Null energy condition violation: Tunneling versus the Casimir effect

Jean Alexandre and Drew Backhouse
Theoretical Particle Physics and Cosmology, King's College London, London WC2R 2LS, United Kingdom

(Received 17 January 2023; accepted 18 April 2023; published 28 April 2023)

We show that tunneling between two degenerate minima, as allowed in a finite volume, leads to a nonextensive symmetric ground state. This results in a null energy condition violation for sufficiently low temperatures, when a continuous set of momenta in the box containing the field is assumed. Taking into account discrete momenta can modify this picture and is achieved via the addition of the Casimir energy to the tunneling-induced ground state energy. Focusing on zero-temperature, these nontrivial effects are found to compete, depending on the typical length scales involved.

DOI: 10.1103/PhysRevD.107.085022

I. INTRODUCTION

Spontaneous symmetry breaking (SSB) is strictly speaking valid for infinite volumes only, where tunneling between degenerate vacua is completely suppressed. On the other hand, for a field confined in a box of finite volume, tunneling between degenerate vacua is allowed, and we study here the energetic consequences.

Involving tunneling in the quantization of a system automatically takes into account the different vacua and is known to lead to a convex effective action [1–9]. This is not the case in the situation of SSB, where the different vacua are decoupled and quantization over a single vacuum does not necessarily lead to convexity. Taking into account several degenerate vacua in the partition function comes with a remarkable energetic feature, generated dynamically: the effective action is nonextensive, as was shown in [10,11] with a semiclassical approximation for the partition function.

The latter works were done in an $O(4)$-symmetric Euclidean spacetime though, and to account for a full description of tunneling one needs a finite spatial volume V and an independent large Euclidean time β. The natural context for these studies is therefore equilibrium field theory at a finite-temperature $T = 1/\beta$. The corresponding quantum mechanics study was done by the authors of [12], involving a gas of instantons/anti-instantons which dominates the partition function in the limit of a small temperature. It is shown there that the null energy condition (NEC) (see [13,14] for reviews) is violated as a consequence of a nonextensive effective action induced by tunneling. The present article extends this study to full four-dimensional quantum fluctuations, and we find that NEC violation occurs in any finite volume for sufficiently low temperatures.

Our study does not, however, deal with high-temperature symmetry restoration, as seen in the Kibble-Zureck mechanism [15,16]. We are instead interested in the low-temperature regime, where tunneling dominates over thermal fluctuations providing an opportunity to violate the NEC, which the Kibble-Zurek mechanism does not.

We first evaluate quantum corrections with continuous momentum for fluctuations above each saddle point, to describe the fundamental dynamical mechanism induced by tunneling. We then take into account the modification arising from discrete momentum in a finite volume, using results known from studies of the Casimir effect (see [17] for a review). The latter is known to be either attractive or repulsive, depending on the geometry of the box containing the field, as well as the boundary conditions the field satisfies on the walls of the box. As a consequence, as far as NEC violation is concerned, the difference between discrete and continuous momentum can play an important role.

In Sec. II we describe the semiclassical approximation in which the partition function is derived, to take into account the different saddle points which are relevant to tunneling: static saddle points and the instanton/anti-instanton dilute gas. Details of the calculations with continuous momentum are given in Appendixes A and B. Section III focuses on the ground state of the effective action, with a nonextensive energy density providing the origin of NEC violation. The maximum effect occurs at zero temperature and is the regime in which we introduce corrections arising from discrete momentum in Sec. IV via the Casimir energy. We find that tunneling and the Casimir effect compete when the typical size of the box containing the field is of the order of the Compton wavelength of the

corresponding particle. For a larger box, the Casimir effect seems to be dominant.

To summarize our results: in the low-temperature regime, the sum of energy density ρ and pressure p can be written in the form

$$\rho + p \simeq A_{\text{finite-T}} + B_{\text{tunneling}} + C_{\text{Casimir}},$$

where

(i) the finite-temperature contribution A is always positive (and vanishes exponentially for $T \to 0$);
(ii) the tunneling contribution B, calculated with continuous momentum, is always negative (and vanishes exponentially for $V \to \infty$);
(iii) the discrete momentum correction C has a sign which depends on the geometry and topology of the finite box containing the field (and vanishes for $V \to \infty$).

As expected, the NEC is satisfied at zero temperature and for infinite volume, where $\rho + p = 0$ for a homogeneous vacuum.

II. SEMICLASSICAL APPROXIMATION

A. Model

Consider a single real scalar field $\phi(t, x)$ in Euclidean space, at finite-temperature $T = 1/\beta$ and in a three-dimensional spatial volume V, described by the Euclidean action

$$\int_0^\beta dt \int_V d^3x \left(\frac{1}{2}(\partial\phi)^2 + \frac{\lambda}{24}(\phi^2 - v^2)^2 + j\phi \right). \quad (1)$$

The finite volume is represented by a physical box containing the scalar field, in which we assume continuous momenta to calculate quantum corrections. Section IV discusses corrections arising from discrete momentum and the boundary conditions the field satisfies at the walls of the box. Finite temperature requires field configurations to have periodic boundary conditions in Euclidean time and, as later discussed, has an impact on the saddle point configurations which are allowed in the partition function.

Introducing the dimensionless variables $\tau \equiv \omega t$ and

$$\omega \equiv v\sqrt{\frac{\lambda}{6}}, \qquad \varphi \equiv \sqrt{\frac{\lambda}{6}}\frac{\phi}{\omega}, \qquad k \equiv \sqrt{\frac{\lambda}{6}}\frac{j}{\omega^3} \quad (2)$$

leads to the bare action

$$S[\varphi] = \frac{\lambda v^4}{12\omega} \int_0^{\omega\beta} d\tau \int_V d^3x \left((\varphi')^2 + \frac{1}{\omega^2}(\nabla\varphi)^2 + \frac{1}{2}(\varphi^2 - 1)^2 + 2k\varphi \right), \quad (3)$$

where a prime represents a derivative with respect to the dimensionless Euclidean time τ. As shown further in this article, the effective action is convex, and we thus focus on the true vacuum, which occurs for vanishing source $j = 0 = k$. As a consequence, no bubbles of true/false vacuum can form as they would have an infinite radius [18,19]. We are therefore interested in time-dependent instantons only, beyond the static and homogeneous saddle points. The corresponding equation of motion is then

$$\varphi'' - \varphi^3 + \varphi - k = 0, \quad (4)$$

where the solutions to this equation, $\varphi_i(\tau, \mathbf{x})$, are the saddle points of the partition function to be introduced below.

B. Static saddle points

Introducing the critical dimensionless source

$$k_c \equiv 2/(3\sqrt{3}) \quad (5)$$

allows us to distinguish two cases.

For $|k| > k_c$, there is only one static and homogeneous (real) solution to the equation of motion (4), and quantization of the theory can therefore be based on one saddle point only, leading to the usual 1PI effective potential.

For $|k| < k_c$, the regime we focus on, there are two such solutions

$$\begin{aligned} \varphi_L(k) &= \frac{2}{\sqrt{3}}\cos(\pi/3 - (1/3)\arccos(k/k_c)), \\ \varphi_R(k) &= \frac{2}{\sqrt{3}}\cos(\pi - (1/3)\arccos(k/k_c)) \\ &= -\varphi_L(-k). \end{aligned} \quad (6)$$

The actions for these configurations are

$$\begin{aligned} S_L &\equiv S[\varphi_L(k)] = B\omega\beta(4k - k^2 + \mathcal{O}(k^3)), \\ S_R &\equiv S[\varphi_R(k)] = S[\varphi_L(-k)], \end{aligned} \quad (7)$$

where

$$B \equiv \frac{\lambda v^4 V}{24\omega}. \quad (8)$$

C. Instanton/anti-instanton gas

In Euclidean time, and with the absence of a source, the motion described by Eq. (4) corresponds to the motion in real time with the upside-down potential $V(\varphi) \equiv -(\varphi^2 - 1)^2/2$, for which the minimum action S_{inst} is obtained by the known solution

$$\varphi_{\text{inst}}(\tau) = \pm \tanh\left(\frac{\tau - \tau_0}{\sqrt{2}} \right), \quad (9)$$

where $0 \leq \tau_0 \leq \omega\beta$ and

$$S_{\text{inst}} \equiv S[\varphi_{\text{inst}}] = \frac{8\sqrt{2}}{3} B. \tag{10}$$

Because of finite temperature though, field configurations should be periodic in Euclidean time, such that one needs to consider an instanton/anti-instanton pair as the basic building block. For two distant "jumps" at τ_1 and τ_2 such that $|\tau_1 - \tau_2| \gg 1$, the configuration can be approximated by [12]

$$\varphi_{\text{pair}}(\tau) \simeq -\tanh\left(\frac{\tau - \tau_1}{\sqrt{2}}\right) \tanh\left(\frac{\tau - \tau_2}{\sqrt{2}}\right), \tag{11}$$

with an action exponentially close to $2S_{\text{inst}}$.

In the presence of a source, the basic building block is in principle either a bounce or a shot (see [20,21] for reviews). However, since we are interested in the limit of vanishing source and periodic boundary conditions, the fundamental saddle point we consider behaves as the function (11). Assuming the jumps occur over a short time in comparison to β, the instanton/anti-instanton pair spends the same time $\beta/2$ exponentially close to each static saddle point, resulting in an action for such a pair φ_{pair} of

$$S_{\text{pair}} \simeq \frac{1}{2} S_L + \frac{1}{2} S_R + 2S_{\text{inst}}. \tag{12}$$

Revisiting the analogy of classical mechanics in the upside-down potential $V(\varphi)$, the other possible saddle points consist of periodic oscillations made of n instanton/anti-instanton pairs, where the value of n depends on how "exponentially close" the oscillations from a static saddle point begins. An example of an exact saddle point is given in Fig. 1(a). Assuming the total Euclidean time β is large enough to leave the structure of pairs intact, the time spent close to one static saddle point is the same as the time spent close to the other and the total action for n pairs is

$$S_{n\,\text{pairs}} \simeq \frac{1}{2} S_L + \frac{1}{2} S_R + 2nS_{\text{inst}}. \tag{13}$$

The latter "crystalline" structure, with n periodic oscillations, corresponds to an exact solution of the equation of motion. For large β, where the average distance between instantons and anti-instantons remains large compared to their width, a translation of each jump leaves the action S_{inst} invariant and the resulting highly degenerate "gas" of instanton/anti-instanton pairs dominates the partition function. An example of an approximate saddle point is given in Fig. 1(b). In this "dilute gas" approximation, the n instanton/anti-instanton pair configurations spend on average an equal time $\beta/2$ close to each static saddle point, with the same total action (13) as for an exact n-pair configuration as a result of the translational invariance of jumps.

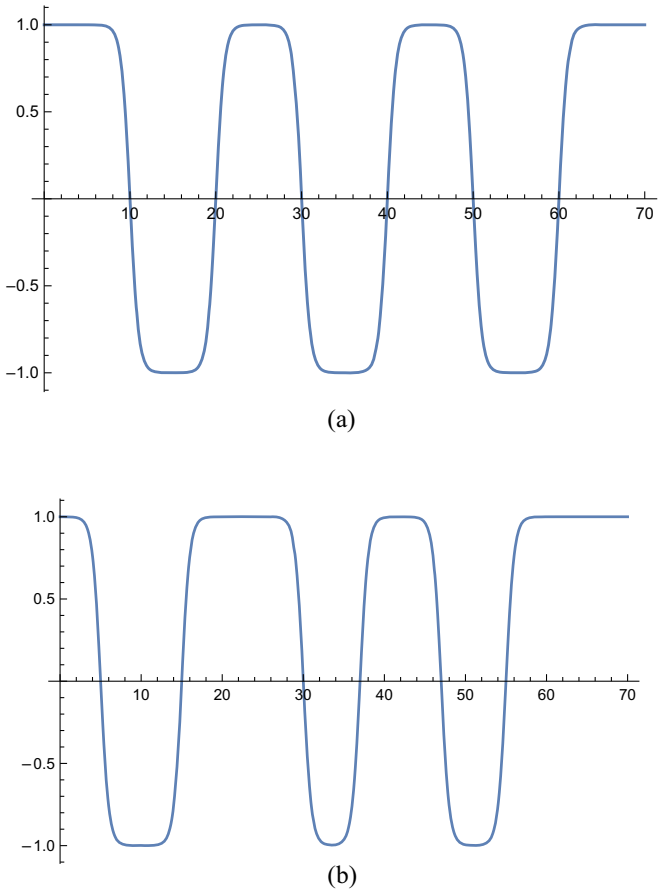

FIG. 1. Examples of exact and approximate saddle points. In the dilute gas approximation, the difference between the corresponding actions is of order $B\omega\beta\exp(-\omega\beta) \ll 1$, and the partition function is dominated by the whole set of approximate saddle points. (a) An exact saddle point conguration with 3 instanton/anti-instanton pairs and action S_3 pairs: the oscillations are periodic. (b) An approximate saddle point conguration with 3 instanton/anti-instanton pairs: the jumps are randomly distributed, but the average distance between them is larger than their width, such that they keep their shape and the action of the conguration is also S_3 pairs.

D. Partition function

The partition function is evaluated in the semiclassical approximation via a sum over the two static saddle points, φ_L and φ_R, and the dilute gas of n instanton/anti-instanton pairs for all possible values of n. Together, with the corresponding one-loop fluctuation factors $F_{L,R}$ and F_n, the semiclassical approximation of the partition function reads

$$Z[k] \simeq F_L(\beta)\exp(-S_L) + F_R(\beta)\exp(-S_R) + \sum_{n=1}^{\infty}\left(\prod_{i=1}^{2n}\int_{\tau_{i-1}}^{\omega_R\beta}\tau_i\right)F_n\exp(-S_{n\,\text{pairs}}). \tag{14}$$

In the latter expression, the product of integrals over the times τ_i where the jumps occur corresponds to the zero-mode of the

fluctuation factor for the saddle point made of n instanton/anti-instanton pairs. Indeed, the translational invariance of the action means that the ith jump can happen at any time $\tau_i \in [\tau_{i-1}, \beta]$. For finite temperature though, there is a maximum number of instanton/anti-instanton pairs; however, the error made in the summation for $n \to \infty$ is negligible since each term is suppressed by $\exp(-nS_{\text{inst}})$. With the fluctuation factors derived in Appendix A and Appendix B, we can write

$$Z[k] = \exp(-\Sigma_L(\beta)) + \exp(-\Sigma_R(\beta)) + \exp(-\Sigma_{\text{gas}}(\beta)), \tag{15}$$

where Σ_L, Σ_R, and Σ_{gas} are the connected graphs generating functionals for the static saddle points and the gas of instanton/anti-instanton pairs, respectively. We note that an instanton or anti-instanton does not lead to any imaginary part in the partition function, unlike a bounce, since the former are monotonous functions of the Euclidean time, such that the fluctuation operator does not have negative eigenvalues [22].

1. Static saddle points

One-loop quantum corrections can be split into two contributions: the zero-temperature corrections, containing all the divergences, and the divergence free finite-temperature dependent corrections. The zero-temperature contribution is calculated in [11] and is expressed in terms of the renormalized parameters. It is mentioned here that, in the case of several saddle points and in order to avoid confusion between loop orders, renormalization should be done at the level of the individual connected graphs generating functionals before performing the Legendre transform. The finite-temperature contribution can be calculated using the Schwinger proper time representation—see Appendix A—and the overall contribution is

$$\begin{aligned}\Sigma_{L,R}(\beta) = B_r\omega_r\beta\Bigg(&(\varphi_{L,R}^2-1)^2 + 4k\varphi_{L,R} \\ &+ \frac{\lambda_r}{96\pi^2}(3\varphi_{L,R}^2-1)^2\ln\left(\frac{3}{2}\varphi_{L,R}^2-\frac{1}{2}\right) \\ &- \frac{\lambda_r(3\varphi_{L,R}^2-1)}{3\pi^2}\sum_{l=1}^{\infty}\frac{K_2\left(l\omega_r\beta\sqrt{3\varphi_{L,R}^2-1}\right)}{(l\omega_r\beta)^2}\Bigg).\end{aligned} \tag{16}$$

In the previous expression, the renormalized parameters are

$$\begin{aligned}\lambda_r &\equiv \lambda - \frac{3\lambda^2}{32\pi^2}\log\left(\frac{\Lambda^2}{\lambda v^2}\right), \\ v_r^2 &\equiv v^2 - \frac{3\Lambda^2}{16\pi^2} + \frac{\lambda v^2}{16\pi^2}\log\left(\frac{\Lambda^2}{\lambda v^2}\right), \\ B_r &\equiv \frac{\lambda_r v_r^4 V}{24\omega_r}, \\ \omega_r &\equiv v_r\sqrt{\frac{\lambda_r}{6}},\end{aligned} \tag{17}$$

and $K_2(z)$ is a modified Bessel function of the second kind with asymptotic behavior

$$K_2(z\to\infty) \simeq e^{-z}\sqrt{\frac{\pi}{2z}}. \tag{18}$$

We note that l does not correspond to Matsubara modes. Also, the temperature-independent part of the expression (16) reproduces the zero-temperature result derived in [11].

2. Gas of instanton/anti-instanton pairs

The evaluation of Σ_{gas} involves the fluctuation factor above each jump and includes a summation over the allowed jump positions in the interval $\tau \in [0, \beta]$ [22]. The additional contribution of quantum fluctuations arises from the "flat" parts of the instanton/anti-instanton configurations, which are exponentially close to each static saddle point for the approximate average time of $\beta/2$ when neglecting the width of each jump compared to β. Performing the resummation over instantons/anti-instantons, we show in Appendix B that the corresponding connected graphs generating functional is then

$$\Sigma_{\text{gas}}(\beta) \simeq \Sigma_L(\beta/2) + \Sigma_R(\beta/2) - \ln(\cosh(\bar{N}) - 1), \tag{19}$$

where

$$\bar{N} \equiv \omega_r\beta\sqrt{\frac{6}{\pi}S_{\text{inst}}}e^{-S_{\text{inst}}}, \tag{20}$$

corresponding to the average number of instanton/anti-instanton pairs at temperature $T = 1/\beta$. In this article we are interested in the limit $\omega_r\beta \gg 1$ for a fixed volume—and thus fixed action S_{inst}—such that we consider the situation where $\bar{N} \gg 1$, corresponding to the full tunneling regime. In the situation where β is fixed and V becomes large we have $\bar{N} \ll 1$, where tunneling is suppressed and the system is better approximated by SSB [12].

III. NONEXTENSIVE GROUND SATE

A. One-particle-irreducible effective action

From the partition function evaluated for a constant source j, the classical field is obtained as

$$\phi_c \equiv -\frac{1}{Z}\frac{\delta Z}{\delta j} \to -\frac{1}{V\beta Z}\frac{\partial Z}{\partial j}, \tag{21}$$

which, in terms of the dimensionless quantities previously introduced, can be written as

$$\varphi_c = -\frac{1}{4B_r\omega_r\beta Z}\frac{\partial Z[k]}{\partial k}. \tag{22}$$

From the expression (15) for the partition function, together with the expressions (16) and (19), the classical field is expanded in powers of the source k,

$$\varphi_c = \left(-f_0 + \frac{\lambda_r}{128\pi^2} f_1\right) k + \mathcal{O}(k^3), \tag{23}$$

where

$$f_0 \equiv \frac{1 + 16B_r\omega_r\beta + \cosh(\bar{N})}{2(1+\cosh(\bar{N}))},$$
$$f_1 \equiv \frac{7 + 32B_r\omega_r\beta + 7\cosh(\bar{N})}{1+\cosh(\bar{N})}. \tag{24}$$

Consistently with the symmetry of the bare potential, the classical field ϕ_c is an odd function of k: the even powers of k cancel out in the expression for φ_c after adding the contribution of the different saddle points, leading to the mapping $k = 0 \Leftrightarrow \varphi_c = 0$.

We then perform the Legendre transform, after expressing the source as a function of the classical field

$$k(\varphi_c) = -\frac{1}{2}\left(g_0 + \frac{\lambda}{16\pi^2} g_1\right)\varphi_c + \mathcal{O}(\varphi_c^3), \tag{25}$$

where

$$g_0 \equiv \frac{4(1+\cosh(\bar{N}))}{1 + 16B_r\omega_r\beta + \cosh(\bar{N})},$$
$$g_1 \equiv \frac{(1+\cosh(\bar{N}))(7 + 32B_r\omega_r\beta + 7\cosh(\bar{N}))}{(1 + 16B_r\omega_r\beta + \cosh(\bar{N}))^2}. \tag{26}$$

The effective action for a constant configuration is finally

$$\Gamma(\varphi_c) = -\ln Z(k(\varphi_c)) - 4B_r\omega_r\beta \int k(\varphi_c)d\varphi_c$$
$$= \Gamma(0) + B_r\omega_r\beta\left(g_0 + \frac{\lambda}{16\pi^2} g_1\right)\varphi_c^2 + \mathcal{O}(\varphi_c^4), \tag{27}$$

where

$$\Gamma(0) = -\ln Z(0)$$
$$= -\ln\left(2e^{-\Sigma_0(\beta)} + e^{-2\Sigma_0(\beta/2)}(\cosh(\bar{N}) - 1)\right), \tag{28}$$

and $\Sigma_0 \equiv \Sigma_L|_{k=0} = \Sigma_R|_{k=0}$. The effective potential U_{eff} is finally given by

$$\Gamma(\phi_c) = V\beta U_{\text{eff}}(\phi_c), \tag{29}$$

and, as expected, it satisfies the following properties:

(i) it is a convex function of ϕ_c, since the mass term is positive;
(ii) the ground state is at $\varphi_c = 0$, or equivalently $k = 0$;
(iii) it has a nontrivial volume-dependence and is thus nonextensive.

For the following studies of NEC violation we focus on the ground state $\varphi_c = 0$.

B. NEC violation

The ground state density ρ and pressure p are obtained from the free energy

$$F = \frac{1}{\beta}\Gamma(0) = -\frac{1}{\beta}\ln Z(0), \tag{30}$$

and their sum can be written as [12]

$$\rho + p = \frac{1}{V}\left(F - T\frac{\partial F}{\partial T}\right) - \frac{\partial F}{\partial V}$$
$$= -T\frac{\partial U_{\text{eff}}(0)}{\partial T} - V\frac{\partial U_{\text{eff}}(0)}{\partial V}. \tag{31}$$

From the expression (28), we obtain for $\omega_r\beta \gg 1$

$$\rho + p \simeq \frac{4\omega_R^{5/2}}{(\sqrt{2}\pi\beta)^{3/2}} e^{-\omega_R\beta/\sqrt{2}} - \frac{\omega_R}{V}\left(S_{\text{inst}} + \frac{1}{2}\right)\sqrt{\frac{6}{\pi}S_{\text{inst}}}\, e^{-S_{\text{inst}}}. \tag{32}$$

On the right-hand side, the first term corresponds to thermal fluctuations and the second term corresponds to tunneling. These terms compete for the overall sign of $\rho + p$ leading to the following cases:

(i) Infinite volume: $\rho + p \geq 0$

In the limit of infinite volume tunneling is suppressed, as seen via the vanishing of the average number (20) of instanton/anti-instanton pairs for any fixed temperature: $\lim_{V\to\infty} \bar{N} = 0$ for fixed β. Hence only thermal fluctuations contribute and

$$\rho + p = \frac{4\omega_R^{5/2}}{(\sqrt{2}\pi\beta)^{3/2}} e^{-\omega_R\beta/\sqrt{2}}, \tag{33}$$

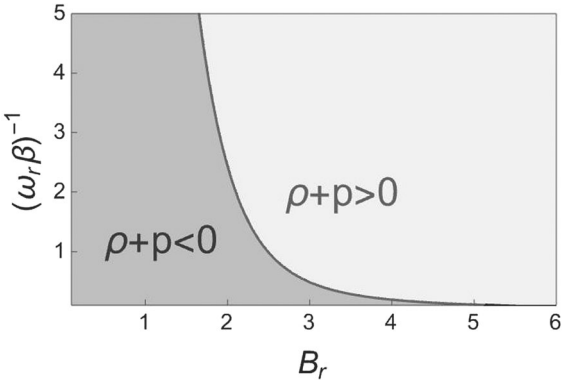

FIG. 2. The boundary between the regions where the NEC is satisfied and where it is violated due to the competition of tunneling and thermal fluctuations. The plot shows the curve $V(T)$ in terms of the dimensionless variables used in this article.

with $\rho + p \to 0$ as the temperature goes to 0 or equivalently $\beta \to \infty$. This result is not surprising: the limit of infinite volume corresponds to SSB and, as expected, the NEC is satisfied.

(ii) Finite volume and zero temperature: $\rho + p < 0$

In this situation, only the tunneling term contributes and

$$\rho + p = -\frac{\omega_R}{V}\left(S_{\text{inst}} + \frac{1}{2}\right)\sqrt{\frac{6}{\pi}S_{\text{inst}}}e^{-S_{\text{inst}}}. \quad (34)$$

The NEC is violated as a consequence of the explicit volume dependence of the effective potential.

(iii) Boundary $\rho + p = 0$

We sketch in Fig. 2 the boundary $V(T)$ between the region where the NEC is satisfied and the region where the NEC is violated.

Finally, we note that NEC violation is suppressed exponentially with the volume, unlike the power-law suppression which is found with $O(4)$-symmetric Euclidean spacetime coordinates [11,23,24].

IV. DISCRETE MOMENTUM CORRECTIONS

We focus here on the ground state obtained for $k = 0$ in the case of zero temperature, where NEC violation arising from tunneling is maximum.

The previous sections ignore quantization of momentum when calculating the connected graphs generating functional for each static saddle point in a finite volume. As we explain below, the evaluation of $\Sigma_{L,R}$ with discrete momentum consists of taking into account the relevant Casimir energy. There is no such contribution from the jumps in the instantons/anti-instantons since the corresponding one-loop corrections do not depend on momentum.

A. Vacuum energy

The Casimir contribution to the connected graphs generating functional is defined as

$$\Sigma_{\text{Cas}} \equiv \Sigma_{L,R}|_{\text{discrete}} - \Sigma_{L,R}|_{\text{continuum}}, \quad (35)$$

where the ultraviolet divergences cancel out since they are identical in the discrete and continuum cases. For zero temperature and vanishing source, the expression (16) gives

$$\Sigma_{L,R}(k = 0, T = 0)|_{\text{continuum}} = \lim_{\beta\to\infty}\Sigma_0(\beta) = 0, \quad (36)$$

such that, instead of Eq. (28), one-loop corrections obtained with discrete momentum lead to

$$\begin{aligned}\Gamma(0) &= -\ln\left(2e^{-\Sigma_{\text{Cas}}} + e^{-\Sigma_{\text{Cas}}}(\cosh(\bar{N}) - 1)\right)\\ &= \Sigma_{\text{Cas}} - \ln(\cosh(\bar{N}) + 1). \end{aligned} \quad (37)$$

The above expression takes advantage of the proportionality between Σ_0 and β in the limit of vanishing temperature, such that

$$2\Sigma_0(\beta/2) \to \Sigma_0(\beta), \quad (38)$$

as $\beta \to \infty$. In the situation of one saddle point, and therefore no tunneling, $\Gamma(0) = \Sigma_{\text{Cas}} = \beta E_{\text{Cas}}$ where E_{Cas} is the Casimir energy corresponding to quantum fluctuations about a single vacua $\pm v$ (where one has approximately quadratic fluctuations with mass $m = \sqrt{2}\omega_r$). Hence

$$U_{\text{eff}}(0) = \frac{E_{\text{Cas}}}{V} - \frac{1}{V\beta}\ln(\cosh(\bar{N}) + 1), \quad (39)$$

and we see the additive nature of the Casimir effect and tunneling contributions, similar to the finite-temperature contribution. The sum of density and pressure reads finally

$$\rho + p = \frac{E_{\text{Cas}}}{V} - \frac{\partial E_{\text{Cas}}}{\partial V} - \frac{\omega_R}{V}\left(S_{\text{inst}} + \frac{1}{2}\right)\sqrt{\frac{6S_{\text{inst}}}{\pi}}e^{-S_{\text{inst}}}. \quad (40)$$

B. Casimir contribution to the NEC

The Casimir energy is highly sensitive to the geometry of the box containing the field, as well as the boundary conditions used on the corresponding surfaces [17]. For a scalar field $\varphi(t, x)$ in the interval $x \in [0, L]$, for example, the possible choices of boundary conditions are defined as follows:

$$\text{Dirichlet}: \ \varphi(t, 0) = \varphi(t, L) = 0, \quad (41)$$

$$\text{Neumann}:\ \partial_x\varphi(t,0)=\partial_x\varphi(t,L)=0,$$
$$\text{Periodic}:\ \varphi(t,0)=\varphi(t,L). \tag{42}$$

For the cases we consider, the asymptotic form of the Casimir effect is identical for both Dirichlet and Neumann boundary conditions. We thus consider mixed boundary conditions, where different subsets of the boundary can possess either Dirichlet or Neumann conditions. For the case of mixed boundary conditions, the Casimir energy is dependent on the size/curvature of the material boundaries, and for the case of periodic boundary conditions, it is dependent on the period length/curvature of the nontrivial spacetime. A "general rule" states that flat geometries lead to exponential suppression of the Casimir energy for $mL \gg 1$, where L is the length scale of the relevant boundaries, and that curved geometries lead to power-law suppression of the Casimir energy for $mR \gg 1$, where R is the radius of curvature of the relevant surfaces. There are exceptions to this general rule though, which are highlighted in the following examples.

(i) Dirichlet boundary conditions, flat boundaries

The original Casimir configuration consists of a scalar field constrained between two parallel, flat mirrors with surface area A and separation a, with the scalar field satisfying Dirichlet conditions on the boundaries. The corresponding Casimir energy is [17]

$$E_{\text{Cas}} \simeq \begin{cases} -\frac{A\pi^2}{1440a^3} & \text{for } am \ll 1 \\ -\frac{A}{8\sqrt{2}}\left(\frac{m}{\pi a}\right)^{3/2} e^{-2ma} & \text{for } am \gg 1 \end{cases} \tag{43}$$

and is always negative.

(ii) Dirichlet boundary conditions, curved boundaries

For dimensional reasons, the Casimir energy for a scalar field confined within the curved boundary of a 2-sphere of radius R with Dirichlet boundary conditions is given in terms of the dimensionless function

$$E_{\text{Cas}} = \frac{1}{R} f(mR) \tag{44}$$

and is found to obey power-law suppression in mR, for $mR \gg 1$ [25].

(iii) Periodic boundary conditions, flat spacetime

For a scalar field confined to the surface of a 3-torus (a rectangular box with periodic boundary conditions), the sign of the Casimir energy depends on the ratio of the lengths of the box and we have [26]

$$E_{\text{Cas}} \simeq -\frac{(mL)^{3/2}}{L}\exp(-mL) \quad \text{for } mL \gg 1, \tag{45}$$

where L is the typical size of the period length.

(iv) Periodic boundary conditions, curved spacetime

For a scalar field confined to the surface of a 3-sphere with radius R, we would expect the asymptotic form to be a power law in R. However, this special case is an exception to the general rule as a consequence of the accidental vanishing of the heat-kernel coefficients (see Sec. 3 of [17] for details). The resulting Casimir energy has instead an exponential asymptotic form, as in the case of flat geometries [27]

$$E_{\text{Cas}} \simeq +\frac{(mR)^{5/2}}{R}\exp(-2\pi mR) \quad \text{for } mR \gg 1. \tag{46}$$

The above examples display how the Casimir effect for a massive scalar field is at most suppressed by the exponential e^{-mL}, where L is a typical size of the boundary containing the field. On the other hand, the tunneling contribution to the NEC, calculated with continuous momentum, is proportional to

$$e^{-S_{\text{inst}}} \sim \exp\left(-\frac{(mL)^3}{\lambda}\right), \tag{47}$$

and is therefore negligible compared to the Casimir contribution in the regime $mL \gg \sqrt{\lambda}$. For $mL \sim \sqrt{\lambda}$ though, tunneling competes with the Casimir effect and can change the sign of $\rho + p$ in the situation where the Casimir energy is positive. As an example, we sketch in Fig. 3 the boundary $R(\lambda)$ between the region where the NEC is satisfied and the region where it is violated, due to the competition between tunneling and the Casimir effect on a 3-sphere.

We note, however, two important points regarding the Casimir examples cited here: (i) they are valid for ideal surfaces only, and a realistic confining mechanism for the

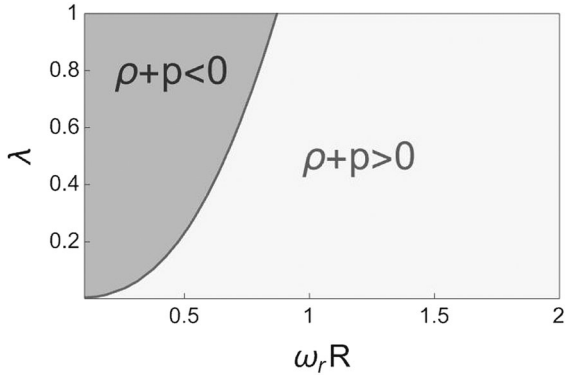

FIG. 3. The boundary between the regions where the NEC is satisfied and where it is violated, due to the competition of the Casimir energy and tunneling at zero temperature on a 3-sphere. The plot shows the curve $R(\lambda)$ in terms of the dimensionless variables used in this article.

scalar field would lead to a modification of the Casimir vacuum energies, especially if the field is confined by an external potential instead of a physical box [28]. (ii) they assume free scalar fields and ignore its self-interactions. On the other hand, the tunneling mechanism described here as follows: (i) necessitates the field to be self-interacting; and (ii) is not sensitive to the geometry/topology of the box containing the field. Hence the conclusions regarding which effect dominates could be modified by a more thorough study, depending on the situation which is considered.

Finally, the average null energy condition is not violated by the present mechanism. Indeed, if we take into account the energy necessary to maintain the confining mechanism, the overall ground state of the system does not violate the NEC [29], consistent with what is expected from causality [30].

V. CONCLUSIONS

Tunneling between degenerate vacua is exponentially suppressed with the volume of the box containing the field, but nevertheless allows the possibility of NEC violation at low temperatures. Taking into account discrete momentum of fluctuations in a finite volume implies this effect is mainly relevant for situations where the typical size of the box is not too large compared to the Compton wavelength of the particle, and where tunneling can lead to an overall NEC violation. A potential application lies in axion physics, where the de Broglie wavelength can be of order 1 kpc [31] with the confinement provided by a gravitational well.

Exponential suppression in the volume could potentially be avoided by a consideration of nondegenerate vacua, where other saddle points with a volume-independent action become relevant, as in the original study of false vacuum decay [18,19]. The resulting effective action would be nonextensive in a certain regime of the classical field, but more studies need to be done for the status of NEC violation in the corresponding vacuum.

Finally, NEC violation could play an important role in early universe cosmology, where tunneling could provide a dynamical mechanism for a cosmological bounce, as explained in [23,24]: as the Universe contracts, tunneling switches on and violates the NEC, which induces a bounce after which tunneling is suppressed as the Universe expands. This scenario necessitates the study of tunneling in a Friedman-Lemaître-Robertson-Walker background though, and is left for future work.

ACKNOWLEDGMENTS

The authors thank Klaus Kirsten for valuable correspondence regarding the Casimir effect, and J. A. thanks Janos Polonyi for enlightening discussions. This work is supported by the Leverhulme Trust (Grant No. RPG-2021-299) and the Science and Technology Facilities Council (Grant No. STFC-ST/T000759/1). For the purpose of Open Access, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

APPENDIX A: FLUCTUATION FACTOR FOR A STATIC SADDLE POINT

The fluctuation factors for the static saddle points are calculated with continuous three-dimensional momenta, introducing the cutoff Λ in the Schwinger proper time representation of the propagator. Introducing the dimensionless Matsubara frequency $\nu_n \equiv 2\pi n/\omega\beta$, we have

$$\begin{aligned}
&\mathrm{Tr}\{\ln(\delta^2 S[\varphi_i])\} \\
&= V\int\frac{d^3p}{(2\pi)^3}\sum_{n=-\infty}^{\infty}\int_{1/\Lambda^2}^{\infty}\frac{\mathrm{d}s}{s}\mathrm{e}^{-4B\omega\beta s(p^2/\omega^2+\nu_n^2+3\varphi_i-1)} \\
&= \frac{V\omega^3}{2\pi^2}\sum_{n=-\infty}^{\infty}\int_{1/X^2}^{\infty}\frac{\mathrm{d}x}{x}\int_0^{\infty}\mathrm{d}qq^2\mathrm{e}^{-x(q^2+\nu_n^2+3\varphi_i-1)} \\
&= \frac{V\omega^3}{8\pi^{3/2}}\sum_{n=-\infty}^{\infty}\int_{1/X^2}^{\infty}\frac{\mathrm{d}x}{x^{5/2}}\mathrm{e}^{-x(\nu_n^2+3\varphi_i-1)} \\
&= \frac{V\omega^3}{8\pi^{3/2}}\int_{1/X^2}^{\infty}\frac{\mathrm{d}x}{x^{5/2}}\mathrm{e}^{-x(3\varphi_i-1)}\vartheta_0\left(\frac{4\pi x}{\omega^2\beta^2}\right),
\end{aligned} \tag{A1}$$

where the dimensionless variables are

$$q \equiv \frac{p}{\omega}, \qquad x \equiv 4B\omega\beta s, \qquad X^2 \equiv \frac{\Lambda^2}{4B\omega\beta}, \tag{A2}$$

and $\vartheta_0(y)$ is the Jacobi function

$$\vartheta_0(y) \equiv \sum_{n=-\infty}^{\infty}\mathrm{e}^{-\pi y n^2}. \tag{A3}$$

Making use of the following property:

$$\vartheta_0(y) = y^{-1/2}\vartheta_0(1/y), \tag{A4}$$

the above becomes

$$\begin{aligned}
&\mathrm{Tr}\{\ln(\delta^2 S[\varphi_i])\} \\
&= \frac{V\omega^4\beta}{16\pi^2}\int_{1/X^2}^{\infty}\frac{\mathrm{d}x}{x^3}\mathrm{e}^{-x(3\varphi_i-1)}\vartheta_0\left(\frac{\omega\beta}{4\pi x}\right) \\
&= \frac{V\omega^4\beta}{16\pi^2}\int_{1/X^2}^{\infty}\frac{\mathrm{d}x}{x^3}\mathrm{e}^{-x(3\varphi_i-1)}\sum_{n=-\infty}^{\infty}e^{-\omega^2\beta^2n^2/4x} \\
&= \lambda\frac{B\omega\beta}{24\pi^2}(I_\Lambda(\varphi_i)+I_T(\varphi_i)),
\end{aligned} \tag{A5}$$

where

$$\begin{aligned}
I_\Lambda(\varphi_i) &\equiv \int_{1/\Lambda^2}^{\infty}\frac{\mathrm{d}x}{x^3}\mathrm{e}^{-x(3\varphi_i-1)}, \\
I_T(\varphi_i) &\equiv 2\sum_{n=1}^{\infty}\int_0^{\infty}\frac{\mathrm{d}x}{x^3}\mathrm{e}^{-x(3\varphi_i-1)-\omega^2\beta^2n^2/4x}.
\end{aligned} \tag{A6}$$

The first integral I_Λ is the temperature-independent divergent integral which, after renormalization, produces the same results as in the zero-temperature case [11]. The second integral I_T is the temperature-dependent contribution corresponding to the finite-temperature corrections. It is finite, which is why the cutoff is taken to infinity in this specific term. This temperature-dependent integral can be written in terms of the modified Bessel functions of the second kind $K_2(z)$ as

$$I_T(\phi_i) = \sum_{n=1}^{\infty} \frac{16(3\varphi_i - 1)}{(n\omega\beta)^2} K_2(n\omega\beta\sqrt{3\varphi_i - 1}). \quad \text{(A7)}$$

Together with the integral I_Λ, the connected graphs generating functional for homogeneous saddle points is given by Eq. (16).

APPENDIX B: FLUCTUATION FACTOR FOR THE INSTANTONS/ANTI-INSTANTONS GAS

We calculate here the contribution $\exp(-\Sigma_{\text{gas}})$ to the partition function (15), following the known approach in studies of tunneling effects [22].

The invariance of the action for n instanton/anti-instanton pairs under the translation of the jumps leads to the degeneracy factor in the partition function

$$\left(\prod_{i=1}^{2n} \int_{\tau_{i-1}}^{\omega\beta} \tau_i\right) = \frac{(\omega\beta)^{2n}}{(2n)!}, \quad \text{(B1)}$$

where $\tau_i \in [\tau_{i-1}, \omega\beta]$ and $\tau_0 = 0$, since successive instanton jumps can only occur after previous ones. Each jump has an associated fluctuation factor $\sqrt{6S_{\text{inst}}/\pi}$, and thus the total fluctuation factor is given by the product of the contributions of the "flat" parts of the n-pairs of instanton/anti-instantons and the n pairs of jumps. On average, each configuration of n instanton/anti-instanton pairs spends the same time $\simeq\beta/2$ close to each static saddle point, such that the expression for F_n is finally

$$F_n = F_L(\beta/2)F_R(\beta/2)\left(\frac{6S_{\text{int}}}{\pi}\right)^n. \quad \text{(B2)}$$

Substituting the above results into the partition function (14), along with the total action (13) for n pairs, yields the total contribution to the partition function due to instanton/anti-instanton pairs

$$\begin{aligned}\exp(-\Sigma_{\text{gas}}) &= e^{-\Sigma_L[\beta/2]}e^{-\Sigma_R[\beta/2]}\sum_{n=1}^{\infty}\frac{(\omega\beta)^{2n}}{(2n)!}\left(\frac{6S_{\text{int}}}{\pi}\right)^n e^{-2nS_{\text{int}}} \\ &= \exp(-\Sigma_L[\beta/2] - \Sigma_R[\beta/2]) \\ &\quad \times \left(\cosh\left(\omega\beta\sqrt{\frac{6S_{\text{int}}}{\pi}}e^{-S_{\text{inst}}}\right) - 1\right), \quad \text{(B3)}\end{aligned}$$

This leads to the expression (19), where the parameters can be replaced by their renormalized version, since the overall expression is already at one-loop.

[1] K. Symanzik, Commun. Math. Phys. **16**, 48 (1970).
[2] S. R. Coleman, R. Jackiw, and H. D. Politzer, Phys. Rev. D **10**, 2491 (1974).
[3] J. Iliopoulos, C. Itzykson, and A. Martin, Rev. Mod. Phys. **47**, 165 (1975).
[4] R. W. Haymaker and J. Perez-Mercader, Phys. Rev. D **27**, 1948 (1983).
[5] Y. Fujimoto, L. O'Raifeartaigh, and G. Parravicini, Nucl. Phys. **B212**, 268 (1983).
[6] C. M. Bender and F. Cooper, Nucl. Phys. **B224**, 403 (1983).
[7] M. Hindmarsh and D. Johnston, J. Phys. A **19**, 141 (1986).
[8] A. D. Plascencia and C. Tamarit, J. High Energy Phys. 10 (2016) 099.
[9] P. Millington and P. M. Saffin, J. Phys. A **52**, 405401 (2019).
[10] J. Alexandre and A. Tsapalis, Phys. Rev. D **87**, 025028 (2013).
[11] J. Alexandre and D. Backhouse, Phys. Rev. D **105**, 105018 (2022).
[12] J. Alexandre and J. Polonyi, Phys. Rev. D **106**, 065008 (2022).
[13] V. A. Rubakov, Phys. Usp. **57**, 128 (2014).
[14] E.-A. Kontou and K. Sanders, Classical Quantum Gravity **37**, 193001 (2020).
[15] T. W. B. Kibble, J. Phys. A **9**, 1387 (1976).
[16] W. H. Zurek, Nature (London) **317**, 505 (1985).
[17] M. Bordag, U. Mohideen, and V. M. Mostepanenko, Phys. Rep. **353**, 1 (2001).
[18] S. R. Coleman, Phys. Rev. D **15**, 2929 (1977); **16**, 1248(E) (1977).
[19] C. G. Callan, Jr. and S. R. Coleman, Phys. Rev. D **16**, 1762 (1977).
[20] A. Andreassen, D. Farhi, W. Frost, and M. D. Schwartz, Phys. Rev. D **95**, 085011 (2017).
[21] W.-Y. Ai, B. Garbrecht, and C. Tamarit, J. High Energy Phys. 12 (2019) 095.
[22] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets* (World Scientific, 2004), 10.1142/5057.
[23] J. Alexandre and K. Clough, Phys. Rev. D **100**, 103522 (2019).

[24] J. Alexandre and J. Polonyi, Phys. Rev. D **103**, 105020 (2021).
[25] K. Kirsten, AIP Conf. Proc. **484**, 106 (1999).
[26] S. C. Lim and L. P. Teo, Ann. Phys. (Amsterdam) **324**, 1676 (2009).
[27] S. G. Mamaev, V. M. Mostepanenko, and A. A. Starobinsky, Zh. Eksp. Teor. Fiz. **70**, 1577 (1976).
[28] N. Graham and K. D. Olum, Phys. Rev. D **67**, 085014 (2003); **69**, 109901(E) (2004).
[29] V. Sopova and L. H. Ford, Phys. Rev. D **66**, 045026 (2002).
[30] T. Hartman, S. Kundu, and A. Tajdini, J. High Energy Phys. 07 (2017) 066.
[31] D. J. E. Marsh, Phys. Rep. **643**, 1 (2016).