Quantum dynamics of Kaluza-Klein theories

Thomas Appelquist and Alan Chodos
J. W. Gibbs Laboratory, Yale University, New Haven, Connecticut 06511
(Received 8 April 1983)

Some of the quantum properties of Kaluza-Klein theories are studied. The classical features of these theories are reviewed, and the quantization of the gravitational field in an arbitrary number of dimensions is described. These results are then applied to a detailed analysis of the five-dimensional Kaluza-Klein model. The fifth dimension is taken to be compact and a quantum effective potential, as a function of the five-five component of the metric, is constructed. It is argued that the one-loop computation is reliable as long as the distance around the fifth dimension is large compared to the Planck length. The effective potential separates into two pieces: an induced cosmological constant, independent of the size of the fifth dimension, and a distance-dependent "Casimir" energy. The cosmological term is subtracted, leaving an attractive Casimir potential which will contract the fifth dimension to a size on the order of the Planck length. Consequences of this result are discussed and some of the ways in which it can be generalized are outlined.

I. INTRODUCTION

One of the most intriguing and elegant ways of unifying gauge field theories with gravitation is also one of the oldest. The suggestion of Kaluza[1] and Klein[2] was that electromagnetism and general relativity could be unified by starting with a five-dimensional version of the latter and then somehow arranging for the fifth dimension to become unobservable. In recent years, the Kaluza-Klein idea has been generalized to higher dimensions in an effort to unify non-Abelian gauge fields with gravitation.[3,4] The possibility of obtaining a realistic four-dimensional theory by starting with a simpler theory in a higher-dimensional space has been actively pursued in the context of supergravity. The construction of the $N=8$ theory by starting with $N=1$ supergravity in 11 dimensions[5] is especially attractive.

We shall take the point of view that any implementation of the Kaluza-Klein idea should regard the extra dimensions as actually existing with some physical size, rather than as only an intermediate device for deriving interesting four-dimensional theories. If so, then these extra dimensions presumably form some compact manifold whose size is extremely small, perhaps not much larger than the Planck length $(\hbar G/c)^{1/2}=1.6\times 10^{-33}$ cm.

In a recent letter,[6] we suggested that the vacuum fluctuations of the higher-dimensional gravitational field might provide a physical mechanism capable of accounting for the extreme smallness of the extra dimensions. An explicit computation was carried out in the five-dimensional prototypical Kaluza-Klein model which showed that these quantum fluctuations give rise to a gravitational Casimir effect[7] which can contract the fifth dimension to a size on the order of the Planck length. It was assumed that the relevant vacuum solution to the field equation of five-dimensional general relativity is the Cartesian product of flat four-space and a compact fifth dimension with associated metric component g_{55}. The vacuum fluctuations about this classical background are then computed in the form of a quantum effective potential as a function of g_{55}. When the distance around the fifth dimension is larger than the Planck length, the loop expansion is reliable and the potential is seen to give rise to an attractive Casimir force.

The purpose of this paper is to describe the quantum physics of the five-dimensional model in more detail and to begin to discuss some of its generalizations. The most important of these is the extension to higher numbers of dimensions in order to accommodate non-Abelian gauge fields. In this paper, our remarks along these lines will be restricted to the classical theory. Some of these have appeared already in the literature[3,4] but we include them here in order to prepare for a study of the quantum properties of these theories.

The effective dimensional reduction which takes place in a Kaluza-Klein theory is analogous to the effective reduction of a high-temperature four-dimensional gauge theory to three dimensions.[8] Such theories have a periodicity in the time coordinate inversely proportional to the temperature. For temperatures much larger than the momentum scales being considered, the time dimension is squeezed out of the problem, leaving an effective massless three-dimensional theory. This classical picture is then modified by quantum loops composed of the nonzero (massive) modes in the Fourier time expansion. Because of the ultraviolet singularities in certain loop computations, they do not completely decouple from the low-momentum physics. In particular, they produce an effective mass (a Debye mass) for the electric components of the gauge field. Thus quantum effects screen the electric field at large distances, completing the reduction to an effective three-dimensional theory. The effective potential computation in Kaluza-Klein theories is very similar to the Debye-screening computation in finite-temperature gauge theories.

In Sec. II, we discuss the classical features of Kaluza-Klein theories. The five-dimensional model and its reduction to four dimensions is described in some detail. We relate the time-independent solution, which will be used as a starting point for the quantum computations, to a time-dependent classical solution (the Kasner solution) that was considered in Ref. 9. Some of the features of higher-dimensional compactification are then described. Finally, returning to five dimensions, we show that the presence of a cosmological term which is naturally generated by the quantum fluctuations, leads to a five-dimensional de Sitter-type solution. With the assumption that the cosmological constant is sufficiently small, this solution reduces to the above time-independent or Kasner solutions for arbitrarily long times.

In Sec. III, some of the properties of quantum gravity in D dimensions are described. The quantization of the five-dimensional theory is carried out in detail, using a particularly useful and physical gauge choice. The quadratic Lagrangian sufficient to carry out a one-loop computation of the effective potential is constructed, the gauge-fixing and ghost terms are derived, and a convenient method of regularization is described.

In Sec. IV, we present the explicit details of the computation of the one-loop effective potential as a function of g_{55}. In Sec. V, this result is interpreted, the analogy with the Casimir effect is explored, and the validity of the loop expansion is discussed. Finally, for the convenience of the reader, we include an Appendix in which some of the relevant features of finite-temperature gauge theories are summarized.

II. CLASSICAL FEATURES OF KALUZA-KLEIN THEORIES

To begin, we focus on the classical theory of general relativity, not in the usual 3 + 1 dimensions but instead in $D=d+1$ dimensions. The action is

$$S=\frac{1}{16\pi G_D}\int d^Dx\sqrt{\bar{g}}R\ , \tag{2.1}$$

where $\bar{g}_{AB}$ is the metric in D-dimensional space, $1\leq A,B\leq D$, R is the scalar curvature computed from $\bar{g}$, and G_D is the D-dimensional version of Newton's constant. It will be convenient for our purposes to work with the Euclidean version of the theory in which the signature of $\bar{g}_{AB}$ is D. Our units are such that $\bar{g}_{AB}$ is dimensionless; since R contains two derivatives and since S must be dimensionless (we have set $\hbar=1$), it follows that the dimension of G_D is $(\text{length})^{D-2}$. For the time being, we omit a cosmological constant in the action, Eq. (2.1). The possible role of such a constant will be discussed later.

The central idea of the Kaluza-Klein approach is that "dimensional reduction" takes place; i.e., the D-dimensional theory becomes effectively equivalent to a four-dimensional theory, with the extra $D-4$ dimensions being somehow rendered unobservable. In the literature one finds two points of view regarding dimensional reduction. On the one hand, some authors use it as a technical device for obtaining rather complicated Lagrangians in four dimensions from simpler Lagrangians in D dimensions. According to this approach, one merely truncates the D-dimensional theory by assuming, without need for further justification, that either the fields are totally independent of the extra $D-4$ coordinates, or else that they depend on them in some particularly simple way that can be integrated out.[10]

It is tempting, however, to suppose that the extra dimensions really exist even though we cannot detect them. This is possible if they form a compact manifold whose size is very small, perhaps as small as the Planck length $(G_D)^{1/(D-2)}$, the only length scale appearing in the action, Eq. (2.1). There is, of course, nothing in the classical theory which would force the compact dimensions to be related to the Planck length. In the classical theory, G_D is simply an overall factor and is irrelevant. In the quantum theory, we will discover a mechanism which might be responsible for the contraction of the extra dimensions to the order of the Planck length. However, there will still remain the question of why the Planck length, which is very different from all other length scales in physics, arises in the first place.

To see in more detail how dimensional reduction works, it is simplest to begin with the five-dimensional model. One chooses coordinates x^μ for (3 + 1)-dimensional space-time ($\mu=0,1,2,3$) and x^5 for the extra dimension which is assumed to be a circle of radius R_5:

$$0\leq x^5<2\pi R_5\ . \tag{2.2}$$

Here R_5 is a parameter with dimensions of length. We observe that, by itself, R_5 has no physical significance; it merely serves to characterize the range of the coordinate x_5. What can have meaning, however, is the distance around the extra dimension:

$$\delta_5=\int_0^{2\pi R_5}dx^5\sqrt{\bar{g}_{55}}\ . \tag{2.3}$$

[Even this could have no significance if the theory were truly generally covariant in a five-dimensional sense. The field equations derived from Eq. (2.1) will be generally covariant, but the boundary condition we have chosen, namely, that the extra dimension be compact, breaks the five-dimensional general covariance.]

To proceed, we parametrize the metric in the form

$$\bar{g}_{AB}=\phi^{-1/3}\begin{bmatrix} g_{\mu\nu}+A_\mu A_\nu\phi & A_\mu\phi \\ A_\mu\phi & \phi \end{bmatrix} \tag{2.4}$$

and we expand each component in a Fourier series

$$\bar{g}_{AB}(x^\mu,x^5)=\sum_{n=-\infty}^{\infty}\bar{g}_{AB}^{(n)}(x^\mu)e^{inx^5/R_5}\ . \tag{2.5}$$

So far, no approximations have been made. Any five-dimensional metric can be written in the form of Eq. (2.4) and any sufficiently smooth function on a finite interval can be expanded in the form of Eq. (2.5). Complete dimensional reduction is then achieved by keeping only the $n=0$ (x^5-independent) mode of $\bar{g}_{AB}$. Upon so doing, one finds

$$S_4=-\frac{1}{16\pi G}\int d^4x(\det\{g_{\mu\nu}\})^{1/2}\left[R^{(4)}+\tfrac{1}{4}\phi F_{\mu\nu}F^{\mu\nu}+\frac{1}{6}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2}\right]. \tag{2.6}$$

Here $R^{(4)}$ is the four-dimensional scalar curvature computed from $g_{\mu\nu}$, $F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$, and $G=G_5/2\pi R_5$. (The zero-mode superscript on the fields has been dropped.) The role of the Weyl factor $\phi^{-1/3}$ in Eq. (2.4) is to ensure that no extra factors of ϕ multiply $R^{(4)}$ in Eq.

(2.6). Thus the first term in S_4 looks precisely like general relativity. This choice is purely cosmetic, however; Eq. (2.6) really represents a scalar-tensor theory of gravitation, and it is not possible to isolate gravitational effects in $g_{\mu\nu}$ alone or in ϕ alone. Indeed, with $A_\mu=0$, the action S_4 describes a particular version of the well-known Brans-Dicke theory[11] (the Brans-Dicke parameter ω is zero in this case). Had we chosen some other Weyl factor, S_4 would have looked different in detail, but all physical predictions would be the same.

Equation (2.6) also illustrates the simplest instance of the Kaluza-Klein "miracle": the appearance of the standard gauge piece $F_{\mu\nu}F^{\mu\nu}$. It is true that in order to make it take the conventional form, one had to be clever in one's choice of parametrization, Eq. (2.4), but it is not true that the gauge fields were put in by hand; the starting point was pure general relativity in five dimensions.

In order to provide a physical mechanism for dimensional reduction, it will be important to work within the full five-dimensional theory [all terms in the Fourier series Eq. (2.5) must, *a priori*, be considered]. Quantum corrections will be computed as an expansion in $\hbar$ about a "vacuum" classical solution to the five-dimensional field equations derived from the action Eq. (2.1). Using the nomenclature of Eq. (2.4), the classical solution of choice is

$$\begin{aligned} g_{\mu\nu}&=\delta_{\mu\nu}\ , \\ A_\mu&=0\ , \\ \phi&=\phi_c=\text{constant}\ , \end{aligned} \tag{2.7}$$

along with the manifold restriction Eq. (2.2). The distance around the fifth dimension, Eq. (2.3), will be determined by the magnitude of ϕ_c which, in turn, will depend on the quantum corrections. The question of why the five-dimensional theory chooses the classical solution, Eq. (2.7), over some other remains unanswered. Nevertheless, the seemingly trivial fact that it *is* a solution is worth emphasizing since corresponding solutions do not always exist in more realistic, higher-dimensional theories.

The classical solution, Eq. (2.7), is not the only one which might be used as a starting point for the five-dimensional Kaluza-Klein theory. In Ref. 9, it was pointed out that there exists a very simple solution to the five-dimensional Einstein equations (the Kasner solution), in which the metric has a time dependence of an especially appropriate form. The manifold is one in which time ranges over the entire real line and all the spatial coordinates x_i $(i=1,2,3)$, x_5 have a finite range, say $0\le x_i,\ x_5<R$. The metric has the form

$$\begin{aligned} &\phi^{-1/3}g_{00}=1,\ \phi^{-1/3}g_{ij}=(t/t_0)^{2p_1}\delta_{ij}\ , \\ &A_\mu=0\ , \\ &\phi^{2/3}=(t/t_0)^{2p_2}\ , \end{aligned} \tag{2.8}$$

where the sum rules

$$\begin{aligned} &3p_1+p_2=1\ , \\ &3p_1{}^2+p_2{}^2=1 \end{aligned} \tag{2.9}$$

must be obeyed. The interesting case is $p_1=\frac{1}{2}$, $p_2=-\frac{1}{2}$ in which the fifth dimension shrinks and the three other spatial dimensions expand with time. (The case $p_1=0$, $p_2=1$ is simply flat space in disguise.)

The evolution of the metric over cosmological time scales might then explain both the smallness of the fifth dimension and the "largeness" of the visible universe. The time-independent solution, Eq. (2.7), with ϕ_c small, can then be regarded as an approximation to the Kasner solution during the present epoch. We shall see in Sec. IV that this approximation will be adequate for the computation of the quantum effective potential. It will be the quantum effects which finally determine the size of the fifth dimension relative to the Planck length.

We next consider how the classical theory generalizes when $D>5$ in order to expose some of the difficulties which are not apparent from the five-dimensional case and which are often not pointed out in the literature. Nevertheless, since the quantum corrections will be computed only for $D=5$ in the following sections, the reader who is specifically interested in those computations may wish to rejoin the discussion in the paragraph following Eq. (2.20).

For the case $D>5$, it is natural to write

$$\bar g_{AB}=W(\phi)\begin{bmatrix} g_{\mu\nu}+B^a_\mu B^b_\nu\phi_{ab} & B^c_\mu\phi_{ca} \\ B^c_\nu\phi_{cb} & \phi_{ab}\end{bmatrix}\ , \tag{2.10a}$$

with inverse

$$\bar g^{AB}=W^{-1}(\phi)\begin{bmatrix} g^{\mu\nu} & -B^a_\lambda g^{\lambda\mu} \\ -B^b_\lambda g^{\lambda\nu} & \phi^{ab}+B^a_\lambda B^b_\sigma g^{\lambda\sigma}\end{bmatrix}\ . \tag{2.10b}$$

Here $W(\phi)$ is an appropriately chosen Weyl factor, and one expects that $g_{\mu\nu}$ $(\mu,\nu=0,1,2,3)$ will play the role of the metric of spacetime, ϕ_{ab} $(a,b=1,2,\cdots,D-4)$ will be scalar fields, and B^a_μ will somehow represent the gauge degrees of freedom. Notice that both $\bar g_{AB}$ and $\bar g^{AB}$ are polynomial in B^a_μ, and that

$$\sqrt{\bar g}=\sqrt{g}\,\sqrt{\det\phi}[W(\phi)]^{D/2}\ . \tag{2.11}$$

Thus the Lagrangian will be polynomial in the gauge fields. (In fact, no higher than quartic terms will occur.)

If one computes the action, Eq. (2.1), based on the parametrization Eq. (2.8), one finds a term

$$\left[-\frac{1}{16\pi G_D}\right]\sqrt{\det\phi}[W(\phi)]^{D/2-1} \times\sqrt{g}\left[\tfrac{1}{4}F^a_{\mu\nu}F^b_{\kappa\lambda}g^{\mu\kappa}g^{\nu\lambda}\phi_{ab}\right]\ . \tag{2.12}$$

[The natural choice for W is $W(\phi)=(\det\phi)^{-1/(D-2)}$; cf. Eq. (2.4).] Here

$$F^a_{\mu\nu}=B^a_{\mu;\nu}-B^a_{\nu;\mu}\ , \tag{2.13}$$

where

$$B^a_{\mu;\nu}=B^a_{\mu,\nu}-B^b_\nu B^a_{\mu,b}\ . \tag{2.14}$$

This is almost the generalization of the five-dimensional result to the non-Abelian case. We must somehow convert the extra terms in $F^a_{\mu\nu}$ into the standard $f^{abc}A^b_\mu A^c_\nu$ form. The way to do this is to assume the existence, on the internal manifold, of a set of vector fields ξ^a_α which obey the relations

$$[\zeta_\alpha,\zeta_\beta]^a \equiv \zeta_\alpha^b \zeta_{\beta,b}^a - \zeta_\beta^b \zeta_{\alpha,b}^a = -f^{\alpha\beta\gamma}\zeta_\gamma^a . \quad (2.15)$$

Here $[\zeta_\alpha,\zeta_\beta]$ is the standard Lie bracket. We also assume that $\zeta_{\alpha,\mu}^a=0$. Setting

$$B_\mu^a = \zeta_\alpha^a A_\mu^\alpha \quad (2.16)$$

it then follows that

$$F_{\mu\nu}^a = \zeta_\alpha^a (A_{\mu,\nu}^\alpha - A_{\nu,\mu}^\alpha) - A_\nu^\beta \zeta_\beta^b (\zeta_\alpha^a A_\mu^\alpha)_{,b} + A_\mu^\beta \zeta_\beta^b (\zeta_\alpha^a A_\nu^\alpha)$$
$$= \zeta_\alpha^a (A_{\mu,\nu}^\alpha - A_{\nu,\mu}^\alpha - f^{\alpha\beta\gamma} A_\mu^\beta A_\nu^\gamma) + \cdots , \quad (2.17)$$

where the ... refers to terms that depend on the derivatives of A_μ^α with respect to the internal coordinates.

A problem with the non-Abelian generalization of the Kaluza-Klein idea is that, whereas in the five-dimensional case the manifold $M^4 \times S^1$, with metric given in Eq. (2.7) solves the Einstein equations and hence serves as an appropriate background for the quantum dynamics, no such manifold exists in the non-Abelian case. One can see this as follows. Let us assume that the manifold we are searching for has a "vacuum" metric of the form

$$\bar{g}_{AB} = W(\phi) \begin{bmatrix} \delta_{\mu\nu} & 0 \\ 0 & \phi_{ab}(y) \end{bmatrix} . \quad (2.18)$$

This is the analog of the vacuum solution, Eq. (2.7), in the Abelian case. Here we use y to denote the coordinates of the internal manifold. Inserting this ansatz into the Einstein equations, one finds, first, that W must be a constant and, second, that

$$R_{ab}(\phi) = 0 . \quad (2.19)$$

$R_{ab}(\phi)$ is the Ricci tensor computed from ϕ_{ab} alone. The solution to Eq. (2.19) is complicated by the fact that we need the vector fields introduced in Eq. (2.15). Indeed, for the vacuum solution, one would like these vectors to be symmetries of the metric ϕ_{ab}:

$$\nabla_a \zeta_{ab} + \nabla_b \zeta_{aa} = 0 , \quad (2.20)$$

where the covariant derivative ∇_a is with respect to the metric ϕ. However, there is a theorem[12] which states that Eqs. (2.15) (with nonvanishing f's), (2.19) and (2.20) are incompatible on a compact Riemannian manifold. If one gives up (2.15), one gives up non-Abelian gauge theories; if one abandons Eq. (2.19), one abandons Einstein's equation; and if one forsakes Eq. (2.20), one has a hard time understanding what it means to have a natural candidate for the vacuum field configuration. In the literature, the usual assumption is that the pure Einstein equations no longer apply, either because of the addition of extra matter fields,[13] or because of the presence of torsion,[14] or because of quantum effects.[15] The danger to be avoided, however, is that in a more complicated theory the assumption of extra dimensions may fail to explain anything, such as the origin of gauge theories. Under such circumstances, one might as well stick to four dimensions.

To close this section, we return briefly to the question of a cosmological constant. We shall see in Sec. IV that, just as in any nonsupersymmetric field theory, the fluctuations of the quantum fields induce a constant, infinite, vacuum energy density, which in a theory of gravity plays the role of a cosmological constant. In the usual spirit of renormalization, it becomes unnatural to expect that the observed cosmological constant should be zero; in fact, its natural size exceeds the observed upper bound by 120 orders of magnitude.

There is no commonly accepted resolution of this inconsistency. For the purposes of this paper, we shall simply take the cosmological constant to vanish as an observational fact, and we shall tune our parameters at each stage of the calculation to ensure that it remains zero, however unnatural an act that may be.

To see the dynamical role of a cosmological term, we look at the five-dimensional Einstein equations in the presence of a cosmological constant Λ. In four dimensions, one has the usual de Sitter and anti-de Sitter solutions, depending on the sign of Λ. In the five-dimensional case, we postulate the form

$$ds^2 = -dt^2 + f(t) d\sigma_k{}^2 + g(t)(dx^5)^2 . \quad (2.21)$$

Here $d\sigma_k{}^2$ represents the line element of a maximally symmetric three-dimensional space; there are three such spaces, of positive ($k=1$), zero ($k=0$), or negative ($k=-1$) constant curvature. We solve the Einstein equations

$$R_{AB} - \tfrac{1}{2}\bar{g}_{AB} R + \Lambda \bar{g}_{AB} = 0 , \quad (2.22)$$

or equivalently

$$R_{AB} = \alpha \bar{g}_{AB} , \quad (2.23)$$

$$\alpha = \tfrac{2}{3}\Lambda . \quad (2.24)$$

It is convenient to set $\phi = \dot{f}/f$; $\gamma = \dot{g}/g$. Equations (2.23) reduce to

$$\alpha = \tfrac{3}{2}\dot{\phi} + \tfrac{3}{4}\phi^2 + \tfrac{1}{2}\dot{\gamma} + \tfrac{1}{4}\gamma^2 , \quad (2.25)$$

$$\alpha = \frac{k}{3f} + \tfrac{1}{2}\dot{\phi} + \tfrac{1}{4}\phi(3\phi+\gamma) , \quad (2.26)$$

$$\alpha = \tfrac{1}{2}\dot{\gamma} + \tfrac{1}{4}\gamma(3\phi+\gamma) . \quad (2.27)$$

Together, Eqs. (2.25) and (2.27) imply that

$$\gamma = \frac{2\dot{\phi}}{\phi} + \phi , \quad (2.28)$$

i.e., that

$$g = g_0 \frac{\dot{f}^2}{f} . \quad (2.29)$$

Then Eq. (2.26) determines f to be

$$f(t) = Ae^{\omega t} + Be^{-\omega t} + \frac{k}{3\alpha} , \quad (2.30)$$

where $\omega^2 = \alpha$ and A and B are constants of integration. If $\alpha > 0$, A and B are each real and arbitrary; if $\alpha < 0$, $B = A^*$ with A arbitrary.

Of course, if Λ is small enough, we can, by setting $k=0$ in Eq. (2.30), return arbitrarily closely to the time-independent solution, Eq. (2.7). Here we show, in addition, that in the same limit one can also recover the Kasner solution, Eq. (2.8).

To achieve this, we look at the case $k=0$, $\alpha > 0$, and

A,B of opposite sign. Then by choosing the origin of time so that $f(0)=0$, we have the solution

$$f(t)=A_0\sinh\omega t \ , \tag{2.31}$$

$$g(t)=(A_0 g_0 \alpha)\left[\frac{\cosh^2\omega t}{\sinh\omega t}\right] . \tag{2.32}$$

To get a nontrivial limit as $\alpha\to 0$, it is necessary to take $A_0=B_0/\omega$ where B_0 remains finite. Then, provided $\omega t \ll 1$, we find

$$f(t)\sim B_0 t \ , \tag{2.33}$$

$$g(t)\sim \frac{B_0 g_0}{t} . \tag{2.34}$$

This is the Kasner-type behavior that we sought. Notice however, that when $\omega t \gg 1$, we have

$$f(t)\sim A_0 e^{\omega t} , \tag{2.35}$$

$$g(t)\sim (A_0 g_0 \alpha) e^{\omega t} . \tag{2.36}$$

Thus, if α is sufficiently small, there can be an arbitrarily long Kasner-type epoch; nevertheless, unless α is strictly zero, the fifth dimension will begin to grow (the changeover from contraction to expansion occurs when $e^{\omega t}=\sqrt{2}+1$), and will ultimately reemerge from the obscurity of the submicroscopic world.

III. QUANTUM GRAVITY IN D DIMENSIONS

A. General remarks

Classical gravitational theories, like all classical dynamical systems, must become inadequate in regimes where quantum effects can no longer be neglected. In the case of gravity, this happens when distance scales on the order of the Planck length are being approached. When the Planck length is reached, however, the effective gravitational self-coupling will become of order 1 and the loop expansion will break down. Thus, unless important quantum effects can be identified at distance scales larger than the Planck length, the conventional loop-expansion approach to the quantization of the gravitational field may not be useful.

Kaluza-Klein theories offer just such a distance scale. The radius of curvature of the extra dimension(s) can, *a priori*, be larger than the Planck length and, as we shall show in the next section, it can provide the scale for an important and reliably computable quantum effect. In order to set the stage for this computation, we briefly review some of the main features of the quantization of general relativity. We emphasize that, even if Einstein's theory is only an effective theory for use at large distances beyond the Planck length, the kind of situation which arises in Kaluza-Klein theories could make the quantization program meaningful and useful.

The path integral quantization of a field theory proceeds by first finding an extremum of the action—a classical solution, as discussed in the previous section—and then computing the quantum corrections as an expansion in $\hbar$. A natural way of implementing this program, especially if general properties such as renormalizability are being studied, is the background-field method. The fields are split into a classical part and a quantum part and the action is then expanded in the quantum fields about arbitrary classical background fields. An expansion to second order is sufficient to generate all one-loop diagrams with any number of external lines attached. Physics is then extracted by putting the external lines on the mass shell, that is, by taking the background fields to satisfy the classical field equations.

In general relativity, the background-field method[16] has been extensively used to construct the on-shell counterterms for the study of renormalizability.[17,18] The result of these studies, as expected from the fact that Newton's constant is dimensionful, is that the theories are nonrenormalizable. This means that not only is one restricted to using them to construct low energy expansions at distances beyond the Planck length, but the coefficients in this expansion must be taken to be arbitrary parameters. We expect this to be the case in higher-dimensional theories as well but we will see that some quantities might be reliably computed at distances beyond the Planck length. In particular, quantum effective potentials which govern the dynamics of the extra compact dimensions can be computed when these dimensions are larger than the Planck length.

B. Quantizing the five-dimensional theory

We shall consider the pure Einstein theory, Eq. (2.1), in five dimensions, taking as a starting classical solution the generalized Minkowski space, Eq. (2.7), with $0\le x^5 < 2\pi R_5$. The effect of an additive cosmological term, which is naturally induced by the quantum corrections, was discussed in Sec. II and we shall briefly return to it at the end of this section.

The metric $\bar{g}_{AB}$ is parametrized as in Eq. (2.4) and quantum fields $h_{\mu\nu}$, A_μ, and ϕ' are then introduced by setting

$$\begin{aligned} g_{\mu\nu}&=\delta_{\mu\nu}+h_{\mu\nu} \ , \\ \phi&=\phi_c+\phi' \ . \end{aligned} \tag{3.1}$$

The action, Eq. (2.1), can then be expanded as an infinite series in the quantum fields. Note that even though these fields have been introduced with an eye toward dimensional reduction, they are at this point still functions of x_μ and x_5.

To determine propagators and to compute one-loop diagrams in the classical background field, it suffices to keep only terms quadratic in the quantum fields. The result of this exercise is

$$S^{(1)}=-\frac{1}{16\pi G_5}\int d^5x\, \mathscr{L}^{(2)} , \tag{3.2}$$

where

$$\mathscr{L}^{(2)}=\mathscr{L}_{\rm Gr}+\mathscr{L}_G+\mathscr{L}_S+\mathscr{L}_5 \tag{3.3}$$

and

$$\begin{aligned} \mathscr{L}_{\rm Gr}=&\tfrac{1}{4}(h_{\mu\nu,\rho}h_{\mu\nu,\rho}-2h_{\mu\nu,\rho}h_{\mu\rho,\nu} \\ &-2h_{\mu\nu,\mu}h_{\rho\rho,\nu}+h_{\rho\rho,\mu}h_{\sigma\sigma,\mu}) \\ &+\frac{1}{4\phi_c}(h_{\mu\nu,5}h_{\mu\nu,5}-h_{\mu\mu,5}h_{\nu\nu,5}) \ , \end{aligned} \tag{3.4}$$

$$\mathscr{L}_G=\tfrac{1}{4}\phi_c(F_{\mu\nu})^2\ , \tag{3.5}$$

$$\mathscr{L}_S=\frac{1}{6\phi_c{}^2}(\partial_\mu\phi')^2\ , \tag{3.6}$$

$$\mathscr{L}_5=-\frac{3}{4\phi_c}\left[\left[\frac{\phi'_{,5}}{\phi_c}\right]^2-h_{\mu\mu,5}\frac{\phi'_{,5}}{\phi_c}\right] -\frac{3}{2}\frac{\phi'_{,5}}{\phi_c}\partial_\mu A_\mu+A_{\mu,5}(h_{\nu\nu,\mu}-h_{\mu\nu,\nu})\ . \tag{3.7}$$

There are several points worth making about this result:

(1) There is no mixing among $h_{\mu\nu}$, A_μ, and ϕ' in the $n=0$ sector. This was, of course, already demonstrated in the full zero-mode Lagrangian, Eq. (2.6), and is a consequence of the variable choice, Eq. (2.4), in particular the presence of the Weyl factor $\phi^{-1/3}$.

(2) The first piece $\mathscr{L}_{\rm Gr}$, Eq. (3.4), is the usual quadratic Lagrangian for the graviton, together with a masslike term for the $n\neq 0$ modes. It can be inverted for $n\neq 0$ to give the unitary form of the massive, spin-2 propagator.

(3) The second two pieces, $\mathscr{L}_G$ and $\mathscr{L}_S$, describe the massless gauge field and scalar field.

(4) The role of the constant, classical background field ϕ_c as the metric for the fifth dimension can be seen in each term.

A simple and natural gauge choice for the quantization of this theory is the cylindrical gauge

$$\bar{g}_{\mu 5,5}(x)=0\ . \tag{3.8}$$

It is analogous to the static gauge in finite temperature gauge theories, discussed in the Appendix. It can be effected by the addition of the gauge-fixing Lagrangian

$$\Delta\mathscr{L}_{\rm gf}=-\tfrac{1}{2}(\Lambda^B\bar{g}_{AB})^2\ , \tag{3.9}$$

where

$$\Lambda^B=(\Lambda^\mu,\zeta\partial_5)\ , \tag{3.10}$$

and Λ^μ is, for the moment, an arbitrary four vector. The cylindrical gauge, Eq. (3.8), is then obtained by taking the limit $\zeta\to\infty$. Only the $n=0$ mode in the Fourier expansion, Eq. (2.5), will remain sensitive to Λ^μ in this limit. Using the variables $h_{\mu\nu}$, A_μ, and ϕ', and referring to the quadratic Lagrangian, Eqs. (3.3)–(3.7), Λ^μ can be used for further gauge fixing within the massless, zero-mode sector.

In the cylindrical gauge, the fields A_μ and ϕ are independent of x_5—they are purely zero mode. Because of this, the $\mathscr{L}_5$ term in the quadratic Lagrangian, Eqs. (3.3) and (3.7), which mixes the fields, must vanish. The underlying reason for this useful simplification is that the cylindrical gauge is *the* physical gauge for the five-dimensional Kaluza-Klein theory. To see this, we note that in N dimensions, an $N\times N$ metric field g_{AB} describes $(N^2+N)/2-2N$ physical degrees of freedom, the subtraction of $2N$ corresponding to the freedom in the choice of a coordinate basis and the fact that g_{A0} are dependent variables. The 5 degrees of freedom in the five-dimensional theory are then accounted for in the following way.[19] In the zero-mode sector, they correspond to the massless graviton, photon, and Brans-Dicke scalar. By contrast, the 5 degrees of freedom in each nonzero mode are those of a single massive, spin-2 particle. Clearly the cylindrical gauge explicitly reflects the physics of the five-dimensional Kaluza-Klein theory.

The absorption of spin-0 and spin-1 degrees of freedom to create a massive spin-2 particle is, of course, a Higgs mechanism. The chosen classical solution [Eqs. (2.2) and (2.7)] involving, as it does, a fixed scale associated with the fifth dimension, clearly does not have all the symmetries of the original Lagrangian. Spontaneous symmetry breakdown ("spontaneous compactification") has taken place for those modes ($n\neq 0$) which are sensitive to the presence of the scale R_5, and the fields A_μ and ϕ are the associated Goldstone fields.

The change in $\bar{g}_{AB}$ induced by an infinitesimal coordinate transformation

$$x^A\to x^A-\epsilon^A(x) \tag{3.11}$$

is

$$\delta\bar{g}_{AB}(x)=\epsilon^C\bar{g}_{AB,C}+\epsilon^C{}_{,A}\bar{g}_{CB}+\epsilon^C{}_{,B}\bar{g}_{AC}\ . \tag{3.12}$$

From this, one finds in the usual way that the DeWitt-Faddeev-Popov ghost factor Δ is

$$\Delta=\det M\ , \tag{3.13}$$

where M is defined by

$$\int M_{AB}(x,x')\epsilon^B(x')dx'=[\bar{g}_{A5,B}\partial_5+\bar{g}_{B5}\partial_A\partial_5+\bar{g}_{AB}(\partial_5)^2+\bar{g}_{AB,5}\partial_5]\epsilon^B(x)\ . \tag{3.14}$$

The gauge condition $\bar{g}_{A5,5}=0$, corresponding to the limit $\zeta\to\infty$ in Eq. (3.10), has been used to simplify this expression. The quantity Δ may be represented as an integral over anticommuting variables:

$$\Delta=\int D\eta\, D\bar{\eta}\exp\left[\int dx\, dx'\bar{\eta}^A(x)M_{AB}(x,x')\eta_B(x')\right]\ . \tag{3.15}$$

We see that the ghost action is already quadratic in the fields η and $\bar{\eta}$, and hence the one-loop approximation is obtained by setting $\bar{g}_{AB}$ in Eq. (3.14) to the background value Eq. (2.7).

Having obtained the one-loop versions of the classical action and the gauge-fixing and ghost terms, we are almost ready to study the quantum theory defined by

$$z=\int D\mu(h)\Delta(h)e^{-[S^{(2)}+S_{\rm gf}]}\ , \tag{3.16}$$

where

$$\bar{g}_{AB}=g_{AB}^{(0)}+h_{AB}$$

and $g_{AB}^{(0)}$ is given in Eq. (2.7). The only remaining issue is what to take for the measure $D\mu(h)$. This question has been discussed at length by Fradkin and Vilkovisky,[20] and by DeWitt[21] and 't Hooft,[22] among others. Its definitive resolution probably awaits a well-defined quantum theory of gravity. However, it is clear that the measure must at least intervene to cancel unwanted terms in physical quantities proportional to $\delta^4(0)$ and, in our case, $\delta^5(0)$. This role of the measure can be seen, for example, in derivatively coupled scalar theories in four dimensions, where Feynman graphs with one or more loops typically contain quartic divergences. These are precisely canceled by a factor in the measure that can be derived by starting with the phase-space path integral where the measure factor is unity and then integrating out the canonical momenta to obtain the standard path integral in configuration space.

In performing our computation, we shall enforce this cancellation by setting $D\mu(h)=Dh$, and at the same time systematically deleting any terms in the effective potential which are proportional to $\delta^4(0)$ or $\delta^5(0)$. This is consistent with having chosen a regularization scheme such as dimensional continuation or zeta-function regularization, in which these singular factors would have been defined to zero from the beginning.

C. Divergences and counterterms

In D-dimensional quantum gravity, the maximum degree of divergence of an L-loop graph is

$$d=(D-2)L+2 . \tag{3.17}$$

This maximum, independent of the number of external lines, is attained only if the two momentum factors at each vertex are taken to be internal loop momenta. The counterterm being computed would then have no derivatives acting on the fields; it would correspond to an induced cosmological constant. In general, an L-loop amplitude will produce terms ranging from finite ones (thank goodness) up to the maximally divergent cosmological constant. The possible divergent terms will be dictated by the local counterterms allowed by D-dimensional general covariance. The actual divergences encountered in a computation will further depend on the regulation scheme employed. We here restrict ourselves to some general remarks about counterterms, imagining that a cutoff Λ is in place on the momentum integrals. A specific such cutoff will be introduced in the next section in the course of describing the effective potential computation.

At one loop, the maximal degree of divergence is $d=D$. For the five-dimensional theory, the leading quintic divergence corresponds to an induced cosmological constant. It must, of course, be fine-tuned away and we have nothing new to add to this age-old problem. A cubic divergence will appear multiplying the scalar curvature R and it can, therefore, be absorbed into a renormalization of Newton's constant. Finally, there can be a linear divergence multiplying the four-derivative operator $(R_{\mu\nu\lambda\sigma})^2$. In an even number of dimensions, this operator must vanish on shell[12] because it is proportional to $R^2-4R_{\mu\nu}{}^2$ plus a total divergence. In five dimensions, it persists and one sees the nonrenormalizability already at one loop. Of course, all these divergences would be defined to be zero using dimensional continuation or zeta-function regularization.

The effective potential is a zero-derivative operator and it is therefore a (finite) partner of the induced cosmological constant. It will be possible to separate these two pieces and then to focus our attention on the finite part. The presence of the other, higher-derivative, counterterms at one loop will not impact on the effective potential. In higher orders, all these terms can be mixed together and this problem will be briefly discussed after computing the effective potential.

IV. COMPUTATION OF THE EFFECTIVE POTENTIAL

Having outlined the general framework, we turn to the explicit evaluation of the effective potential. We begin with

$$Z[\phi_c]=\int Dh\,\delta(h_{A5,5})\Delta(h)e^{-S[h]} , \tag{4.1}$$

where we have expanded the metric as in Eq. (3.1). We have chosen to take the $\zeta\to\infty$ limit of the gauge-fixing Lagrangian, Eqs. (3.9) and (3.10), immediately; hence the delta functional appears in Eq. (4.1). To compute the effective potential to one loop, we use[23]

$$Z[\phi_c]=\exp[-V_{\text{eff}}(\phi_c)(g^{(0)})^{1/2}\int d^5x] . \tag{4.2}$$

This differs from the usual expression in that we have extracted a factor $(g^{(0)})^{1/2}$ in defining V_{eff}. If we did not do this, V_{eff} would not be generally covariant, and hence would be without physical significance.

For a one-loop computation, we need only the terms in S that are quadratic in the quantum fields. The relevant Lagrangian has been given in Eqs. (3.4)—(3.6). We can see explicitly from these equations how it is that the zero modes (the scalar ϕ', the vector A_μ, and the x_5-independent piece of the tensor $h_{\mu\nu}$) do not contribute to V_{eff}. We observe that by scaling the integration variables

$$\phi'\to\phi_c\phi' , \tag{4.3a}$$

$$A_\mu\to\frac{1}{\sqrt{\phi_c}}A_\mu , \tag{4.3b}$$

we can remove the dependence on ϕ_c completely from the zero-mode terms. The price we pay for this is that we generate a Jacobian factor in the path integral, e.g.,

$$D\phi'\to[\prod\phi_c]D\phi'=(D\phi')\exp\left[\sum_x \ln\phi_c\right] .$$

Using the relationship $(d^4x)\delta^4(0)=1$, we have

$$D\phi'\to\exp\left[\delta^4(0)\int d^4x\,\ln\phi_c\right](D\phi') . \tag{4.4}$$

This ill-defined Jacobian factor is precisely the sort of thing that must be absorbed into the measure of integration, as discussed in Sec. III. The same holds true for our rescaling of A_μ. Thus we conclude that the zero modes contribute a ϕ_c-independent factor to $Z[\phi_c]$, which will be absorbed into the overall normalization of the path integral [see Eq. (4.15) below].

Therefore, we need only concern ourselves with the massive-mode contribution to the path integral, and possibly with the ghost-determinant contribution as well. We shall first evaluate the massive mode integral, and then show that the ghost factor in fact makes no contribution.

To evaluate the massive-mode path integral, it is convenient to diagonalize the derivatives by means of Fourier transformation. Denoting by $\tilde{S}$ the contribution of these modes to the action, we have

$$\tilde{S}=\frac{1}{8}\sum_{n=-\infty}^{\infty}\int\frac{d^4k}{(2\pi)^4}h^{(n)}_{\mu\nu}(k)\{[\tfrac{1}{2}(\delta^{\mu\lambda}\delta^{\nu\sigma}+\delta^{\mu\sigma}\delta^{\nu\lambda})-\delta^{\mu\nu}\delta^{\lambda\sigma}](k^2+m_n{}^2)+(\delta^{\mu\nu}k^\lambda k^\sigma+\delta^{\lambda\sigma}k^\mu k^\nu)$$
$$-\tfrac{1}{2}(\delta^{\mu\lambda}k^\nu k^\sigma+\delta^{\mu\sigma}k^\nu k^\lambda+\delta^{\nu\lambda}k^\mu k^\sigma+\delta^{\nu\sigma}k^\mu k^\lambda)\}h^{(-n)}_{\lambda\sigma}(-k)\,, \tag{4.5}$$

where

$$M_n{}^2=\frac{1}{\phi_c}\frac{n^2}{R_5{}^2}\,. \tag{4.6}$$

(It is easier to include the $n=0$ piece even though it will not contribute to V_{eff}.) If we abbreviate this expression by

$$\tilde{S}=\frac{1}{2}\sum_n\int\frac{d^4k}{(2\pi)^4}h^{(n)}_{\mu\nu}(k)\mathscr{S}^{\mu\nu\lambda\sigma}_{(n;k)}h^{(-n)}_{\lambda\sigma}(-k)\,, \tag{4.7}$$

then we have the standard result that

$$\int Dh_{\mu\nu}e^{-\tilde{S}}=(\det\mathscr{S})^{-1/2}\,. \tag{4.8}$$

We evaluate the determinant by finding the eigenvalues of $\mathscr{S}$ and taking their product. For each value of k and n, the index pair $(\mu\nu)$ takes on 10 values. Thus we can think of $\mathscr{S}$ as a 10×10 matrix labeled by k and n. (It is the fact that in momentum space, $\mathscr{S}$ is block diagonal with finite-dimensional blocks that allows us to evaluate $\det\mathscr{S}$ this way.)

Let $p^{\mu(j)}$ be vectors satisfying $p^{\mu(j)}k_\mu=0$ $(j=1,2,3)$. Then by explicit computation we find the following

(a) $4\mathscr{S}^{\mu\nu\lambda\sigma}v^{(j)}_{\lambda\sigma}=M_n{}^2v^{\mu\nu(j)}$,

where

$$v^{(j)}_{\lambda\sigma}=k_\lambda p^{(j)}_\sigma+k_\sigma p^{(j)}_\lambda,\quad j=1,2,3 \tag{4.9}$$

(i.e., there are three independent v's corresponding to the three independent choices for p^μ).

(b) $4\mathscr{S}^{\mu\nu\lambda\sigma}w^{(i)}_{\lambda\sigma}=\lambda^{(i)}w^{\mu\nu(i)},\quad i=1,2$,

where

$$\lambda^{(i)2}+2(k^2+M_n{}^2)\lambda^{(i)}-3M_n^{(4)}=0 \tag{4.10}$$

and $w^{\mu\nu(i)}$ are two linearly independent combinations of $\delta^{\mu\nu}$ and $k^\mu k^\nu$. Note that Eq. (4.10) implies that $\lambda^{(1)}\lambda^{(2)}=-3M_n{}^4$.

(c) $4\mathscr{S}^{\mu\nu\lambda\sigma}X^{(j)}_{\lambda\sigma}=(k^2+M_n{}^2)X^{\mu\nu(j)},\ j=1,2$, (4.11)

where the $X^{(j)\mu\nu}$ are appropriate linear combinations of $\delta^{\mu\nu}$, $k^\mu k^\nu/k^2$, and $p^{(j)\mu}p^{(j)\nu}/p^{(j)2}$. Although there would appear to be three $X^{(j)}$'s, corresponding to the three possible choices of $p^{\mu(j)}$, this is reduced to two because of the completeness relation

$$\delta^{\mu\nu}=\frac{k^\mu k^\nu}{k^2}+\sum_{j=1}^{3}\left[\frac{p^{\mu(j)}p^{\nu(j)}}{p^{(j)2}}\right].$$

(d) $4\mathscr{S}^{\mu\nu\lambda\sigma}Y^{(j)}_{\lambda\sigma}=(k^2+M_n{}^2)Y^{(j)\mu\nu},\quad j=1,2,3$, (4.12)

where $Y^{\mu\nu(j)}$ correspond to the three independent ways of choosing $p^{\mu(l)}$ and $p^{\mu(k)}$ in the expression

$$Y^{\mu\nu}=p^{\mu(l)}p^{\nu(k)}+p^{\nu(l)}p^{\mu(k)},\ (l\neq k)\,.$$

Thus the product of all the eigenvalues is given by

$$d(k,n)=-3[M_n{}^2(k^2+M_n{}^2)]^5\,. \tag{4.13}$$

There are five eigenvalues independent of k, corresponding to the five nonpropagating modes in the field $h_{\mu\nu}(k)$ (comprising a massive spin-1 and two spin-0 modes). There are also five eigenvectors with eigenvalue $(k^2+M_n{}^2)$; these are the five propagating components of spin 2. The fact that $d(k,0)=0$ is a signal that further gauge fixing is necessary in the zero-mode sector (i.e., one must choose coordinates in ordinary four-dimensional relativity). Since the zero modes do not contribute to V_{eff}, we need not concern ourselves with this problem.

We can now write, formally,

$$(\det 4\mathscr{S})^{-1/2}=\exp\left[-\frac{1}{2}\sum_n\sum_k\ln[-3M_n{}^{10}(k^2+M_n{}^2)^5]\right]\,. \tag{4.14}$$

Our first step is to normalize this expression by dividing by its value when $\phi_c=1$. That is, we compute the ratio

$$\frac{Z[\phi_c]}{Z[\phi_c=1]}=\exp\left\{-\frac{1}{2}\sum_n\sum_k\ln\left[\frac{1}{\phi_c{}^2}\frac{\phi_c(kR_5)^2+n^2}{(kR_5)^2+n^2}\right]^5\right\}\,. \tag{4.15}$$

This is just a specification of the arbitrary overall normalization of Z, and is without physical significance. It corresponds to adding an arbitrary constant to $\sqrt{g}\,V_{\text{eff}}$; by contrast, a cosmological constant, which *is* physical, is a constant term in V_{eff} itself. Next, we use the relationship

$$\left[\int d^4x\right]\frac{d^4k}{(2\pi)^4}=1 \tag{4.16}$$

to write

$$\frac{Z[\phi_c]}{Z[\phi_c=1]}=\exp\left\{-\frac{5}{2}\sum_{n=-\infty}^{\infty}\left[\frac{1}{2\pi R_5}\right]\int\frac{d^4k}{(2\pi)^4}\left[\ln\left[\frac{\phi_c(kR_5)^2+n^2}{(kR_5)^2+n^2}\right]-2\ln\phi_c\right]\left[\int d^5x\right]\right\}\,. \tag{4.17}$$

Once again, we drop the second term since it is of the form $(\text{const})\times(\ln\phi_c)\times\delta^5(0)$, which can be absorbed by rescaling of the integration variable $h_{\mu\nu}(x)$ by an appropriate power of ϕ_c. This leaves us with

$$(g^{(0)})^{1/2}V_{\rm eff}(\phi_c)-V_{\rm eff}(\phi_c=1)=\frac{1}{2\pi R_5}\,\tfrac{5}{2}\sum_n\int\frac{d^4k}{(2\pi)^4}\ln\left[\frac{\phi_c(kR_5)^2+n^2}{(kR_5)^2+n^2}\right]. \tag{4.18}$$

$g^{(0)}$ is the determinant of the classical solution $g_{AB}{}^0$. Its value is $\phi_c{}^{-2/3}$.

Before proceeding further with our analysis of Eq. (4.18), let us return to the contribution of the ghost determinant, Eq. (3.14). Transforming to momentum space, we find from Eq. (3.14)

$$\begin{aligned}M_{AB}^{\text{one-loop}}(k,n)&=-k_5[k_A\bar{g}_{B5}^{(0)}+k_5\bar{g}_{AB}^{(0)}]\\&=-k_5\phi_c{}^{-1/3}[\phi_c k_A\delta_{B5}+k_5\delta_{AB}+k_5\delta_{A5}\delta_{B5}(\phi_c-1)]\,.\end{aligned} \tag{4.19}$$

Here we have set $k_5=+n/R_5$. The eigenvalue equation for M reads

$$-k_5\phi_c{}^{-1/3}[\phi_c k_A v_5+k_5v_A+k_5(\phi_c-1)\delta_{A5}v_5]=\lambda v_A\,. \tag{4.20}$$

If $v_5=0$, then this equation is satisfied with

$$\lambda=-k_5{}^2\phi_c{}^{-1/3}\,.$$

If $v_5\neq 0$ then setting $A=5$ we have

$$\lambda=-k_5\phi_c{}^{-1/3}[2\phi_c k_5]\,.$$

Thus we have four eigenvectors with eigenvalue $-k_5{}^2\phi_c{}^{-1/3}$, and one with eigenvalue $-2k_5{}^2\phi_c{}^{2/3}$. Their product is, therefore,

$$d_G=-2\left[\frac{n}{R_5}\right]^{10}\phi_c{}^{-2/3} \tag{4.21}$$

and the ratio $d_G(\phi_c)/d_G(\phi_c=1)$ is simply $\phi_c{}^{-2/3}$. Hence to one loop the ghost determinant contributes only a scaling term (const)$\times\ln\phi_c\times\delta^5(0)$ to the logarithm of $Z(\phi_c)/Z(\phi_c=1)$ and so does not affect $V_{\rm eff}(\phi_c)$.

Returning to Eq. (4.18), we use the formula

$$\sum_{n=-\infty}^{\infty}f_n=\int_{-\infty}^{\infty}dz\,f(z)+\int_{-\infty+i\epsilon}^{\infty+i\epsilon}dz\frac{[f(z)+f(-z)]}{e^{-2\pi iz}-1} \tag{4.22}$$

to perform the mode sum. This yields

$$\begin{aligned}\phi_c{}^{-1/3}V_{\rm eff}(\phi_c)-V_{\rm eff}(\phi_c=1)&=\tfrac{5}{2}\int\frac{d^5k}{(2\pi)^5}\ln\left[\frac{\phi_ck^2+k_5{}^2}{k^2+k_5{}^2}\right]\\&\quad+\frac{5}{2\pi(R_5)^5}\int\frac{d^4q}{(2\pi)^4}\int_{-\infty+i\epsilon}^{\infty+i\epsilon}dz\ln\left[\frac{\phi_cq^2+z^2}{q^2+z^2}\right]\frac{1}{e^{-2\pi iz}-1}\,.\end{aligned} \tag{4.23}$$

In the first term, we have defined the integration variable $k_5=z/R_5$; k^2, however, continues to denote the four-dimensional product $k_\mu k^\mu$. In the second term, we have defined the dimensionless variable $q_\mu=k_\mu R_5$.

Let us focus for the moment on the second term. By contour methods, one can easily establish the formula

$$\int_{-\infty}^{\infty}dz\,H(z)\ln\left[\frac{z^2+a^2}{z^2+b^2}\right]=2\pi\int_b^a dx\,H(ix) \tag{4.24}$$

provided $H(z)$ is analytic in the upper half plane, and that

$$H(z)\ln\left[\frac{z^2+a^2}{z^2+b^2}\right]\leq\frac{c}{|z|^{1+\epsilon}}$$

as $|z|\to\infty$ in the upper half plane. We have defined the logarithm to have its cut on the negative real axis, and have taken a,b, to be the positive square roots of a^2,b^2. In our case, $H(z)=1/(e^{-2\pi iz}-1)$ so the analyticity and boundedness properties are satisifed, and we have for the second term

$$\frac{5}{(2\pi)(R_5)^5}\int\frac{d^4q}{(2\pi)^4}\left[2\pi\int_q^{q\sqrt{\phi_c}}dx\frac{1}{e^{2\pi x}-1}\right]. \tag{4.25}$$

The x integration is easily done. Defining $w_\mu=2\pi q_\mu$, we have

$$\frac{5}{(2\pi R_5)^5}\left[\frac{1}{\phi_c}-1\right]\int\frac{d^4w}{(2\pi)^4}\ln(1-e^{-w})\,, \tag{4.26}$$

i.e.,

$$\phi_c^{-1/3}V_{\rm eff}(\phi_c)-V_{\rm eff}(\phi_c=1)=\tfrac{5}{2}\int\frac{d^5k}{(2\pi)^5}\ln\left[\frac{\phi_c k^2+k_5^2}{k^2+k_5^2}\right]+\frac{5}{(2\pi R_5)^5}\left[\frac{1}{\phi_c^{\,2}}-1\right]\int\frac{d^4w}{(2\pi)^4}\ln(1-e^{-w})\ . \tag{4.27}$$

In the first term, we observe that the k_5 integral is convergent, and can be evaluated straightaway using Eq. (4.24). The first term then becomes

$$\tfrac{5}{2}\int\frac{k\,d^4k}{(2\pi)^4}2\pi(\sqrt{\phi_c}-1)\ . \tag{4.28}$$

Clearly this is quintically divergent. We define the integral by assuming the domain of integration to be the interior of a large sphere in momentum space: $k\le\Lambda$.

However, general covariance in the four-dimensional space parametrized by x^μ prevents us from assuming that Λ is independent of ϕ_c. We can see this as follows: The coordinate system we are using, in which the background four-dimensional metric is $\phi_c^{-1/3}\delta_{\mu\nu}$, is related to the standard Euclidean-space coordinate system x_0^μ in which the background metric is just $\delta_{\mu\nu}$, by

$$x^\mu=\phi_c^{1/6}x_0^\mu\ . \tag{4.29}$$

Now

$$\int\frac{d^4k}{(2\pi)^4}=\frac{\Lambda^4}{32\pi^4}=\delta^4(0)\ .$$

Thus, under the scaling, Eq. (4.29), we must have

$$\Lambda=\phi_c^{-1/6}\Lambda_0\ , \tag{4.30}$$

where Λ_0 is the corresponding cutoff in the standard coordinate system. Therefore, in evaluating the expression (4.28), we must cut off the first term with $\Lambda=\phi_c^{-1/6}\Lambda_0$, and the second term with Λ_0. This gives

$$\frac{\Lambda_0^{\,5}}{8\pi}(\phi_c^{-5/6}\phi_c^{1/2}-1)=\frac{\Lambda_0^{\,5}}{8\pi}(\phi_c^{-1/3}-1)\ . \tag{4.31}$$

Finally, returning to Eq. (4.27), we can extract $V_{\rm eff}(\phi_c)$:

$$V_{\rm eff}(\phi_c)=\frac{\Lambda_0^{\,5}}{8\pi}+\frac{5\beta}{(2\pi\phi_c^{1/3}R_5)^5}\ , \tag{4.32}$$

where β is shorthand for the integral appearing in Eq. (4.27). The difference taken in Eq. (4.27) would allow for the presence of an additional constant in

$$\int d^5x\,(g^{(0)})^{1/2}V_{\rm eff}=\int d^5x\,\phi_c^{-1/3}V_{\rm eff}$$

but, as noted following Eq. (4.15), this can be absorbed into the normalization of the functional integral. It is physically irrelevant and cleanly separable from the effective potential.

Our last numerical chore is to compute β:

$$\begin{aligned}\beta&=\int\frac{d^4w}{(2\pi)^4}\ln(1-e^{-w})\\&=-\frac{1}{8\pi^2}\sum_{n=1}^{\infty}\frac{1}{n}\int_0^\infty w^3e^{-nw}dw\\&=-\frac{3}{4\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^5}\ .\end{aligned} \tag{4.33}$$

Therefore

$$\begin{aligned}5\beta&=-\frac{15}{4\pi^2}\zeta(5)\\&\cong-0.394.\end{aligned}$$

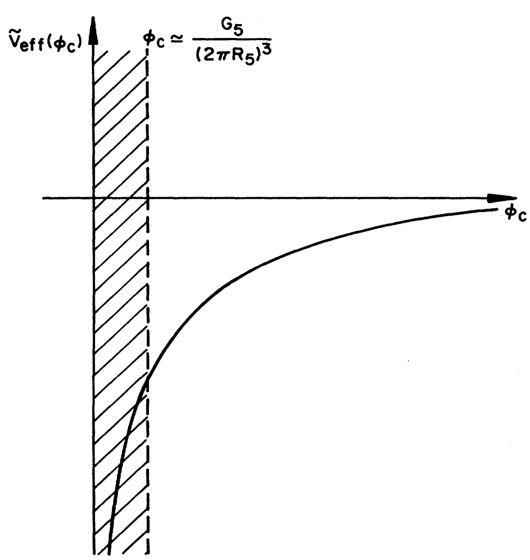

FIG. 1. The one-loop Kaluza-Klein effective potential

V. SUMMARY AND DISCUSSION

Our result for the effective potential consists of two pieces: the first is positive, divergent, and independent of the distance L_5 around the fifth dimension; the second is negative, convergent, and depends on L_5 in the manner illustrated in Fig. 1. This state of affairs is reminiscent of the well-known Casimir effect in electrodynamics. There one considers the zero-point energy of the free electromagnetic field between two infinite parallel perfectly conducting plates separated by a distance a, and one finds that the energy per unit volume has the form

$$E(a)=E_0-\frac{c}{a^4}\ . \tag{5.1}$$

Here E_0 is cutoff dependent, but independent of the separation a; the second term is finite and represents a measurable attractive force between the plates. E_0 is just the vacuum energy density of free space, and is to be subtracted out in computing physical quantities.

Our computation is the direct gravitational analog of the usual Casimir energy. The infinite term is an induced cosmological constant, and according to the philosophy put forward in Sec. II, we subtract it from the observable effective potential $\tilde V_{\rm eff}$:

$$\tilde V_{\rm eff}=V_{\rm eff}-\frac{1}{8\pi}\Lambda_0^{\,5}\ . \tag{5.2}$$

(Of course, if we had chosen to use dimensional continuation or zeta-function regularization, the infinite term would have been defined to zero *ab initio*.) The remaining term is then to be interpreted as an observable energy per unit volume, which tends to make the "distance between the plates"—in this case, the distance around the fifth dimension—contract.

From Fig. 1, however, one sees that there is no apparent limit to the contraction: $V_{\rm eff}$ tends to $-\infty$ as L_5 tends to zero. This is equally true of the Casimir case, but there one expects on physical grounds that a natural cutoff will appear when the structure of the plates becomes important, i.e., when a becomes of order the interatomic separation.

The analogous cutoff in the gravitational case is the Planck length. What we have computed is the first nonvanishing term in a loop expansion, which is an expansion in powers of the reciprocal of the constant multiplying the action, i.e., an expansion in powers of G_5. Now G_5 has dimension (length)^3; hence the dimensionless expansion parameter must be

$$\gamma=\frac{G_5}{{L_5}^3}\equiv\frac{G_4}{{L_5}^2}\ , \tag{5.3}$$

i.e., the loop expansion should have the form

$$\widetilde{V}_{\rm eff}=\widetilde{V}_{\rm eff}^{(1)}(1+\alpha_2\gamma+\alpha_3\gamma^2+\cdots)\ .$$

Here $\widetilde{V}_{\rm eff}^{(1)}$ is the one-loop term of Eq. (5.2) and α_n is a dimensionless factor which can be evaluated by computing the finite part of the n-loop contribution to $V_{\rm eff}$; presumably it is of order 1. Hence the loop expansion will make sense (and our one-loop result will be reliable) only if

$$\gamma\ll 1\ , \tag{5.4}$$

i.e.,

$$L_5\gg(G_4)^{1/2}=L_P\ ,$$

where L_P is the observed Planck length

$$L_P=1.6\times10^{-33}\ \text{cm}. \tag{5.5}$$

Thus our one-loop result loses its validity if $L_5\lesssim L_P$. We conclude, therefore, that the Casimir force tends to push the fifth dimension down to a size on the order of the Planck length, but we cannot say what happens after that. For the theory to make sense, it must presumably be the case that dynamics at the Planck scale or below will stabilize the extra dimension at this size.

A Planck-size stabilization would have two interesting consequences. First of all, the massive spin-2 modes are charged and their charge is given by a multiple of the ratio of the Planck length to the circumference of the compact dimension. Thus a dynamics which determine this distance relative to the Planck length will, in turn, determine the electric charge of the massive modes. If this charge is to be a multiple of the fine structure constant $\alpha=1/137$, then the extra dimension would have to be somewhat larger than the Planck length. A stable minimum of the effective potential would also give rise to a mass for the Brans-Dicke scalar, presumably on the order of the Planck mass. This would then screen the scalar out of the effective four-dimensional theory in the same way that finite temperature Debye-screens the long-range electric field in a gauge theory.

The five-dimensional model we have examined in detail is the simplest possible Kaluza-Klein scheme. It invites generalization in a variety of ways. One can, in the first place, have any number d of toroidally compactified dimensions in a D-dimensional space-time. The computation of the effective potential is not significantly more complicated than the $D=5$, $d=1$ case we have considered in this paper. In particular, if one looks at a five-dimensional theory with $d=2$, one can interpret the second compactified dimension as Euclidean time, in which case the distance around it has the interpretation of an inverse temperature β. This case has been analyzed by Rubin and Roth,[24] who find that the attractive Casimir force competes with the thermal pressure: If the ratio β/L_5 is sufficiently large, the fifth dimension will contract; otherwise, it expands. Once can also investigate how the Casimir energy affects $d>1$ compactified spatial dimensions: this will be the subject of a future publication.[25]

Of at least equal importance, and of much greater difficulty, is the problem of extending our analysis to the non-Abelian cases discussed classically in Sec. II. In order to achieve a spontaneously compactified vacuum, it may well be necessary to add additional fields, either ordinary matter[13] or perhaps spinning matter,[14,26] which will give rise to torsion. Even if a sensible classical Kaluza-Klein theory can be constructed, however, the task of performing the mode sum is likely to be very much more complicated.

There is also the problem of incorporating fermions in a realistic way. This may involve the introduction of supersymmetry; the possible close connection between Kaluza-Klein ideas and supersymmetry has been especially emphasized by Witten,[27] and the various possibilities for realizing 11-dimensional supergravity have been discussed by Duff and collaborators.[28]

Finally, we remark that if some version of the Kaluza-Klein idea is right, it must impact on the evolution of the early universe. Some suggestions for cosmology involving Kaluza-Klein theories have already been advanced[9,29]; it remains to be seen whether this approach can be naturally unified with standard big-bang and inflationary cosmological ideas.

APPENDIX A: SOME FEATURES OF FINITE TEMPERATURE GAUGE THEORIES

A system in thermal equilibrium at temperature T is described by the partition function

$$Z=\sum_i e^{-E_i\beta}\ , \tag{A1}$$

where $\beta=1/T$. Functional integral expressions for the partition function and for thermal expectation values can be derived using standard methods. Euclidean "time" runs between $t=0$ and $t=\beta$, and the fields are either periodic or antiperiodic in this interval depending on whether they describe bosons or fermions. For gauge theories, the Faddeev-Popov gauge-fixing procedure is employed with the ghost field periodic.

The Feynman rules for the quantum loop expansion are formed from the zero-temperature Euclidean Feynman rules by the replacement

$$\int \frac{d^4k}{(2\pi)^4} \rightarrow T \sum_{n=\pm\text{integer}} \int \frac{d^3k}{(2\pi)^3} , \tag{A2}$$

where the discrete energy sum is over $k_4 = 2n\pi T$ for bosons and $k_4 = (2n+1)\pi T$ for fermions. The ultraviolet divergences are the same as those of the zero-temperature theory. The zero-temperature counterterms can, therefore, be used to remove the divergences and the temperature T can be used as the scale to define the coupling constant.

At distance scales much larger than $1/T$, the theory simplifies by, in effect, reducing to the same theory in one less dimension. With the external momenta in Green's functions small compared to T, everything except the $n=0$ mode for the bosons can be expected to decouple, leaving only three-dimensional momentum integrals. The only exception to this is the self-mass $\pi_{44}^{ab}(q=0)$ of the A_4 component of the gauge field, which is ultraviolet divergent without the inclusion of the $n\neq 0$ modes. When they are included, one finds

$$\pi_{44}^{ab}(q=0) = (N + N_f/2)\frac{g^2T^2}{3} \tag{A3}$$

for an SU(N) gauge theory with N_f fermions. Thus, a color-electric Debye-screening mass is generated, meaning that at distances much larger than $1/gT$, the fourth (electric) component of the gauge field decouples. With only the three spatial components of the gauge field remaining and with only three-dimensional integrations to be done, the reduction to an effective three-dimensional theory is then complete.

A natural gauge choice for the implementation of this program is the static gauge

$$\partial_4 A_4^a = 0 . \tag{A4}$$

It can be obtained by the addition of the gauge-fixing term

$$\Delta\mathscr{L}_{\text{gf}} = \tfrac{1}{2}(\Lambda_\mu A_\mu^a)^2 , \tag{A5}$$

where $\Lambda_\mu = (-i\vec{\Lambda}, \lambda\partial_4)$ with $\vec{\Lambda}$ an arbitrary three-vector. The corresponding ghost Lagrangian is

$$\Delta\mathscr{L}_{\text{ghost}} = \eta^{a\dagger}\Lambda_\mu(\partial_\mu\eta^a + gf^{abc}A_\mu^b\eta^c) . \tag{A6}$$

The static gauge is then obtained in the limit $\lambda\rightarrow\infty$. It has the following properties:

(1) The A_4 propagator is static ($n=0$ only) with

$$iD_{44}^{ab}(k) = -i\delta^{ab}/k^2 . \tag{A7}$$

At one loop, a mass is generated by the Debye-screening mechanism.

(2) The A_i propagator has a nonstatic contribution ($k_0 = 2n\pi T, n\neq 0$):

$$n\neq 0\text{: } iD_{ij}^{ab}(k) = -i\delta^{ab}\frac{\delta_{ij} + k_ik_j/k_0}{k^2 + k_0{}^2} . \tag{A8}$$

It also has a static ($n=0$) contribution which depends on $\vec{\Lambda}$, that is, on further gauge fixing within the static sector. In the O(3)-covariant gauge $\vec{\Lambda} = -(i/\zeta)\vec{\partial}$,

$$n=0\text{: } iD_{ij}^{ab}(\vec{k}) = -i\frac{\delta_{ab}}{\vec{k}^2}\left[\delta_{ij} + (\zeta-1)\frac{k_ik_j}{\vec{k}^2}\right] . \tag{A9}$$

(3) Only the static ($n=0$) Faddeev-Popov ghost survives. Its propagator in O(3)-covariant gauge is

$$iG^{ab}(\vec{k}) = -i\delta^{ab}\zeta/k^2 , \tag{A10}$$

and it couples only to the static A_i propagator [Eq. (A9)], with strength g/ζ.

The static gauge is a natural choice for infrared studies at finite temperature because, from the point of view of the effective three-dimensional theory, it is physical. This can be seen by first noting that the four-dimensional finite-temperature gauge field describes 2 physical degrees of freedom for each value of n. For $n\neq 0$, they correspond to the 2 degrees of freedom of the massive three-vector field $A_i^{(n\neq 0)}$. The propagator for this field appears automatically in the physical, Proca-Wentzel, from [Eq. (A8)]. For $n=0$, the 2 degrees of freedom are accounted for by the massless gauge field $A_i^{(n=0)}$ and the three-dimensional scalar A_4. Each describes 1 physical degree of freedom.

Note added. After this work was completed, we became aware of a paper by An Ing and Chen Shi [proceedings of the Third Marcel Grossman Meeting (unpublished)] in which the same effective potential is computed, for related although somewhat different reasons. We thank E. Witten for bringing this work to our attention.

ACKNOWLEDGMENTS

We thank our colleagues at Yale for many useful comments and for their encouragement. We are especially grateful to Steve Detweiler and Vince Moncrief. We acknowledge the hospitality of the Lewes Center for Physics where some of this work was done. Research was supported in part by the U. S. Department of Energy under Contract No. DE-AC02-76ER03075.

---

[1]Th. Kaluza, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl. 966 (1921).

[2]O. Klein, Z. Phys. 37, 895 (1926).

[3]B. DeWitt, in *Lectures at 1963 Les Houches School, Relativity, Groups and Topology,* edited by B. Dewitt and C. DeWitt (Gordon and Breach, New York, 1964), published separately under the title *Dynamical Theory of Groups and Fields* (Gordon and Breach, New York, 1965); R. Kerner, Ann. Inst. Henri Poincaré 9, 143, (1968), A. Trautman, Rep. Math. Phys. 1, 29 (1970); Y. M. Cho, J. Math. Phys. 16, 2029 (1975); Y. M. Cho and P. G. O. Freund, Phys. Rev. D 12, 1711 (1975); Y. M. Cho and P. S. Jang, *ibid.* 12, 3789 (1975); L. N. Chang, K. I. Macrae, and F. Mansouri, *ibid.* 13, 235 (1976).

[4]C. A. Orzalesi, Fortschr. Phys. 29, 413 (1981); E. Witten, Nucl. Phys. B186, 412 (1981); C. Wetterich, Phys. Lett. 113B, 377 (1982); S. D. Unwin, *ibid.* 103B, 18 (1981); P. G. O. Freund and M. A. Rubin, *ibid.* 97B, 233 (1980).

[5]E. Cremmer and B. Julia, Phys. Lett. 80B, 48 (1978); Nucl. Phys. B159, 141 (1979).

[6]T. Appelquist and A. Chodos, Phys. Rev. Lett. 50, 141 (1983).

[7]H. B. G. Casimir, Proc. K. Ned. Akad. Wet. 51, 793 (1948).

[8]T. Appelquist and R. Pisarski, Phys. Rev. D 23, 2305 (1981); S.

Nadkarni, *ibid.* 27, 917 (1983).
[9]A. Chodos and S. Detweiler, Phys. Rev. D 21, 2167 (1980).
[10]J. Scherk and J. H. Schwarz, Nucl. Phys. B153, 61 (1979).
[11]C. Brans and R. H. Dicke, Phys. Rev. 124, 925 (1961); see also S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972), Sec. 7.3.
[12]V. Moncrief, private communication.
[13]J. F. Luciani, Nucl. Phys. B135, 111 (1978).
[14]C. A. Orzalesi and M. Pauri, Phys. Lett. 107B, 186 (1981).
[15]C. Wetterich, Ref. 4.
[16]B. S. DeWitt, Phys. Rev. 162, 1195 (1967); 162, 1239 (1967).
[17]G. 't Hooft and M. Veltman, Ann. Inst. Henri Poincaré 20, 69 (1974).
[18]S. Deser and P. van Nieuwenhuizen, Phys. Rev. D 10, 401 (1974); 10, 411 (1974).
[19]A. Salam and J. Strathdee, Ann. Phys. (N.Y.) 141, 316 (1982), Appendix 5.
[20]E. S. Fradkin and G. A. Vilkovisky, Phys. Rev. D 8, 4241 (1973).
[21]B. S. DeWitt, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, London, 1979), especially Sec. 14.5.2 and pp. 734 and 735.
[22]G. 't Hooft, in *Recent Developments in Gravitation* (Plenum, New York, 1979), especially Sec. 3.
[23]R. Jackiw, Phys. Rev. D 9, 1686 (1974).
[24]M. A. Rubin and B. D. Roth, University of Texas Report No. UTTG-2-83, (unpublished).
[25]T. Appelquist, A. Chodos, and E. Myers, Phys. Lett. (to be published).
[26]M. Duff and C. A. Orzalesi, Phys. Lett. 122B, 37 (1983).
[27]E. Witten, Ref. 4.
[28]M. J. Duff, CERN Report No. TH3451-CERN, 1982 (unpublished) and references therein.
[29]P. G. O. Freund, Nucl. Phys. B209, 146 (1982).