-wall bubbles. The high-temperature critical bubble corresponds to the maximum of E, while the bubble created by quantum tunneling at $T = 0$ corresponds to the value of R where the curve crosses the horizontal axis.

As shown in Fig. 12.9, the critical bubble corresponds to the maximum of this function, while the bubble nucleated by purely quantum tunneling is one with $E = 0$.

We can view the three-dimensional saddle point configuration as a four-dimensional τ-independent bounce solution that extends for a period β. Equation (12.53) then applies equally well here, with the distinction between thermally assisted tunneling and a purely thermal transition being whether or not the bounce has a τ-dependence.

The problem can also be attacked by path integral methods. The partition function can be written as an integral over paths with periodic boundary conditions, which in the field theory case takes the form

$$Z = e^{-\beta F} = \sum_j e^{-\beta E_j} = \int [d\phi] e^{-S_E(\phi)} \,. \tag{12.56}$$

Because we are interested in a system that is a thermal distribution of false-vacuum-type states, the path integral is restricted to configurations on the false vacuum side of the potential barrier. The individual energies E_j and the free energy F then develop imaginary parts. Assuming these to be small, we have

$$\Gamma = -2 \operatorname{Im} F = -2Z^{-1} \sum_j e^{-\beta E_j} \operatorname{Im} E_j \,. \tag{12.57}$$

Thus, the imaginary part of the free energy gives us a thermally weighted average of the decay rate.

As before, the path integral can be evaluated as a sum of contributions from the stationary points and approximate stationary points of the action. When the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

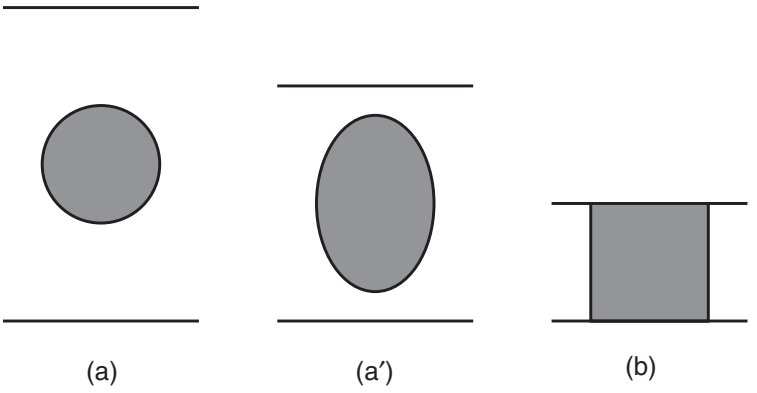


Fig. 12.10. Schematic illustration of the periodic bounces for finite temperature transitions. In all three cases the solid horizontal lines indicate the hypersurfaces $\tau = 0$ and $\tau = \beta$, the bounds of one period. Bounces (a) and (a′) describe thermally assisted tunneling and illustrate the deformation of the shaded true vacuum region with increasing temperature. Bounce (b) is τ-independent and describes thermal fluctuation over the potential energy barrier. Such bounces dominate at high temperature. At lower temperatures they remain solutions of the Euclidean field equation, but are unacceptable because they have multiple negative modes.

dilute-gas approximation is valid, these exponentiate and the imaginary part of the free energy can be read off from the coefficient of β in the exponent, leading to an expression for the pre-exponential factor analogous to that obtained in the zero-temperature case.

We have already seen that two distinct types of bounce can occur. At very low temperature (i.e., large β), the bounce corresponding to thermally assisted tunneling closely approximates the O(4)-symmetric zero-temperature bounce. As the temperature is increased this bounce retains an O(3) symmetry, but (except in the thin-wall approximation) is deformed in the τ-direction. Indeed, by considering such a bounce and its periodic recurrences, we see that the true vacuum region of the bounce is elongated, as shown in Fig. 12.10, eventually "colliding" with itself when the inverse temperature is comparable to the bounce radius.

The τ-independent bounce corresponding to a purely thermal transition is a solution of the Euclidean field equation at all temperatures.[7] However, at low temperatures it is not only subdominant, in that its action is greater than that of the bounce for thermally assisted tunneling, but it is also unacceptable because it has more than one negative mode. This can be seen by noting that the three-dimensional normal modes about the critical bubble solution $\phi_{\rm bubble}(\mathbf{x})$ include a negative mode $\eta_-(\mathbf{x})$ for which

[7] There is no similar guarantee for the τ-dependent bounce. The absence of such a solution would indicate that the expression in Eq. (12.50) had no minimum and was instead a monotonically decreasing function of E in the range $E_{\rm fv} < E < E_{\rm top}$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

$$\left[-\boldsymbol{\nabla}^2 + V''(\phi_{\rm bubble})\right]\eta_- = -\mu^2\eta_- \,. \tag{12.58}$$

The four-dimensional normal modes about the corresponding bounce are eigenfunctions of

$$S'' = \left[-\frac{\partial^2}{\partial\tau^2} - \boldsymbol{\nabla}^2 + V''(\phi_{\rm bubble})\right] \tag{12.59}$$

that are periodic in τ. In particular, these include modes of the form $\cos(2\pi n\tau/\beta)\eta_-(\mathbf{x})$, with eigenvalues $(2\pi n/\beta)^2 - \mu^2$. If $T = 1/\beta < \mu/2\pi$, there are multiple negative modes.

Although the existence of multiple negative modes necessarily implies that a solution is not the bounce with minimum action, the converse need not be true. For example, in the thin-wall approximation one can easily verify that there is a temperature range in which the τ-independent bounce has only a single negative mode but nevertheless has a higher action than the τ-dependent bounce.

As a final point, note that the expansion of the bubble after nucleation is more complicated at nonzero temperature. In particular, its expansion within the thermal medium never approaches the speed of light.

12.5 Including gravity: bounce solutions

The effects of gravity on bubble nucleation have been neglected so far. While this is justified in a wide range of applications, there are situations where gravitational effects are clearly significant. The most obvious is when a transition involves mass scales close to the Planck scale, but gravitational effects can also come into play at lower mass scales if the bubble nucleates with a size large enough to be sensitive to the curvature of spacetime. Even when gravitational effects are unimportant for bubble nucleation itself they can be quite significant for the evolution of the bubble after nucleation.

With gravity ignored, the WKB and path integral approaches led us in a straightforward manner from first principles to expressions for the nucleation rate. The extension of these arguments to include gravity is a nontrivial matter, in large part because of the difficulties associated with globally defining energy in a curved spacetime, and to date no complete derivation of an expression for the nucleation rate has been produced. Nevertheless, one can argue from analogy, as was first done by Coleman and De Luccia [276]. As we will see, this leads to an expression that agrees with the flat spacetime result[8] in the weak gravity limit. However, some issues of interpretation arise when gravitational effects are large. I will show in Sec. 12.6 how these can be clarified by considering a limiting case in which the nucleation rate can be derived [277].

[8] Here, and in similar contexts, "flat spacetime" refers to the case where gravitational effects are ignored and the spacetime is treated as a fixed flat spacetime in both the true and false vacua.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

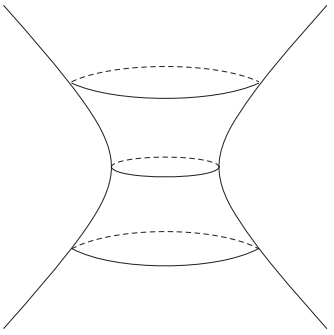


Fig. 12.11. De Sitter spacetime viewed as a hyperboloid. Time runs vertically. Horizontal spacelike slices through the hyperboloid are three-spheres.

It may be best to begin with a brief description of the initial false vacuum state, which requires some comment when gravity is included. The scalar field is of course spatially uniform, equal to $\phi_{\rm fv}$ everywhere. The nature of the spacetime depends on the value of the potential in the false vacuum. If $V_{\rm fv} = 0$, the false vacuum spacetime is simply flat Minkowski space, and requires little further discussion. If it is nonzero, then $V_{\rm fv}$ is equivalent to a nonzero cosmological constant, and the false vacuum is either a de Sitter or an anti-de Sitter spacetime, according to whether $V_{\rm fv}$ is positive or negative, respectively.

As illustrated in Fig. 12.11, de Sitter spacetime can be viewed as the hyperboloid[9]

$$-v^2 + x^2 + y^2 + z^2 + w^2 = \Lambda^2 \tag{12.60}$$

embedded in a five-dimensional spacetime with metric

$$ds^2 = dv^2 - dx^2 - dy^2 - dz^2 - dw^2 \,. \tag{12.61}$$

Here Λ is related to the vacuum energy density V by

$$\Lambda = \sqrt{\frac{3}{\kappa V}} \,, \tag{12.62}$$

where

$$\kappa = 8\pi G = \frac{8\pi}{M_{\rm Pl}^2} \,. \tag{12.63}$$

There are several common choices of coordinates for this spacetime. The entire spacetime can be covered by a single coordinate patch in which the hypersurfaces of constant time are three-spheres and the metric takes the form

$$ds^2 = dt^2 - \cosh^2(t/\Lambda)\, d\Omega_3^2 \,, \tag{12.64}$$

[9] It may appear that the "equator" of the hyperboloid is somehow distinguished. This is an artifact of the visualization. In fact, by a symmetry transformation one can bring any point in the spacetime to the equator.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

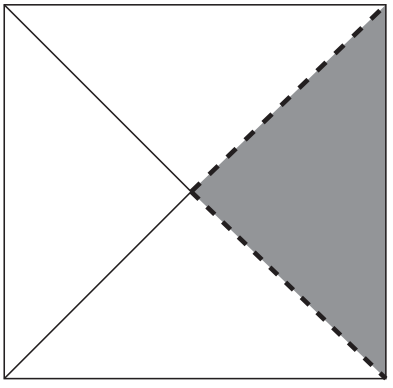


Fig. 12.12. The Penrose diagram of de Sitter spacetime. The shaded triangular region on the right represents the static patch. The vertical boundary on the right corresponds to $\tilde{r} = 0$. The dashed diagonal lines correspond to $\tilde{r} = \Lambda$, with the upper and lower ones being at $\tilde{t} = \infty$ and $\tilde{t} = -\infty$, respectively.

where $d\Omega_3$ is the line element on the unit three-sphere. This describes a closed Robertson–Walker spacetime[10] that expands exponentially fast at large time, with Hubble parameter $H = \Lambda^{-1}$. A consequence of this exponential expansion is the existence of an event horizon, beyond which an observer cannot see. This is nicely captured by a second set of coordinates for which the metric takes the static form

$$ds^2 = \left(1 - \frac{\tilde{r}^2}{\Lambda^2}\right) d\tilde{t}^2 - \left(1 - \frac{\tilde{r}^2}{\Lambda^2}\right)^{-1} d\tilde{r}^2 - \tilde{r}^2 d\Omega_2^2 \,. \tag{12.65}$$

The "static patch" covered by these coordinates is bounded by the horizon at $\tilde{r} = \Lambda$ and includes half of the three-sphere equator of the de Sitter hyperboloid. It corresponds to the causal diamond in the Penrose diagram of Fig. 12.12.

The existence of a black hole horizon gives the black hole thermal properties characterized by the Hawking temperature. In a similar fashion, the presence of the de Sitter horizon leads to the Gibbons–Hawking temperature [278]

$$T_{\rm dS} = \frac{H}{2\pi} = \frac{1}{2\pi\Lambda} \,. \tag{12.66}$$

This temperature will play an important role in our later discussion.

With a negative vacuum energy density V we obtain anti-de Sitter space. A convenient way of expressing its metric is

$$ds^2 = \left(1 + \frac{r^2}{L^2}\right) dt^2 - \left(1 + \frac{r^2}{L^2}\right)^{-1} dr^2 - r^2 d\Omega_2^2 \,, \tag{12.67}$$

[10] One can also introduce foliations that give flat or open Robertson–Walker coordinates, but these do not cover the entire spacetime.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

where

$$L = \sqrt{-\frac{3}{\kappa V}}\,. \tag{12.68}$$

There is no horizon, and therefore no temperature. Note, however, the behavior of the volume of a sphere as its radius increases. On a hypersurface of constant time, the two-sphere $r = R$ has area $4\pi R^2$, and encloses a volume

$$\mathcal{V}(R) = 4\pi \int_0^R dr\, \frac{r^2}{\sqrt{1 + (r^2/L^2)}}\,. \tag{12.69}$$

For $R \gg L$, this volume only grows as R^2, rather than as R^3, so the ratio of volume to area is bounded and never greater than $L/2$.

Let us turn now to our tunneling problem. Coleman and De Luccia argued that with gravity included the nucleation rate should still be based on a bounce solution, but with the Euclidean action now including a gravitational contribution. In other words,

$$\frac{\Gamma}{\mathcal{V}} = Ae^{-B} = Ae^{-[S(\phi_b) - S(\phi_{\rm fv})]}\,, \tag{12.70}$$

where ϕ_b is a bounce solution of the Euler–Lagrange equations following from a Euclidean action

$$S = \int d^4x\sqrt{g}\left[-\frac{1}{2\kappa}\,R + \frac{1}{2}g^{ab}\,\partial_a\phi\,\partial_b\phi + V(\phi)\right] + S_{\rm bdy} \tag{12.71}$$

that includes a Euclideanized Einstein–Hilbert term with the metric g_{ab} understood to have a positive definite Euclidean signature. The last term on the right-hand side, $S_{\rm bdy}$, is the Euclidean version of the Gibbons–Hawking boundary action [279] that must be added to remove the second derivative terms from the action.

We know that in flat spacetime, at least for a theory with a single scalar field, the bounce of minimum action has an O(4) symmetry. Although the corresponding result in curved spacetime has not been proven, it is widely believed to be true. Let us assume that this is the case. Any O(4)-symmetric Euclidean metric can be written in the form

$$ds^2 = d\xi^2 + \rho(\xi)^2 d\Omega_3^2\,. \tag{12.72}$$

The curvature scalar is then

$$R = -\frac{6}{\rho^2}\left(\rho\rho'' + \rho'^2 - 1\right), \tag{12.73}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

with primes denoting differentiation with respect to ξ, and the action is

$$\begin{aligned} S &= 2\pi^2 \int_{\xi_{\min}}^{\xi_{\max}} d\xi \left\{ \rho^3 \left[\frac{1}{2}\phi'^2 + V(\phi) \right] + \frac{3}{\kappa} \left(\rho^2 \rho'' + \rho\rho'^2 - \rho \right) \right\} + S_{\text{bdy}} \\ &= 2\pi^2 \int_{\xi_{\min}}^{\xi_{\max}} d\xi \left\{ \rho^3 \left[\frac{1}{2}\phi'^2 + V(\phi) \right] - \frac{3}{\kappa} (\rho\rho'^2 + \rho) \right\} \\ &\quad + \frac{6\pi^2}{\kappa} \left(\rho^2 \rho' \right) \Big|_{\xi=\xi_{\min}}^{\xi=\xi_{\max}} + S_{\text{bdy}} \,. \end{aligned} \tag{12.74}$$

The boundary term from the integration by parts is exactly canceled by the Gibbons–Hawking term, and so we have

$$S = 2\pi^2 \int_{\xi_{\min}}^{\xi_{\max}} d\xi \left\{ \rho^3 \left[\frac{1}{2}\phi'^2 + V(\phi) \right] - \frac{3}{\kappa} (\rho\rho'^2 + \rho) \right\} . \tag{12.75}$$

The scalar field equation is

$$\phi'' + \frac{3\rho'}{\rho}\phi' = \frac{dV}{d\phi} \,, \tag{12.76}$$

while the $\xi\xi$-component of the Euclidean Einstein equation gives

$$\rho'^2 = 1 + \frac{\kappa}{3}\rho^2 \left(\frac{1}{2}\phi'^2 - V \right) . \tag{12.77}$$

Another useful equation,

$$\rho'' = -\frac{\kappa}{3} \left(\phi'^2 + V \right) \rho \,, \tag{12.78}$$

is obtained by differentiating Eq. (12.77) and then using Eq. (12.76). By combining Eqs. (12.75) and (12.77), we can express the action for a solution of the field equations as

$$S = 4\pi^2 \int_{\xi_{\min}}^{\xi_{\max}} d\xi \left[\rho^3 V - \frac{3}{\kappa}\rho \right] . \tag{12.79}$$

The boundary conditions for the field equations depend on the topology of the solution which, as is often the case in general relativity, cannot be specified in advance. Subtracting Eqs. (12.77) and (12.78) shows that $\rho'^2 - \rho''\rho \geq 1$, from which we can conclude that ρ cannot have a nonzero minimum and must have at least one zero. The invariance of the metric under translation of ξ allows us to put one such zero at $\xi = \xi_{\min} = 0$. If this is the only zero of ρ, then ξ ranges from 0 to ∞ and the solution has the topology of R^4. The boundary conditions for Eq. (12.76) are then that $\phi'(0) = 0$, to avoid a singularity at the origin, and that $\phi(\infty) = \phi_{\text{fv}}$, just as in flat spacetime. However, it is possible that ρ has a second zero at some finite $\xi_{\max}$. In fact, this is always the case if $V(\phi) \geq 0$ everywhere and is not identically zero [126]. The solution then has the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

topology of a four-sphere, and the boundary conditions for Eq. (12.76) are that $\phi'(0)$ and $\phi'(\xi_{\max})$ both vanish. In this case there is no requirement that there be any point where $\phi = \phi_{\rm fv}$; in fact, although in the bounce solution $\phi - \phi_{\rm fv}$ can become exponentially small, there is no point on the four-sphere where it precisely vanishes.

With these preliminaries completed, let us study the Euclidean solutions, beginning with the case of a de Sitter false vacuum. We need both the bounce solution and the Euclidean version of the false vacuum. The latter is easily obtained. With ϕ constant and $V_{\rm fv}$ positive, Eq. (12.77) gives

$$\rho = \Lambda_{\rm fv} \sin(\xi/\Lambda_{\rm fv}) \tag{12.80}$$

and $\xi_{\max} = \pi\Lambda_{\rm fv}$. With this form for ρ, Eq. (12.72) gives the standard round metric on a four-sphere of radius $\Lambda_{\rm fv}$, which is the Euclidean counterpart of the Lorentzian hyperboloid. Using Eq. (12.79), we find that its action is

$$S(\phi_{\rm fv}) = -\frac{24\pi^2}{\kappa^2 V_{\rm fv}} = -\frac{8\pi^2}{\kappa}\,\Lambda_{\rm fv}^2\,. \tag{12.81}$$

Let us turn now to the bounce itself, beginning with a limit where gravitational effects might be expected to be small. Suppose that $V(\phi)$ is such that (1) the variation in V in the range between $\phi_{\rm fv}$ and $\phi_{\rm tv}$ is much smaller than $M_{\rm Pl}^4$ and (2) the radius of the flat-spacetime bounce (i.e., of the region where ϕ is appreciably different from $\phi_{\rm fv}$) is much less than $\Lambda_{\rm fv}$. Using Eq. (12.77), the first condition tells us that $\rho' \approx 1$, and hence $\rho \approx \xi$, in the region $0 \le \rho \ll \Lambda_{\rm fv}$. Equation (12.76) then reduces to the flat spacetime Eq. (12.14), so that in this region the bounce is essentially the same as in flat spacetime. Our second condition then tells us that ϕ becomes exponentially close to $\phi_{\rm fv}$ (although never quite equal to it) before leaving this region, so that outside this region ρ is well approximated by Eq. (12.80), except for a small shift $\xi \to \xi - \Delta$. Here Δ is of the order of $\mathcal{M}/M_{\rm Pl}^2$, where $\mathcal{M}$ is a mass scale characterizing the size and shape of the barrier in $V(\phi)$ that separates the two vacua. The Euclidean action is finite and negative, but subtracting the false vacuum action gives a positive value for B that differs from the flat spacetime result only by a fractional amount of order $\mathcal{M}^2/M_{\rm Pl}^2$.

We can also consider the thin-wall limit in which the bounce solution consists of a region of pure true vacuum, a region of pure false vacuum, and a relatively thin wall separating the two. The action of such a thin-wall configuration can be written as the sum of three terms, one from the wall and one each from the true vacuum and false vacuum regions.

If the $3\rho'\phi'/\rho$ term in Eq. (12.76) can be neglected, the matter field contribution to the wall action density σ is given by the flat-space expression, Eq. (12.30). The contribution from the second half of the integrand in Eq. (12.75) vanishes in this limit, so if $\rho = \bar\rho$ at the wall, the wall action is

$$S_{\rm wall} = 2\pi^2 \bar\rho^3 \sigma\,, \tag{12.82}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

In calculating the contribution from the vacuum regions, we need to keep in mind that the region bounded by a wall with curvature radius ρ can be either more or less than half of a four-sphere. Using Eq. (12.79), we obtain

$$S_{\rm vol} = -\frac{4\pi^2}{\kappa}\Lambda^2\left[1 \mp \left(1 - \frac{\bar{\rho}^2}{\Lambda^2}\right)^{3/2}\right], \tag{12.83}$$

where the upper (lower) sign applies if the region is less (more) than a hemisphere.

It might seem that there are four cases to consider, with the true vacuum and false vacuum regions each being either more or less than a hemisphere. However, the two cases in which the true vacuum is more than a hemisphere are ruled out because ρ'' would have to be positive at the wall, in contradiction with Eq. (12.78). Thus, we only need to consider two cases: Case A, with more than a hemisphere of false vacuum, and case B, with less than a hemisphere of false vacuum. The weak gravity limit is a limiting case of case A.

After subtracting the action of the pure false vacuum we obtain

$$B = 2\pi^2\bar{\rho}^3\sigma + \frac{4\pi^2}{\kappa}\Lambda_{\rm fv}^2\left[1 \mp \left(1 - \frac{\bar{\rho}^2}{\Lambda_{\rm fv}^2}\right)^{3/2}\right] - \frac{4\pi^2}{\kappa}\Lambda_{\rm tv}^2\left[1 - \left(1 - \frac{\bar{\rho}^2}{\Lambda_{\rm tv}^2}\right)^{3/2}\right], \tag{12.84}$$

where the upper and lower signs correspond to case A and case B, respectively. The bounce corresponds to a stationary point of this function with respect to $\bar{\rho}$. It turns out that for any value of the parameters such a stationary point occurs for only one of the two cases, with the bounce being of type A or of type B according to whether

$$\frac{\epsilon}{3\sigma} - \frac{\kappa\sigma}{4}$$

is positive or negative, respectively. In either case, the stationary point occurs when

$$\begin{aligned}\frac{1}{\bar{\rho}^2} &= \frac{1}{\Lambda_{\rm fv}^2} + \left(\frac{\epsilon}{3\sigma} - \frac{\kappa\sigma}{4}\right)^2 \\ &= \frac{1}{\Lambda_{\rm tv}^2} + \left(\frac{\epsilon}{3\sigma} + \frac{\kappa\sigma}{4}\right)^2 .\end{aligned} \tag{12.85}$$

Substituting this value for $\bar{\rho}$ into Eq. (12.81) gives [280]

$$\begin{aligned}B = \bar{\rho}^3\left[2\pi^2\sigma - \frac{4\pi^2}{\kappa}\Lambda_{\rm fv}^2\left(\frac{\epsilon}{3\sigma} - \frac{\kappa\sigma}{4}\right)^3 + \frac{4\pi^2}{\kappa}\Lambda_{\rm tv}^2\left(\frac{\epsilon}{3\sigma} + \frac{\kappa\sigma}{4}\right)^3\right] \\ + \frac{4\pi^2}{\kappa}\left(\Lambda_{\rm fv}^2 - \Lambda_{\rm tv}^2\right) .\end{aligned} \tag{12.86}$$

The signs in this expression are the same for both the type A and the type B solutions, because the explicit sign change in Eq. (12.84) is compensated by an

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

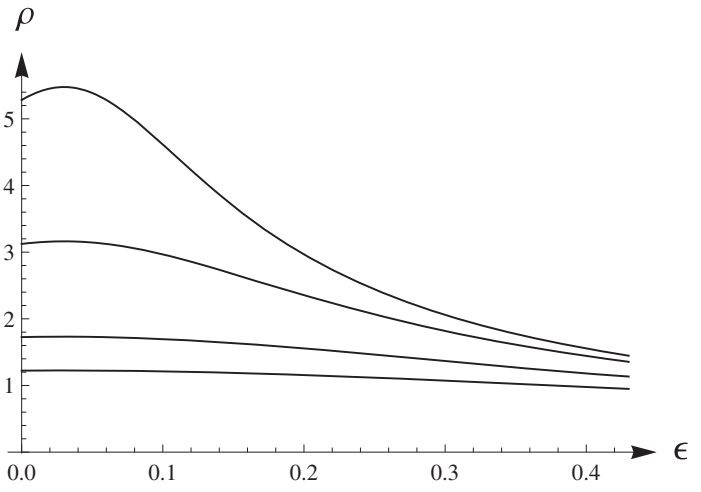


Fig. 12.13. The behavior of $\bar{\rho}$ as a function of ϵ in the thin-wall regime for four values of $\Lambda_{\rm fv}$. In units with $\kappa = 1$, these have $\sigma = 0.2$ and $V_{\rm fv} = 0.1$, 0.3, 1.0, and 2.0 (reading from the top). Note that at large ϵ the true vacuum is anti-de Sitter.

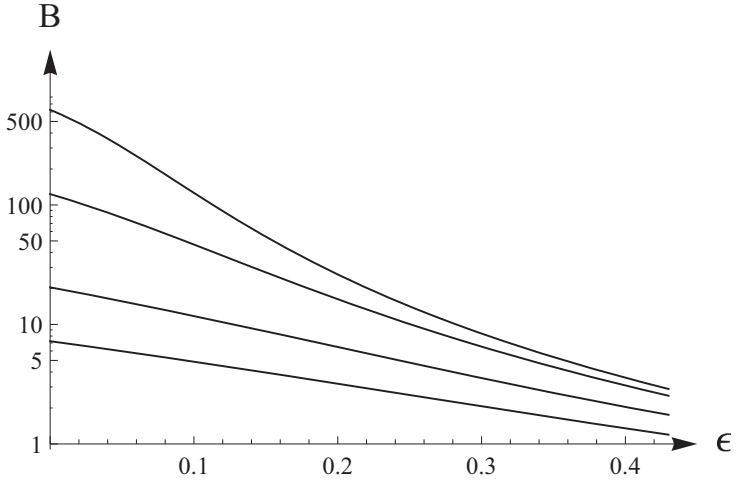


Fig. 12.14. The behavior of B as a function of ϵ in the thin-wall regime for four values of $\Lambda_{\rm fv}$. In units with $\kappa = 1$, these have $\sigma = 0.2$ and $V_{\rm fv} = 0.1$, 0.3, 1.0 and 2.0 (reading from the top). At large ϵ the true vacuum is anti-de Sitter.

implicit sign change coming from the fact that the second term in the square brackets has a different sign in the two cases.

The behaviors of $\bar{\rho}$ and B as functions of ϵ are illustrated in Figs. 12.13 and 12.14. Note that $\bar{\rho}$ reaches its maximum value, $\Lambda_{\rm fv}$, at a nonzero value of ϵ and then decreases as ϵ is reduced below this critical value. This should be compared

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

to the situation with gravity ignored, where the bounce radius is a monotonically decreasing function of ϵ. While it is perhaps not immediately obvious from Eq. (12.86), we see that B continues to be a monotonic function of ϵ although, in contrast with the flat spacetime case, it remains finite as ϵ approaches zero.

For a generic potential that fits neither of the above special cases the situation is more complex than in the corresponding flat spacetime case. It has already been noted that with an everywhere positive potential the bounce is topologically a four-sphere. One consequence is that the second term in Eq. (12.76), which corresponds to a damping term in the analogous upside-down potential problem in flat spacetime, changes sign and becomes an "antidamping" term for part of the interval. This invalidates the overshoot–undershoot argument that guaranteed the existence of a bounce in the absence of gravity. Indeed, there are situations in which there is no "Coleman–De Luccia" bounce; i.e., no O(4)-symmetric bounce with nonconstant ϕ crossing the top of the potential barrier. To understand this, note that the width of the bubble wall is largely determined by the shape of the barrier in $V(\phi)$, with a natural scale for this width being set by the curvature of V at $\phi = \phi_{\rm top}$, the top of the barrier. A very flat potential barrier corresponds to a low mass scale, or a large distance scale, and thus to a thick wall. In flat spacetime this is not a problem, and one can even have a "thin-wall" bounce with a thick wall by arranging for the bubble radius to be large enough.[11] However, it is definitely a problem with gravity included, because the compact Euclidean de Sitter space places an upper limit on the size of the bounce. The critical test is whether $|V''(\phi_{\rm top})| \equiv |V''_{\rm top}|$ is greater or less than $4/\Lambda^2_{\rm top}$ [281–284]. If $V''_{\rm top}$ is greater, then a Coleman–De Luccia bounce is guaranteed to exist. If not, then there may be no Coleman–De Luccia bounce. Even if one does exist, it may be quite different from what one might naïvely expect. We have already seen that in flat spacetime there are bounces in which ϕ never gets near its true vacuum value. In curved spacetime one finds bounces in which ϕ is never close to its false vacuum value either.

On the other hand, there are potentials that admit additional solutions in which ϕ passes back and forth over the potential barrier several times [283, 285]. These "oscillating bounces" have no counterparts in the nongravitational case.

Even when a potential is so flat that no Coleman–De Luccia bounce exists, there is always a second type of solution, first pointed out by Hawking and Moss [286]. This is a homogeneous solution with $\phi = \phi_{\rm top}$ everywhere and

$$\rho = \Lambda_{\rm top} \sin(\xi/\Lambda_{\rm top})\,. \tag{12.87}$$

Its Euclidean action is

$$S_{\rm HM} = -\frac{24\pi^2}{\kappa^2 V_{\rm top}}\,, \tag{12.88}$$

[11] Indeed, one might well have termed it the "large-bubble approximation".

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

leading to a tunneling exponent

$$\begin{aligned} B_{\rm HM} &= S_{\rm HM} - S(\phi_{\rm fv}) \\ &= \frac{24\pi^2}{\kappa^2}\left(\frac{1}{V_{\rm fv}} - \frac{1}{V_{\rm top}}\right) . \end{aligned} \tag{12.89}$$

The Hawking–Moss solution is much more akin to the high-temperature thermal production of a critical bubble than to a quantum tunneling bounce, a point that will elaborated on more fully in Sec. 12.6. Such a process becomes physically plausible when we recall that de Sitter space has a nonzero temperature $T_{\rm dS}$. In fact, the thermal nature of de Sitter space makes another sort of tunneling process possible.

The discussion of the Coleman–De Luccia bounces so far has been phrased in terms of the production of a true vacuum bubble, centered at the point $\xi = 0$, in a false vacuum background represented by the region centered at $\xi = \xi_{\rm max}$. However, one side of a sphere is not intrinsically different from the other. We could just as well interchange the roles of $\xi = \xi_{\rm max}$ and $\xi = 0$ and view the bounce as describing the production of a false vacuum bubble in a true vacuum background [287]. In the absence of gravity such a process, even if it could occur, would have negligible impact because the volume pressure and the surface tension would both act to cause the false vacuum bubble to collapse. In de Sitter spacetime, on the other hand, the cosmic expansion can protect such a bubble from collapse if it nucleates with a sufficiently large radius.

For the nucleation of a true vacuum bubble, the exponent B was the difference between the bounce action and the Euclidean action of the initial false vacuum. We should therefore expect B for the nucleation of a false vacuum bubble to be given by the difference between the bounce action and the Euclidean action of the true vacuum. The ratio of these rates would then be [287]

$$\begin{aligned} \frac{\Gamma_{\rm fv\to tv}}{\Gamma_{\rm tv\to fv}} &\approx \frac{e^{-[S(\phi_b)-S(\phi_{\rm fv})]}}{e^{-[S(\phi_b)-S(\phi_{\rm tv})]}} \\ &= e^{S(\phi_{\rm fv})-S(\phi_{\rm tv})} . \end{aligned} \tag{12.90}$$

Let us suppose that the true and false vacua are almost degenerate, with $V_{\rm fv} - V_{\rm tv} \ll V_{\rm fv}$. To leading approximation, the two vacua can be treated as having a common horizon length Λ and temperature $T_{\rm dS} = 1/(2\pi\Lambda)$, so that the exponent in the last line above becomes

$$\begin{aligned} S(\phi_{\rm fv}) - S(\phi_{\rm tv}) &= \frac{24\pi^2}{\kappa^2}\left(\frac{1}{V_{\rm tv}} - \frac{1}{V_{\rm fv}}\right) \\ &\approx -\frac{\frac{4\pi}{3}\Lambda^3(V_{\rm fv} - V_{\rm tv})}{T_{\rm dS}} \\ &\approx \frac{\Delta U_{\rm hor}}{T_{\rm dS}} , \end{aligned} \tag{12.91}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

where $\Delta U_{\rm hor}$ is the potential energy difference between a horizon volume of false vacuum and one of true vacuum.[12] A false vacuum bubble will expand after nucleation, raising the possibility that true vacuum bubbles might nucleate and expand within it, and then false vacuum bubbles within them, and so forth. One might expect to eventually reach an equilibrium in which the relative volumes of true and false vacua were related by a Boltzmann factor, with

$$\frac{\mathcal{V}_{\rm fv}}{\mathcal{V}_{\rm tv}} \approx \frac{\Gamma_{\rm fv\to tv}}{\Gamma_{\rm tv\to fv}} \approx e^{-\Delta U_{\rm hor}/T_{\rm dS}} \,. \tag{12.92}$$

In actuality, matters are more complicated. One has to specify the spacelike hypersurface on which $\mathcal{V}_{\rm fv}$ and $\mathcal{V}_{\rm tv}$ are being calculated. With a uniform de Sitter metric one would naturally take this to be a hypersurface of constant t. However, with differing vacuum energies the global spacetime is not purely de Sitter and the choice of hypersurface is hardly self-evident. Finding the correct way of evaluating this and similar ratios has become known as the measure problem, and is a major issue in theories that predict eternal inflation.

The discussion thus far has focused on the case of an initial de Sitter vacuum and implicitly assumed that the final vacuum was also de Sitter. The situation is rather similar for decay from de Sitter to anti-de Sitter or de Sitter to Minkowski. In both cases the bounce retains its four-sphere topology. If the thin-wall approximation is appropriate, the true vacuum region is described by

$$\rho = \begin{cases} \xi \,, & \text{Minkowski,} \\ \\ L_{\rm tv} \sinh(\xi/L_{\rm tv}) \,, & \text{anti-de Sitter,} \end{cases} \tag{12.93}$$

with the anti-de Sitter length scale L defined as in Eq. (12.68). Equations (12.84), (12.85), and (12.86) are still valid, provided that one makes the substitutions $\Lambda_{\rm tv}^2 \to -L_{\rm tv}^2$ or $\Lambda_{\rm tv} \to \infty$ for anti-de Sitter or Minkowski true vacua, respectively. In the case of a Minkowski true vacuum these results can also be written as [276]

$$\bar{\rho} = \frac{\bar{\rho}_0}{1 + (\bar{\rho}_0/2\Lambda_{\rm fv})^2} \tag{12.94}$$

and

$$B = \frac{B_0}{\left[1 + (\bar{\rho}_0/2\Lambda_{\rm fv})^2\right]^2} \,, \tag{12.95}$$

where $\bar{\rho}_0$ and B_0 are the flat spacetime expressions given in Eqs (12.34) and (12.35).

[12] There are subtleties in going from the second line to the third in this equation. Because the space is curved, the horizon volume is not given simply by the flat space formula $(4\pi/3)\Lambda^3$. However, when redshift factors are properly taken into account, the curvature effects cancel out in the calculation of the horizon energy difference. This will become clearer from the discussion in the next section.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

Because the bounces for these transitions, like those that describe de Sitter to de Sitter transitions, are topologically four-spheres, one could imagine interchanging the initial and final states and viewing these as describing transitions upward from an initial Minkowski or anti-de Sitter vacuum to a de Sitter one. However, to obtain the tunneling exponent B we must subtract the action for the initial state from the action of the bounce. For both the Minkowski and anti-de Sitter cases the vacuum action is equal to $-\infty$. Subtracting this from a finite bounce action gives $B = \infty$, and thus a vanishing bubble nucleation rate.[13] This result is of course not surprising, since there is no Gibbons–Hawking temperature in a spacetime with negative or vanishing cosmological constant.

We can also consider the nucleation of bubbles of a lower vacuum from an initial anti-de Sitter or Minkowski vacuum. Both the Euclideanized initial vacuum and the bounce solution itself are infinite in extent, with the topology of R^4 and with the scalar field equal to its false vacuum value at Euclidean infinity.

In the thin-wall limit the equations from the de Sitter case can again be used, with the replacements $\Lambda^2 \to -L^2$ or $\Lambda \to \infty$ as appropriate. With the upper sign in Eqs.(12.83) and (12.84) both the bounce and the initial false vacuum are infinite in extent. B is the (finite) difference between the two formally infinite actions. For the particular case of anti-de Sitter bubbles nucleating in Minkowski space, the analogues of Eqs. (12.94) and (12.95) are

$$\bar{\rho} = \frac{\bar{\rho}_0}{1 - (\bar{\rho}_0/2L_{\rm tv})^2} \tag{12.96}$$

and

$$B = \frac{B_0}{\left[1 - (\bar{\rho}_0/2L_{\rm tv})^2\right]^2} \,. \tag{12.97}$$

The analogues of the type B bounces that correspond to the lower signs in Eqs. (12.83) and (12.84) have finite false vacuum regions and thus finite actions. Subtracting the negatively infinite initial state actions from these gives $B = \infty$. Hence, they do not lead to bubble nucleation. Thus, in the thin-wall approximation there is no tunneling if [276]

$$\epsilon \leq \frac{3}{4}\,\kappa\sigma^2 \,. \tag{12.98}$$

This might seem surprising, since one might expect to always be able to tunnel from a vacuum of higher energy to one of lower energy. However, it must be remembered that the bounce does not describe tunneling from vacuum to vacuum, but rather from a vacuum to a state with a bubble surrounded by vacuum. There is a conserved energy if the spacetime is asymptotically flat, so a

[13] Note that this result depends crucially on the proper treatment of the boundary terms. If we had omitted the Gibbons–Hawking term in Eq. (12.71), we would have obtained either zero (in the Minkowski case) or $+\infty$ (in the anti-de Sitter case) for the initial state action.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

bubble that nucleates in Minkowski spacetime must have zero total energy; i.e., it must be large enough that the negative vacuum energy in the interior compensates for the positive energy in the bubble wall. With gravity neglected one can always make the bubble large enough that this is the case. When gravitational effects are included this is not necessarily so, because of the peculiar feature of anti-de Sitter spacetime that the volume enclosed by a three-sphere only grows as R^2 at large radius. If the false vacuum is anti-de Sitter rather than Minkowski, the Poincaré symmetry generator corresponding to the total energy must be replaced by a generator of the O(3,2) anti-de Sitter symmetry, but a similar result is obtained [288].

12.6 Interpretation of the bounce solutions

As derived for flat spacetime, the bounce solution plays a dual role. It not only determines the exponent in the bubble nucleation rate, but it also describes the optimal tunneling path through configuration space and gives the most likely configuration of the bubble at nucleation. We saw in Sec. 12.1 that a sequence of spacelike slices through the bounce gives a family of configurations along this optimal path.

A similar understanding of the bounce seems natural for the bounces with R^4 topology that govern decay from a Minkowski or anti-de Sitter false vacuum. However, the situation is not so straightforward for the bounces associated with bubble nucleation in a de Sitter vacuum. Because these bounces are finite in extent, there is no hypersurface at $\tau = -\infty$ to be identified with the initial false vacuum configuration. Indeed, ϕ never quite reaches its false vacuum value anywhere on the bounce. For "small bounces", such as that illustrated in Fig. 12.15, these may seem minor points. The slice through the lower part of the four-sphere in that figure gives a close approximation to the initial vacuum. The equatorial three-sphere contains a nucleated bubble very similar to that obtained without gravity. When mapped onto the equator of the de Sitter hyperboloid, this seems to be a plausible initial configuration for the Lorentzian evolution after nucleation. However, the identification of the initial configuration is more problematic for the large-bounce solution shown in Fig. 12.16. The difficulty is even greater for the Hawking–Moss solution. Not only is there no reference to the initial state, but identification of the equatorial slice with the configuration after the transition would imply an implausible process in which the field over all of de Sitter space fluctuated to the top of the potential barrier.

The difficulties here can be attributed in part to the fact that the formalism we have been using was obtained by arguing from analogy, rather than by an explicit derivation of the answer to a precisely posed question. Some insight can be obtained by considering a limiting case in which the Coleman–De Luccia prescription for the bounce follows straightforwardly from the results that we have already obtained [277]. To this end, consider the limit where the variation

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

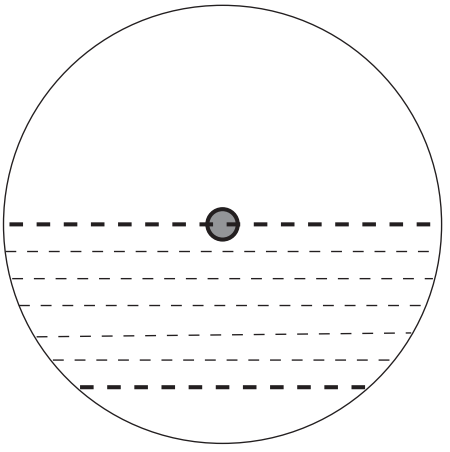


Fig. 12.15. A plausible (but incorrect) curved spaced analogue of Fig. 12.3 for the case when the radius of the shaded true vacuum region is much less than the de Sitter horizon length.

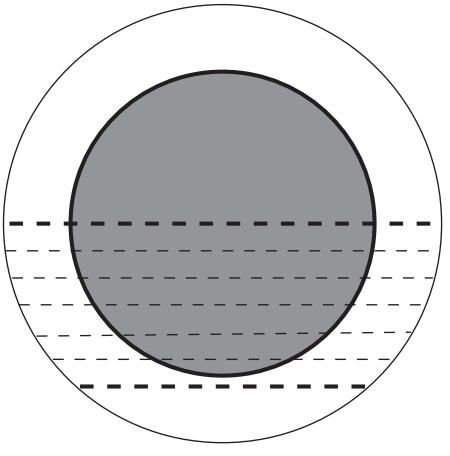


Fig. 12.16. The extension of Fig. 12.15 to the case when the radius of the true vacuum region is comparable to the de Sitter horizon length.

of $V(\phi)$ in the interval between $\phi_{\rm fv}$ and $\phi_{\rm tv}$ is small compared to its actual value. To a first approximation, the tunneling can then be viewed as taking place within a fixed de Sitter background. Working in this fixed background approximation, let us focus on the fields within the static patch, the region covered by the metric of Eq. (12.65). In terms of the static coordinates (with the tildes omitted) the matter part of the real-time Lorentzian action on this patch is

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

$$S_L = \int_{-\infty}^{\infty} dt \int_0^{\Lambda} dr \int d^2\Omega \sqrt{-\det g} \left[\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \, \partial_\nu \phi - V(\phi) \right]$$
$$= \int_{-\infty}^{\infty} dt \int_0^{\Lambda} dr \int d^2\Omega \sqrt{\det h} \left[\frac{1}{2P(r)} \left(\frac{\partial \phi}{\partial t} \right)^2 - \frac{1}{2} P(r) \, h^{ij} \partial_i \phi \, \partial_j \phi - P(r) V(\phi) \right] , \tag{12.99}$$

where $h_{ij} = g_{ij}$ and

$$P^2(r) \equiv g_{tt} = 1 - \frac{r^2}{\Lambda^2} . \tag{12.100}$$

Although this action had its origins in a field theory in a curved spacetime, we can equally well choose to view it as describing a theory, on a curved three-dimensional space of finite extent, whose interactions happen to have a peculiar position dependence that is described by the various factors of $P(r)$. Because the background spacetime is static, there is a well-defined energy functional

$$E = \int d^3x \sqrt{\det h} \left[\frac{1}{2P(r)} \left(\frac{\partial \phi}{\partial t} \right)^2 + \frac{1}{2} P(r) \, h^{ij} \partial_i \phi \, \partial_j \phi + P(r) V(\phi) \right] , \tag{12.101}$$

for this theory, and so there is no impediment to applying the methods of Secs. 12.1–12.4. In view of the nonzero de Sitter temperature, it seems reasonable to view this as thermal tunneling and to look for bounces that are periodic in Euclidean time with period $\beta = 1/T_{\rm dS} = 2\pi\Lambda$. These will be solutions of the Euler–Lagrange equations that follow from the Euclidean action

$$S_E = \int_{-\pi\Lambda}^{\pi\Lambda} d\tau \int_0^{\Lambda} dr \int d^2\Omega \sqrt{\det h} \left[\frac{1}{2P(r)} \left(\frac{\partial \phi}{\partial \tau} \right)^2 + \frac{1}{2} P(r) \, h^{ij} \partial_i \phi \, \partial_j \phi + P(r) V(\phi) \right] . \tag{12.102}$$

As illustrated in Fig. 12.17, these bounces can be understood as a sequences of field configurations, each of which is on a three-dimensional ball bounded by the horizon two-sphere $r = \Lambda$. The configurations are labeled by a Euclidean time τ that runs from $-\pi\Lambda$ to $\pi\Lambda$, with the hypersurfaces $\tau = -\pi\Lambda$ and $\tau = \pi\Lambda$ identified.

As in flat spacetime, there are two modes of finite temperature decay. For the case of thermally assisted tunneling, which is shown in Fig. 12.17, we can take the hypersurface at $\tau = -\pi\Lambda$ to give the approximately false vacuum initial configuration, and the hypersurface $\tau = 0$ to give the configuration immediately after tunneling. This bounce appears to bear little resemblance to that of Coleman and De Luccia. However, we can define a Euclidean metric

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

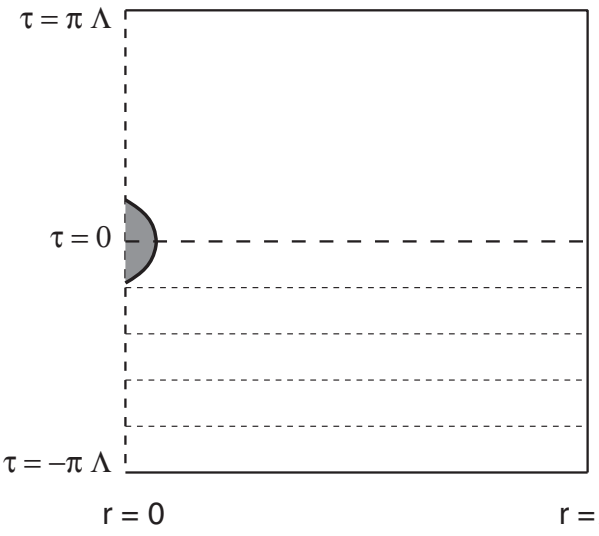


Fig. 12.17. The Coleman–De Luccia bounce, as it appears in static de Sitter coordinates. The horizontal lines represent hypersurfaces of constant τ. These are three-dimensional balls with their centers at $r = 0$ (the dotted line on the left) and their outer edges at $r = \Lambda$ (the solid line on the right). The shaded region is the true vacuum part of the bounce.

$$\begin{aligned}\tilde{g}_{ab}\,dx^a dx^b &= P^2 d\tau^2 + h_{ij} dx^i dx^j \\ &= \left(1 - \frac{r^2}{\Lambda^2}\right) d\tau^2 + \left(1 - \frac{r^2}{\Lambda^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\,.\end{aligned} \tag{12.103}$$

With r ranging from 0 to Λ and $\tau = -\pi\Lambda$ and $\tau = \pi\Lambda$ identified, this is just the round metric on a four-sphere. The Euclidean action of Eq. (12.102) can now be rewritten as

$$S_E = \int d^4x \sqrt{\det \tilde{g}} \left[\frac{1}{2}\tilde{g}_{ab}\,\partial_a\phi\,\partial_b\phi + V(\phi)\right] \tag{12.104}$$

and the action for the decay rate is

$$\frac{\Gamma}{\mathcal{V}} \sim e^{-[S_E(\phi_b) - S_E(\phi_{\rm fv})]}\,. \tag{12.105}$$

The last two equations differ from the corresponding ones in the Coleman–De Luccia formulation in that the Einstein–Hilbert curvature term has been omitted from the Euclidean action. However, in the fixed background approximation that we are using this term is the same for both the bounce and the pure false vacuum, and so cancels between the two terms in the exponent.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

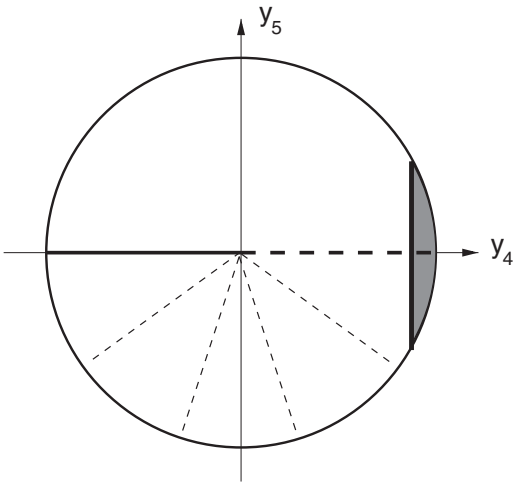


Fig. 12.18. The bounce of Fig. 12.17 mapped onto a four-sphere by the coordinate transformation of Eq. (12.106). This identifies the three-spheres $r = \Lambda$ for all values of τ. The solid and dashed lines correspond to those in Fig. 12.17.

The interpretation of the metric in terms of a four-sphere can be seen more explicitly by defining

$$\begin{aligned} y_1 &= r \sin\theta \cos\phi \\ y_2 &= r \sin\theta \sin\phi \\ y_3 &= r \cos\theta \\ y_4 &= \sqrt{\Lambda^2 - r^2}\, \cos(\tau/\Lambda) \\ y_5 &= \sqrt{\Lambda^2 - r^2}\, \sin(\tau/\Lambda) \end{aligned} \tag{12.106}$$

so that the Euclidean space of the bounce is just the four-sphere $\sum y_a^2 = \Lambda^2$ in a five-dimensional Euclidean space. In terms of these variables the constant τ slices of Fig. 12.17 become the "radial" slices shown in Fig. 12.18. Each of these slices represents a three-dimensional ball, with the center of the ball, $r = 0$, lying on the outer edge of the diagram (where $y_4^2 + y_5^2 = \Lambda^2$), and the outer boundary of the ball, the horizon two-sphere at $r = \Lambda$, being mapped to the two-sphere[14] defined by $y_4 = y_5 = 0$. Contrary to what one might have guessed—and to what has often been assumed in the literature—the configuration after tunneling does not correspond to the full equatorial slice of the bounce. Instead, half of this slice gives the configuration before tunneling and the other half the configuration after

[14] Notice that this mapping takes the horizon two-spheres at all values of τ to a single two-sphere. Although this could have led to a singularity in either the metric or the scalar field as viewed on the four-sphere, this turns out to not be the case for the O(4)-symmetric bounce.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

tunneling. In neither case are we given a full spatial hypersurface, but instead only a constant time slice through the static patch that is bounded by the horizon. The nature of the spacetime beyond the horizon never enters the problem, except insofar as it provides the thermal background represented by $T_{\rm dS}$.

The second mode of decay, fluctuation to a saddle point over the barrier, is represented by a τ-independent bounce. One such example is the Hawking–Moss solution in which the field throughout the entire horizon volume—not the entire universe—is thermally excited to the top of the potential energy barrier.[15] This manifestly τ-independent bounce is completely analogous to the high-temperature τ-independent bounce in flat spacetime. This thermal interpretation is further supported by the expression for the tunneling exponent given in Eq. (12.89). In the fixed background approximation, with $V_{\rm top} - V_{\rm fv} \ll V_{\rm fv}$, we have [see also Eq. (12.91)]

$$B_{\rm HM} \approx \frac{\left(\frac{4\pi}{3}\Lambda^3\right)(V_{\rm top} - V_{\rm fv})}{T_{\rm dS}}\,. \tag{12.107}$$

The numerator is precisely the jump in the energy of Eq. (12.101) needed to take a uniform false vacuum horizon volume to a configuration with the field everywhere at $\phi_{\rm top}$. [Note that the metric factors in the integrand of Eq. (12.101) cancel in such a way that one ends up with the result that would be obtained by a naïve calculation that used a flat space volume calculation and ignored redshift factors.]

Perhaps surprisingly, the Hawking–Moss solution is not the only τ-independent bounce. By reorienting the Coleman–De Luccia bounce of Fig. 12.18 we can obtain the τ-independent bounce solution illustrated in Fig. 12.19 [277]. This corresponds to the purely thermal creation of a true vacuum region that intersects the horizon. Thus, the same Euclidean solution can correspond to either thermally assisted quantum tunneling or to a purely thermal process. This is perhaps not so surprising if we recall that the de Sitter temperature itself is a quantum effect.

There is another class of solutions whose form emphasizes the significance of inferring the configuration after bubble nucleation from just half of the equatorial slice. As was noted in Sec. 12.4, in flat spacetime the dominant bounces at high temperature are independent of the periodic Euclidean time. Hence, assuming O(3) symmetry of the spatial slices, these bounces have an O(2)×O(3) symmetry. In a de Sitter context, there is a natural generalization of these bounces, with the temperature given by $T_{\rm dS}$. The O(2)×O(3) symmetry is retained, although now the constant τ slices only extend out to the horizon. The scalar field is the same on all of these slices, which correspond to the horizontal dotted lines

[15] The stochastic interpretation [296–298] of the Hawking–Moss bounce may be viewed as describing a mechanism by which the thermal distribution of field configurations is established.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

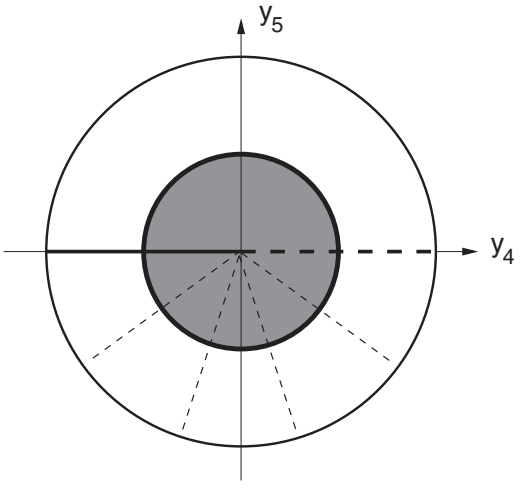


Fig. 12.19. The bounce of Fig. 12.18 rotated to give a τ-independent bounce.

in Fig. 12.17 or to the radial dotted lines in Fig. 12.18. A full equatorial slice would then contain two copies of a true vacuum bubble. Interpreting this full slice as giving the configuration after the transition would lead to the surprising conclusion that the bounce described the production of a pair of bubbles [301]. On the other hand, restricting to the horizon region described by the half-slice yields the more plausible production of a single bubble [302].

In contrast with the flat spacetime case where there is a demonstration that the bounce of lowest action has O(4) symmetry, there is no similar result when gravity is included. These O(2)×O(3)-symmetric bounces thus provide a useful test of whether the O(4)-symmetric bounces remain dominant in curved spacetime. For all the cases that have been examined, the O(4)-symmetric bounce has the lower action [301, 302], although there are parameter ranges where the two types of bounces both merge into the Hawking–Moss solution and thus have equal action in the limit.

The gravitational effects on the pre-exponential factor have not been discussed thus far. When the fixed-background approximation is appropriate one can simply follow the methods used in flat spacetime and obtain an expression involving a functional determinant of the second variation of the matter action. As before, one must seek a bounce with a single negative mode. Multiple negative modes indicate that the bounce is not even a local minimum of B and must therefore be discarded. This turns out to be the case both for the Hawking–Moss solution when $V''_{\rm top} > 4/\Lambda^2_{\rm top}$ (i.e., precisely when an ordinary Coleman–De Luccia bounce is guaranteed to exist), as well as for all of the oscillating bounce solutions [289, 290].

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

Outside this approximation the situation is less clear. First, there is as yet no derivation leading to a precise expression for the decay rate. Second, even if one assumes a similar form for the pre-exponential factor, matters are complicated by the fact that the dynamics of the gravitational and matter fields are no longer decoupled and must both contribute nontrivially to the functional determinant. However, gravity, like any gauge theory, is a constrained system. If these constraints are not properly taken into account when determining the normal modes, spurious negative modes can appear. Some partial results have been obtained, with demonstrations of the existence of a negative mode under certain conditions [291–294], but also arguments suggesting the absence of a negative mode in other regimes [295]. More work is needed to clarify matters here.

12.7 Curved spacetime evolution after bubble nucleation

The discussion of decay from a de Sitter vacuum in Sec. 12.6 was framed in the context of a fixed de Sitter background metric. We saw that the configuration of the field within the static patch after nucleation was given by one-half of an equatorial slice of the bounce. This suggests that in the more general case the initial data for classical evolution after bubble nucleation is given by that part of an equatorial slice that lies within the horizon of some given point P.

Because the initial configuration outside the horizon is not specified, the future evolution is only determined within the causal diamond of P. It might then seem to be best to work with coordinates that were naturally restricted to this region; e.g., the static de Sitter coordinates in the fixed background approximation. However, for the case of an O(4)-symmetric Coleman–De Luccia bounce it turns out to be more convenient to work with coordinates that extend beyond the causal diamond, keeping in mind that only a portion of the resulting classical evolution is reliable.

An O(4)-symmetric bounce has a Euclidean metric of the form

$$\begin{aligned} ds^2 &= d\xi^2 + \rho(\xi)^2 d\Omega_3^2 \\ &= d\xi^2 + \rho(\xi)^2 \left(d\tau'^2 + \cos^2\tau'\, d\Omega_2^2 \right) . \end{aligned} \tag{12.108}$$

The equatorial hypersurface corresponds to $\tau' = 0$. For the case of an exactly de Sitter Euclidean metric, one-half of this hypersurface coincides with the $\tau = 0$ hypersurface of the metric of Eq. (12.103), although for nonzero values the hypersurfaces of constant τ' and of constant τ do not coincide. On the hypersurface $\tau' = \tau = 0$, we have $d\phi/d\tau' = d\phi/d\tau = 0$.

As was the case for flat spacetime, the regions inside and outside the light cone of the origin (i.e., the point $\xi = 0$) must be treated separately. Outside the light cone the metric can be written as

$$\begin{aligned} ds^2 &= -d\xi^2 - \rho(\xi)^2 d\Omega_S^2 \\ &= -d\xi^2 - \rho(\xi)^2 \left(-d\tau'^2 + \cosh^2\tau'\, d\Omega_2^2 \right) \end{aligned} \tag{12.109}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

with the Euclidean $d\Omega_3^2$ replaced by the Lorentzian $d\Omega_S^2$, which is the metric on a unit hyperboloid with a spacelike normal vector. As a result of this change, $\rho = 0$ specifies not just a point, but rather the full light cone of the origin. The metric function ρ and the scalar field obey

$$\rho'^2 = 1 + \frac{\kappa}{3}\rho^2\left(\frac{1}{2}\phi'^2 - V\right) \tag{12.110}$$

and

$$\phi'' + \frac{3\rho'}{\rho}\phi' = \frac{dV}{d\phi}\,, \tag{12.111}$$

with primes denoting differentiation with respect to ξ. These are the same as the equations for the bounce. Therefore, in complete parallel with the flat spacetime results, $\phi(\xi)$ and $\rho(\xi)$ have the same functional form in this region as they did in the bounce.

The metric inside the light cone of the origin can be written in the open Robertson–Walker form

$$ds^2 = dt^2 - a(t)^2 d\Omega_T^2\,, \tag{12.112}$$

where $d\Omega_T^2$, the metric on a unit hyperboloid with a timelike normal vector, was given in Eq. (12.48). The scale factor obeys the Friedmann equation

$$\dot{a}^2 = 1 + \frac{\kappa}{3}a^2\left(\frac{1}{2}\dot{\phi}^2 + V\right) \tag{12.113}$$

(with overdots denoting differentiation with respect to t) and the scalar field satisfies

$$\ddot{\phi} + \frac{3\dot{a}}{a}\dot{\phi} = -\frac{dV}{d\phi}\,. \tag{12.114}$$

The zero of a, which can be taken to occur at $t = 0$, occurs on the light cone. The interior and exterior metrics join smoothly across this boundary.[16]

Just as in the absence of gravity, in the thin-wall approximation the bubble wall lies in the region outside the light cone. If there is not a well-defined thin wall at the time of nucleation, one develops over time and becomes thinner as the bubble expands. Also as before, the trajectory of the bubble wall asymptotically approaches a light cone. In Minkowski spacetime, any pair of light cones eventually intersects. This is not so in de Sitter spacetime, because of the presence of horizons. As a result, a given pair of bubbles expanding in a de Sitter false vacuum will never collide if they are nucleated at spacetime points that are too far apart. Consequently, if the bubble nucleation rate is too low compared

[16] For decay from a de Sitter vacuum there is a second zero of ρ at $\xi_{\max}$. If the continuation from Euclidean to Lorentzian signature were carried out over the entire equatorial slice, there would be a second light cone boundary at $\xi = \xi_{\max}$, with a second open Robertson–Walker region on the other side. However, as we have seen, there is no physical justification for extending the continuation to the region near this second zero.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

to the cosmic expansion, the bubbles do not coalesce to form a uniform region of the equilibrium true vacuum phase. In particular, the regions of true vacuum do not even percolate if the dimensionless ratio $\gamma = (\Gamma/\mathcal{V})H^{-4} = (\Gamma/\mathcal{V})\Lambda^4$ is less than a critical value $\gamma_{\rm cr}$ that has been shown [126] to lie in the range $0.24 > \gamma_{\rm cr} > 1.1 \times 10^{-6}$.

If γ is less than this critical value, the true vacuum regions remain confined to isolated bubbles or clusters of bubbles. We saw in Sec. 12.3 that the interior of an uncollided flat spacetime bubble was essentially pure vacuum, with almost all of the energy from the latent heat of transition going into the acceleration of the bubble walls. A similar effect occurs in curved spacetime. It can be quantified by considering the entropy density inside the bubble. As in flat spacetime, the scalar field in the interior of the bubble eventually begins to oscillate about its true vacuum value. After some time the energy in these oscillations is thermalized, generating an entropy density S. An adiabatically expanding Robertson–Walker universe is characterized by the dimensionless quantity a^3S; in our universe, observations indicate that this is greater than 10^{90}. By contrast, it was shown in Ref. [126] that for a generic potential with a de Sitter or Minkowski true vacuum a^3S in the interior of the bubble would be of order unity, implying that a slow first-order transition would not lead to an acceptable cosmology. However, Linde [299] and Albrecht and Steinhardt [300] realized that a non-generic potential could give very different results. If $V(\phi)$ has a suitably flat region between the potential barrier and the true vacuum minimum at $\phi_{\rm tv}$, the scale factor can grow exponentially before ϕ enters the oscillatory regime. As a result, this "new inflation" scenario can yield values of a^3S that easily satisfy the observational constraint.

Although this scenario would allow the observed universe to evolve inside a single bubble, bubbles nucleated near this reference bubble would collide with it, producing a bubble cluster. The exponentially expanding de Sitter spacetime in the exterior false vacuum region leads to a very unusual distribution of bubble sizes in this cluster [126], with almost all bubbles being exponentially smaller than the largest bubble.[17] The extent to which the distribution of bubbles in the cluster reveals any indication of the onset of inflation in the exterior region, and indeed whether a typical observer would see any observable traces of the bubble collisions, have been the subject of a number of recent papers [304–309]

One can also consider the case where the true vacuum has a negative vacuum energy and so is an anti-de Sitter spacetime. In this case the interior region has a very different fate [276]. If the thin-wall approximation were exact, ϕ would be precisely equal to $\phi_{\rm tv}$ and Eq. (12.113) would take the form

$$\dot{a}^2 = 1 + \frac{\kappa}{3}a^2 V_{\rm tv} = 1 - \frac{a^2}{L^2} \tag{12.115}$$

[17] By the de Sitter analogue of a Lorentz transformation, one can choose a frame in which the three largest bubble in the cluster are of equal size.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013

with the solution

$$a = L \sin(t/L)\,. \tag{12.116}$$

Like the zero on the light cone at $t = 0$, the zero of a at $t = \pi L$ is a coordinate singularity, not a physical one. However, Eq. (12.114) shows that a real singularity in ϕ, and hence a physical singularity in the metric, will develop unless $\dot{\phi}$ is precisely zero whenever $a = 0$. The corrections to the thin-wall limit, even though exponentially small, give ϕ a time-dependence such that for generic initial conditions $\dot{\phi}$ is nonzero at the zero of a, and a singularity ensues.

This difficulty persists even when the thin-wall approximation is not applicable [310]. Equation (12.113) implies that

$$\dot{a}^2 \geq 1 - \frac{a^2}{L^2}\,, \tag{12.117}$$

so that a second zero of a is inevitable once a starts decreasing and becomes small compared to L. Except for exceptional values of $\phi(0)$, $\dot{\phi}$ is nonvanishing at this second zero, and one is inevitably driven to a singularity within a finite time after the nucleation of the bubble.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 18 Nov 2026 at 09:54:10, subject to the Cambridge Core terms of use , available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139017787.013