

The $SU(5)$ Potential in DeSitter Space

BRUCE ALLEN

*Department of Physics, University of California,
Santa Barbara, California 93106*

Received July 13, 1984

The one-loop effective potential for a minimal $SU(5)$ theory is calculated on a curved DeSitter background spacetime. The stability of its extrema in the following subgroups is investigated: $SU(4) \times U(1)$, $SU(3) \times SU(2) \times U(1)$, $SU(3) \times U(1) \times U(1)$, $SU(2) \times SU(2) \times U(1) \times U(1)$. A combination of analytic and numerical methods is used to obtain phase diagrams for the model. In the inflationary universe, the curvature effects do not prevent a slide into the $SU(4) \times U(1)$ extremum. © 1985 Academic Press, Inc.

1. INTRODUCTION

This is the second of two papers about vacuum energy in DeSitter space. The first paper [1] showed how the vacuum energy could be explicitly calculated for any gauge theory in DeSitter space, and illustrated the method with a simple $U(1)$ gauge theory. In this paper, the methods of the first paper are applied to a more complicated and realistic theory, the $SU(5)$ gauge theory.

The inflationary model of the early universe [2] predicts that the stars and all the matter about us resulted from a phase transition soon after the big bang. During this phase transition, real observable matter was formed from latent vacuum energy, and the equation of state of the universe changed. Also, and most relevant to this paper, a symmetry that existed between different types of elementary particles disappeared.

A phase transition probably took place in the very early universe, because a common feature of gauge models of the fundamental interactions is that they have simple high energy behavior. At low temperatures these theories describe a veritable catalog of different particles and interactions. However, at high temperatures, the more fundamental underlying symmetries appear. These fundamental symmetries are described by a gauge group. The process in which these fundamental symmetries disappear as the energy scale decreases is called spontaneous symmetry breaking. If the early universe was very hot, then symmetry breaking must have taken place as it cooled.

Suppose that we are given a model of the fundamental interactions in which symmetry breaking can take place in several ways. We can assume that the universe began in the most symmetric (high-temperature) phase. If the model of the fun-

damental interactions contains free parameters (i.e., coupling constants, masses, etc.) then for different choices of these parameters, the universe will evolve to different broken-symmetry phases as it cools and expands. In short, it is possible to determine the phase to which the universe eventually evolves if all the parameters are specified.

The simple way to study this question is with a potential function. The different broken-symmetry phases of a gauge theory are characterized by different values of the fields in that theory. The potential is a function of the fields, which associates to each broken-symmetry phase an energy density (energy/3-volume). Typically, each broken-symmetry phase corresponds to a local minimum of the potential function. The problem is just like the mechanical problem of a ball rolling on a hill. We start the universe off in one phase, and find out where it ends up.

An interesting study of this sort has been carried out by Breit, Gupta and Zaks [3]. They discovered that the Coleman–Weinberg $SU(5)$ theory, on which the new inflationary universe models are based, had a serious flaw. The universe always evolved from the $SU(5)$ symmetric phase toward the wrong broken-symmetry phase, the $SU(4) \times U(1)$ phase. This is because the potential function has a ridge that separates the $SU(4) \times U(1)$ phase from the desired $SU(3) \times SU(2) \times U(1)$ phase. This ridge means that the universe can only evolve into the $SU(3) \times SU(2) \times U(1)$ phase by tunnelling into it. This would produce a very inhomogeneous universe today, and can be ruled out.

Their analysis used the potential function for a Coleman–Weinberg $SU(5)$ theory, calculated in flat Minkowski spacetime. However, this potential function is not self-consistent, because the vacuum energy gives the spacetime a constant positive curvature. The self-consistent calculation of the $SU(5)$ potential in this constant curvature DeSitter space is the subject of this paper. It is possible to carry out the analysis of Breit, Gupta and Zaks with this curved-space potential. However, such a study would involve the introduction of two additional parameters, and we will argue later that it is likely to give results similar to their flat-space analysis.

The minimal $SU(5)$ theory that we study is described in the second section. Unfortunately this theory has recently been ruled out because the proton lifetime is so long [4]. However, the methods of this paper can be used to calculate the effective potential for any other non-Abelian gauge theory, like $SO(10)$. In the second section, after giving the Lagrangian of the $SU(5)$ theory and its gauge-transformation properties, the one-loop effective potential is calculated symbolically, in terms of functional determinants. The potential depends upon the background Higgs fields via a mass matrix. This mass matrix has a simple physical interpretation in the various broken-symmetry phases.

In the third section, we discuss the broken-symmetry phases of $SU(5)$. Without loss of generality, the background Higgs field is diagonalized, reducing the number of degrees of freedom from 24 to 4. We then choose five phases, or directions in group space, to focus attention on. These five phases correspond to the following subgroups of unbroken symmetries: $SU(5)$, $SU(4) \times U(1)$, $SU(3) \times SU(2) \times U(1)$, $SU(2) \times SU(2) \times U(1) \times U(1)$ and $SU(3) \times U(1) \times U(1)$. In order to facilitate our

later discussions, a notation is introduced which clearly indicates the relative stability or metastability of these five phases.

In the fourth section the curved-space potential, given in terms of symbolic functional determinants, is explicitly evaluated. The results of the first paper [1] are used to express the symbolic determinants in terms of a special function called $A(z)$. Starting in this section, we treat only the Coleman–Weinberg sector of the theory [5]. This is the sector in which the contribution to the potential from closed-scalar Higgs-field loops can be neglected in comparison to the contribution from closed-vector gauge-field loops.

The fifth section contains the most potentially confusing aspect of the paper. The potential up to this point contains an undetermined mass parameter, which was introduced in the path-integral measure. We show how the value of this parameter in curved space is unambiguously fixed by taking the flat-space limit of the calculation. This procedure is essential to Coleman–Weinberg theory, in which a dimensionless parameter is replaced by a dimensional one. To simplify the later sections, we replace the old coupling constants by new linear combinations of them. The new coupling constants are chosen so that the mass introduced by dimensional transmutation is the physical mass M_x of the vector gauge fields in the $SU(3) \times SU(2) \times U(1)$ broken-symmetry phase in flat space.

In the sixth section, by calculating the Higgs fields' masses, we examine the stability of the five broken-symmetry phases. As we said earlier, each of these five phases corresponds to an extremum of the potential function. Now the potential is a function of four variables, and consequently an extremum can be either a minimum, a maximum, or a saddle point. The Higgs fields' masses at an extremum can be used to determine which of these it is. Fortunately, the formula for the Higgs masses have simple flat-space limits. In the case of the symmetric $SU(5)$ phase, all the Higgs fields' masses are the same, and can be expressed in a simple closed form.

In the seventh section, the results of the previous sections are used to generate several phase diagrams which summarize some important information about the nature of symmetry breaking in curved space. Because all of the formula in the preceding sections involved special functions, the results in this last section have been obtained numerically. However, there are several simple results concerning the stability of certain phases which can be obtained analytically.

Throughout this paper, we use units where $\hbar = c = 1$, so that mass = (length) $^{-1}$. The gravitational constant, when used, is defined through the Planck mass $M_p = G^{-1/2}$.

2. THE MINIMAL $SU(5)$ MODEL, AND ITS EFFECTIVE POTENTIAL

The $SU(5)$ theory was discovered by Georgi and Glashow [6]. It is a Yang–Mills gauge model which incorporates strong, weak and electromagnetic forces. One beautiful feature of the theory [7] is that the running gauge-coupling constants of these three fundamental forces all converge to a single value $\approx 1/42$ at an

energy $M_{GUT} \cong 10^{15}$ GeV. This is the energy scale at which the underlying gauge symmetry of the theory is spontaneously restored.

Experimentally, the $SU(5)$ model has not been entirely successful. It correctly predicts the Weinberg–Salam electroweak mixing angle θ_w to within experimental accuracy [4]. However, it also predicts that the proton can decay, with a characteristic lifetime of $\sim 10^{30}$ years. A number of recent experiments have failed to observe such decay, so that the theory has now fallen into disfavor.

In this paper, we only study one sector of the complete $SU(5)$ model. Our interest is high-energy behavior, just before and after the breaking of fundamental $SU(5)$ symmetry. For this reason our model is incomplete. It does not incorporate the 5 of Higgs scalar fields which break electroweak symmetry at much lower energies or fermionic matter fields which become quarks in the low-energy theory. Consequently our $SU(5)$ Lagrangian takes a very simple form.

We give the Lagrangian in its Euclidean form, defined on a spacetime whose metric $g_{\mu\nu}$ has positive signature $(+, +, +, +)$. In field theory it is usually necessary to calculate quantities on an imaginary time manifold, and analytically continue the results back to real Lorentzian time. Since the effective potential function is time independent, we can obtain it directly from a Euclidean calculation. The Euclidean Lagrangian is

$$L = \frac{1}{4} \text{trace}(F_{\mu\nu} F^{\mu\nu}) + \frac{1}{2} \text{trace}(D_\mu \varphi)(D^\mu \varphi) + V_0(\varphi) \tag{2.1}$$

real and positive semi-definite.

There are two fields: a gauge vector field A^μ and a Higgs scalar field φ , both in the adjoint representation of $SU(5)$. They will be represented as traceless 5×5 matrices, which can be expanded in terms of an $SU(5)$ Lie algebra basis. We will denote the 24 basis elements by λ_a . We use Latin letters as $SU(5)$ group indices running from 1 to 24 and Greek letters as spacetime indices running from 0 to 3. The λ_a are a linearly independent set of 24 traceless 5×5 matrices, which are self-adjoint, so $\lambda_a^\dagger = \lambda_a$, and orthonormal, so $\text{trace}(\lambda_a \lambda_b) = \frac{1}{2} \delta_{ab}$. The Hermitian adjoint operation \dagger is complex-conjugation followed by matrix transposition. The matrices form a real adjoint representation of $SU(5)$, and the gauge field and Higgs field can be expanded in terms of them

$$A^\mu(x) = A_a^\mu(x) \lambda_a, \quad \varphi(x) = \varphi_a(x) \lambda_a. \tag{2.2}$$

Throughout this paper, the summation convention applies to group-space indices as well as spacetime ones. The component fields $A_a^\mu(x)$ and $\varphi_a(x)$ are real, so that $\varphi^\dagger = \varphi$ and $A_\mu^\dagger = A_\mu$.

The Euclidean Lagrangian is real, because the field tensor $F_{\mu\nu}$ and gauge-covariant derivative $D_\mu \varphi$ are both self-adjoint.

$$F_{\mu\nu} = F_{\mu\nu}^\dagger = \nabla_\mu A_\nu - \nabla_\nu A_\mu - ig[A_\mu, A_\nu] \tag{2.3}$$

$$D_\mu \varphi = (D_\mu \varphi)^\dagger = \partial_\mu \varphi - ig[A_\mu, \varphi]. \tag{2.4}$$

The dimensionless coupling constant g must be real, since each commutator is anti-self-adjoint, for example, $[A_\mu, \varphi]^\dagger = -[A_\mu, \varphi]$. The tree potential $V_0(\varphi)$ must also be real. In general it can be a quartic polynomial of the form

$$V_0(\varphi) = a \text{ trace } \varphi^2 + b \text{ trace } \varphi^3 + A_2(\text{trace } \varphi^2)^2 + A_4 \text{ trace } \varphi^4. \quad (2.5)$$

There are no terms of the form $(\text{trace } \varphi)^2$ or $(\text{trace } \varphi) \text{ trace } \varphi^3$, because $\text{trace } \varphi \equiv 0$. The subscript in V_0 indicates that it is the “zero-loop” or “tree” potential.

In this paper, we use two potential functions, the tree potential $V_0(\varphi)$ and the one-loop effective potential. The one-loop effective potential incorporates lowest-order quantum fluctuation effects. For the sake of economy, we will refer to it simply as “the potential” and denote it by $V(\varphi)$.

Before proceeding, we are going to restrict the form (2.5) of the tree potential. In order that the action be invariant under $\varphi \rightarrow -\varphi$, we set $b \equiv 0$, so that $L(A^\mu, \varphi) = L(A^\mu, -\varphi)$. We are also going to restrict the form of a . In general, it is of the form $a = \frac{1}{2}(m^2 + \xi R)$, where m^2 is the mass² of the Higgs field in flat-space in the $SU(5)$ symmetric phase, ξ is a number and R is the scalar curvature of the spacetime manifold. In order to permit the Coleman–Weinberg mechanism to operate, we set $m^2 \equiv 0$. The Lagrangian now contains *no* dimensional parameters. The masses of the Higgs field and gauge field will be spontaneously generated by radiative corrections, which are the source of symmetry breaking.

In flat space, where $R = 0$, the $a \text{ trace } \varphi^2$ term in the tree potential vanishes. One might suppose that the choice of $\xi = 0$ is natural, since it appears to remove any direct coupling between the Higgs field and the gravitational field. However, it will turn out that this term is generated by one-loop corrections. Even if one sets $\xi = 0$ in the tree potential, we will see that a nonzero value of ξ is induced by one-loop effects. This means that this term *must* be included in the bare Lagrangian, if one is to have a sensible renormalizable theory.

The Lagrangian is invariant under the action of local $SU(5)$ transformations. Suppose that $\rho(x)$ is a 5×5 matrix which satisfies $\rho^\dagger(x) \rho(x) = I$ so that it is an element of $SU(5)$. Under the action of ρ , the fields are transformed to

$${}^\rho A^\mu = \rho A^\mu \rho^{-1} - ig^{-1}(\nabla^\mu \rho) \rho^{-1} \quad (2.6)$$

$${}^\rho \varphi = \rho \varphi \rho^{-1} \quad (2.7)$$

and it is straightforward to show that $L({}^\rho A^\mu, {}^\rho \varphi) = L(A^\mu, \varphi)$. The unitary condition on ρ ensures that the fields remain self-adjoint.

In the remainder of this section, we are going to derive a symbolic expression for the one-loop effective potential (hereafter called “the potential”). The method is identical to the one used for a simpler $U(1)$ gauge theory in the first paper [1]. We use a four-sphere of radius a as the Euclidean background spacetime manifold. This space of constant positive curvature ($R = 12/a^2$) replaces flat Minkowski space ($R = 0$) in the presence of a constant background energy density. Of course, in the

limit as the potential $V \rightarrow 0$, the radius $a \rightarrow \infty$ and the curvature $R \rightarrow 0$. However, we will ignore the relationship between the potential and the curvature, and simply regard the curvature as an independent parameter.

To calculate the potential to one-loop order, we expand the action around a constant background Higgs field φ_0

$$\varphi = \varphi_0 + \tilde{\varphi} \tag{2.8}$$

and keep only the terms quadratic in the fluctuating fields $\tilde{\varphi}$ and A^μ . To this order, the Lagrangian is

$$\begin{aligned} L = & \frac{1}{4} \text{trace } F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} \text{trace}(D^\mu \tilde{\varphi})(D_\mu \tilde{\varphi}) + ig \text{trace}(\nabla_\mu A^\mu)[\varphi_0, \tilde{\varphi}] \\ & + \frac{1}{2} M_{ab}^2 A^\mu{}_a A_{\mu b} + V_0(\varphi_0) + \tilde{\varphi}_\alpha \left. \frac{\partial V_0}{\partial \varphi_a} \right|_{\varphi_0} + \frac{1}{2} \tilde{\varphi}_a \tilde{\varphi}_b \left. \frac{\partial^2 V_0}{\partial \varphi_a \partial \varphi_b} \right|_{\varphi_0}. \end{aligned} \tag{2.9}$$

The fourth term is especially important to us, because it is the mass term for the gauge fields. The mass-matrix M_{ab}^2 is a function of the background Higgs field φ_0 .

$$M_{ab}^2(\varphi_0) \equiv -g^2 \text{trace}[\lambda_a, \varphi_0][\lambda_b, \varphi_0] \tag{2.10}$$

It is a symmetric 24×24 matrix, with real nonnegative eigenvalues m_e^2 , which are the masses² of the gauge fields A^μ .

Now we have to choose a gauge for our calculation. 'tHooft's background field gauge provides a convenient gauge-fixing term. It is

$$L_{\text{gauge}} = \frac{1}{2} \alpha \text{trace}(\nabla_\mu A^\mu - ig\alpha^{-1}[\varphi_0, \varphi])^2 \tag{2.11}$$

where α is a positive real number. The crossterm in (2.11) cancels the third term in the one-loop Lagrangian (2.9). The resulting total Lagrangian $L_{\text{total}} = L + L_{\text{gauge}}$ is

$$\begin{aligned} L_{\text{total}} = & \frac{1}{2} A_T^{\mu a} [\delta_{ab}(-g_{\mu\nu} \square + R_{\mu\nu}) + g_{\mu\nu} M_{ab}^2] A_T^{\nu b} \\ & + \frac{1}{2} \alpha A_L^{\mu a} [-\delta_{ab} \nabla_\mu \nabla_\nu + \alpha^{-1} g_{\mu\nu} M_{ab}^2] A_L^{\nu b} \\ & + \frac{1}{2} \tilde{\varphi}^a \left[-\delta_{ab} \square + \alpha^{-1} M_{ab}^2 + \left. \frac{\partial^2 V_0}{\partial \varphi_a \partial \varphi_b} \right|_{\varphi_0} \right] \tilde{\varphi}_b. \end{aligned} \tag{2.12}$$

The gauge field has been decomposed into transverse and longitudinal parts

$$A^\mu = A_L^\mu + A_T^\mu. \tag{2.13}$$

They satisfy $\nabla_\mu A_T^\mu = 0$ and $A_L^\mu = \partial^\mu \chi$, where χ is some scalar function. The transverse part A_T^μ is the irreducible spin-1 "physical part" of the gauge field, whereas the longitudinal component A_L^μ is a spin-0 artifact.

The one-loop Lagrangian includes only terms quadratic in the fields. The constant term $V_0(\varphi_0)$, and the linear term $\tilde{\varphi}_a(\partial V_0/\partial\varphi_a)|_{\varphi_0}$ have both been dropped from (2.12). The constant term can be restored later. This is because the potential is the sum of the tree potential and the one-loop contribution, which we will calculate from (2.12). The linear term is cancelled by introducing a current $J_a = -\partial V_0/\partial\varphi_a$, which couples to the scalar field. This current allows us to “hold” the Higgs field fixed at nonequilibrium values of φ_0 . The action density then equals the potential, because there is no kinetic energy due to a changing field.

The gauge-fixing term (2.11) introduces a factor which is called the Fadeev–Popov ghost. This factor Δ must be included in the path integral in order to make the measure independent of our choice of a gauge-fixing term. Fortunately there is a simple procedure [8] for finding Δ .

The gauge-fixing term L_{gauge} “damps out” the path integral unless it is nearly zero. We are going to examine the effects of gauge transformations about this point. To begin, define a 5×5 matrix Q with components Q_a

$$Q \equiv Q_a \lambda_a = \alpha^{1/2} (\nabla_\mu A^\mu - ig\alpha^{-1} [\varphi_0, \varphi]) \quad (2.14)$$

and choose fields A^μ, φ for which $Q \equiv 0$. Now suppose a small gauge transformation $\rho = I + \varepsilon_a \lambda_a$ acts on the fields. Then to first order in the small parameter ε ,

$$Q_a = ig^{-1} \alpha^{1/2} (-\delta_{ab} \square + X_{ab}^\mu \partial_\mu + \alpha^{-1} Y_{ab}) \varepsilon_b, \quad (2.15)$$

where the matrices X_{ab}^μ and Y_{ab} are

$$X_{ab}^\mu = ig \text{trace} [\lambda_b, \lambda_a] A^\mu \quad (2.16)$$

$$Y_{ab} = -g^2 \text{trace} [\varphi, \lambda_a] [\varphi, \lambda_b]. \quad (2.17)$$

Because the trace of Q^2 is the gauge-fixing term, it turns out that the Fadeev–Popov factor Δ is the determinant of the second-order operator in (2.15).

In the one-loop approximation X_{ab}^μ does not contribute, and Y_{ab} reduces to the mass matrix M_{ab}^2 , defined in (2.10). The ghost factor becomes

$$\Delta = \text{Det} \alpha^{1/2} \mu^{-2} (-\delta_{ab} \hat{\square} + \alpha^{-1} M_{ab}^2) \text{Det} \left[\frac{\alpha^{-1/2} \mu^{-2} M_{ab}^2}{\text{Erf}(\alpha^{-1/2} g^{-1} a^2 M_{ab}^2)} \right]. \quad (2.18)$$

The first factor is a symbolic functional determinant on scalars, and the hat indicates that its zero-mode is to be omitted. The second factor comes from these zero-modes. In an earlier paper [1] we showed that the bounded integration over the zero-mode of the gauge group yields an error function, because the Euclidean manifold S^4 is compact. This second determinant is just an ordinary product of 24 eigenvalues $m_e^2/\text{Erf}(m_e^2)$. The regularization mass μ comes from the measure of the integration over all gauge transformations.

Putting together the quadratic parts of the Lagrangian (2.12) and the Fadeev–Popov Jacobian Δ (2.18) we arrive at the following expression for the potential.

$$\begin{aligned}
 V(\varphi_0) = & V_0(\varphi_0) - \Omega^{-1} \log \text{Det} \left[\frac{\alpha^{-1/2} \mu^{-2} M_{ab}^2}{\text{Erf}(\alpha^{-1/2} g^{-1} a^2 M_{ab}^2)} \right] \\
 & + \Omega^{-1} \left[\frac{1}{2} \log \text{Det} \mu^{-2} S_T + \frac{1}{2} \log \text{Det} \mu^{-2} S_L \right. \\
 & \left. + \frac{1}{2} \log \text{Det} \mu^{-2} S_H - \log \text{Det} \mu^{-2} S_G \right] \tag{2.19}
 \end{aligned}$$

where $\Omega = (8/3) \pi^2 a^4$ is the spacetime 4-volume. The second-order differential operators which appear in the determinants are

$$\begin{aligned}
 \text{Transverse} \quad S_T &= \delta_{ab} (-g_{\mu\nu} \square + R_{\mu\nu}) + g_{\mu\nu} M_{ab}^2 \\
 \text{Longitudinal} \quad S_L &= \alpha (-\delta_{ab} \nabla_\mu \nabla_\nu + \alpha^{-1} g_{\mu\nu} M_{ab}^2) \\
 \text{Higgs} \quad S_H &= -\delta_{ab} \square + \alpha^{-1} M_{ab}^2 + \left. \frac{\partial^2 V_0}{\partial \varphi_a \partial \varphi_b} \right|_{\varphi_0} \\
 \text{Ghost} \quad S_G &= \alpha^{1/2} (-\delta_{ab} \hat{\square} + \alpha^{-1} M_{ab}^2). \tag{2.20}
 \end{aligned}$$

These operators act on representations which are adjoints of $SU(5)$ and scalars or vectors of $SO(5)$.

We now specialize to Landau gauge by sending $\alpha \rightarrow \infty$. Because of its linearity for small x , $\text{Erf}(x) \sim 2\pi^{-1/2} x + O(x^2)$, the ordinary determinant in (2.19) becomes $(g^{-1} a^2 \mu^2)^{24}$. Several of the mass terms in (2.20) vanish, and the ghost determinant cancels the longitudinal one. The final result is that the potential is

$$\begin{aligned}
 V(\varphi_0) = & V_0(\varphi_0) + \frac{1}{2\Omega} \log \text{Det} \mu^{-2} [\delta_{ab} (-g_{\mu\nu} \square + R_{\mu\nu}) + g_{\mu\nu} M_{ab}^2(\varphi_0)] \\
 & + \frac{1}{2\Omega} \log \text{Det} \mu^{-2} \left[-\delta_{ab} \square + \left. \frac{\partial^2 V_0}{\partial \varphi_a \partial \varphi_b} \right|_{\varphi_0} \right]. \tag{2.21}
 \end{aligned}$$

Terms like $a^{-4} \log(g^{-1} a^2 \mu^2)^{24}$ which do not depend upon φ_0 have been dropped. These terms contribute to the conformal anomaly, and hence determine the value of $V(\varphi_0)$ when $\varphi_0 \equiv 0$. However, they do not affect the phase structure of the potential, which we will be studying.

Basically the final result is straightforward. The one-loop $SU(5)$ potential is simply the sum of 24 one-loop potentials for scalar electrodynamics. The gauge-field's masses are determined by the 24 eigenvalues m_e^2 of $M_{ab}^2(\varphi_0)$ and the Higgs field's masses are determined by the 24 eigenvalues of $(\partial^2 V_0 / \partial \varphi_a \partial \varphi_b)|_{\varphi_0}$. In later sections, we will explicitly evaluate this symbolic expression for the potential in DeSitter space, and study some of its properties.

3. BROKEN-SYMMETRY PHASES OF $SU(5)$

If the vacuum energy density $V(\varphi)$ is minimized by a nonvanishing Higgs field φ , then the field can remain stable at that value. (For clarity from this point on, we will denote the classical background Higgs field by φ rather than by φ_0 .) In general, such a Higgs field will break some, but not all, of the $SU(5)$ symmetry. This means that under the action of a general local $SU(5)$ gauge transformation, the Higgs field will be changed. However, there will be a subgroup of gauge transformations, contained in $SU(5)$, for which the Higgs field is invariant, so ${}^\rho\varphi = \varphi$. Each such distinct subgroup will be referred to as a broken-symmetry phase of the $SU(5)$ theory. As we will see, the macroscopic physical properties of the theory, like the number of massive fields, depend upon which phase one is in.

To classify some of the different phases, we will begin by examining the mass matrix (2.10). The potential (2.21) depends upon φ through (1) the mass matrix $M_{ab}^2(\varphi)$ and (2) the tree-potential $V_0(\varphi)$. We will show shortly that the tree potential $V_0(\varphi)$ can be expressed as a function of the mass matrix M_{ab}^2 . This means that the potential depends upon the Higgs field only through the mass matrix. In fact it is often convenient to regard the potential as a function of the mass matrix and not as a function of the Higgs field.

In general, the Higgs field φ is a traceless Hermitian 5×5 matrix. This means that we can always find an $SU(5)$ gauge transformation ρ which makes ${}^\rho\varphi$ diagonal, traceless and real. This transformation does not change the tree potential since $V_0(\varphi) = V_0({}^\rho\varphi)$. It also does not change the eigenvalues of M_{ab}^2 . So, without loss of generality, we can assume that φ is diagonal,

$$\varphi \equiv \text{diag}(\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5) \quad (3.1)$$

and traceless $\varphi_1 + \varphi_2 + \varphi_3 + \varphi_4 + \varphi_5 \equiv 0$. The potential is really a function of four independent variables, not 24.

The mass matrix can be expressed in terms of the new variables φ_i (in this section, let i, j, k, \dots , run from 1 to 5). We are only interested in finding the eigenvalues of $M_{ab}^2(\varphi)$ since they are the physical masses² of the 24 gauge fields A^μ . To calculate the eigenvalues, it is convenient to use an $SU(5)$ basis λ_a which diagonalizes M_{ab}^2 . In this basis, four of the λ_a 's are real, diagonal and traceless. Ten of the λ_a 's are completely zero, except for two real, equal, off-diagonal entries. The remaining ten λ_a 's are completely zero, except for two imaginary, opposite, off-diagonal entries. In this basis, M_{ab}^2 is diagonal, and its eigenvalues turn out to be

$$\text{Eigenvalues of } M_{ab}^2(\varphi) = m_e^2 = \frac{1}{2} g^2 (\varphi_i - \varphi_j)^2 \quad (3.2)$$

where i and j range independently from 1 to 5. It is clear from (3.2) that of the 24 eigenvalues, at least four are exactly zero. This is because any choice of φ_i leaves unbroken a residual $U(1) \times U(1) \times U(1) \times U(1)$ symmetry.

Let us express the tree potential $V_0(\varphi)$ in terms of the mass matrix. The trick is to do it in the basis for which M_{ab}^2 is diagonal. The result can then be written in a basis-independent way. For instance, the trace of M_{ab}^2 in our diagonal basis is

$$\begin{aligned} \text{trace } M^2 &= \sum_e m_e^2 = \frac{1}{2} g^2 \sum_{i,j=1}^5 (\varphi_i - \varphi_j)^2 \\ &= \frac{1}{2} g^2 (10 \text{ trace } \varphi^2 - 2(\text{trace } \varphi)^2) = 5g^2 \text{ trace } \varphi^2 \end{aligned} \tag{3.3}$$

since $\text{trace } \varphi = 0$. In a similar fashion, the trace of $M_{ab}^4 = M_{ac}^2 M_{cb}^2$ is

$$\begin{aligned} \text{trace } M^4 &= \sum_e m_e^4 = \frac{1}{4} g^4 \sum_{i,j=1}^5 (\varphi_i - \varphi_j)^4 \\ &= \frac{1}{2} g^4 (5 \text{ trace } \varphi^4 + 3(\text{trace } \varphi^2)^2). \end{aligned} \tag{3.4}$$

Since the tree potential (2.5) is a function only of $\text{trace } \varphi^2$ and $\text{trace } \varphi^4$, it can be expressed as a function of these traces of the mass matrix.

Where are the extrema of the potential? This problem has been analyzed by Kim [9–12], who studied the most general possible gauge-invariant potentials which were quartic polynomials in the fields. In DeSitter space, the one-loop corrections render the potential nonpolynomial. However, some of Kim’s results still apply, since the potential is gauge invariant.

Kim’s analysis is based on the following insight. The orbits of φ under the action of the group can be specified by four orbit parameters: a modulus $\|\varphi\| = \text{trace } \varphi^2$ and three “angles” $\theta_1 = \text{trace } \varphi^3 / \|\varphi\|^{3/2}$, $\theta_2 = \text{trace } \varphi^4 / \|\varphi\|^2$ and $\theta_3 = \text{trace } \varphi^5 / \|\varphi\|^{5/2}$. It turns out that, for a given modulus, the orbit space $(\theta_1, \theta_2, \theta_3) \in \mathcal{R}^3$ is compact. It looks like a solid tetrahedron with inward-sagging faces and inward-sloping edges. The cusps at the vertices are stationary points under the action of the maximal little groups of $SU(5)$, and the faces are stationary points under the action of the maxi-maximal little groups of $SU(5)$. When the potential is a quartic polynomial, one can show that its extrema lie on the boundary of the orbit-space. For that case, the potential is monotonic as one moves outward in orbit space, and hence the extrema lie on the vertices.

For this reason, we will concentrate our attention on four special directions in group space. They are the two maximal little groups, and the two maxi-maximal little groups of $SU(5)$. Bocharov et al. [13–15] have studied the most general possible potential which is a function of the mass matrix M_{ab}^2 , and discovered that it must have extrema in these four directions. It is clear from Kim’s analysis why this is so. The potential is a function of the orbit parameters, and they are extremized in these four directions. While we have no guarantee that all the extrema must lie on the boundary of the orbit space, there is no evidence that any other extrema exist.

TABLE I
The Five Phases of $SU(5)$ Which We Studied

Phase	Invariant subgroup	Dimension	Higgs direction φ
1	$SU(3) \times SU(2) \times U(1)$	12	(2, 2, 2, -3, -3)
2	$SU(4) \times U(1)$	16	(1, 1, 1, 1, -4)
3	$SU(3) \times U(1) \times U(1)$	10	(0, 0, 0, 1, -1)
4	$SU(2) \times SU(2) \times U(1) \times U(1)$	8	(1, 1, -1, -1, 0)
5	$SU(5)$	24	(0, 0, 0, 0, 0)

Note. The five phases of $SU(5)$ correspond to maximal and maxi-maximal little groups of $SU(5)$. The dimension of a subgroup is the number of unbroken symmetries. The Higgs direction preserves only the listed subgroup of symmetries.

Shown in Table I are the phases of $SU(5)$ which we will focus on. They will be called phases 1, 2, 3, 4 and 5. Phase 5 is the unbroken symmetry $SU(5)$ phase. Phases 1 and 2 are maximal little groups and phases 3 and 4 are maxi-maximal little groups. Also shown is the diagonal Higgs field direction, which extremizes the potential in each phase. For notational convenience, we will often use $A = 1, \dots, 5$ to denote one of these phases.

In each phase, the number of gauge fields which become massive equals the number of broken symmetries. The number of distinct masses can be determined from the eigenvalues of the mass-matrix M_{ab}^2 .

For later use, we are going to define a set of integer constants C_κ^A which contain information about the different phases. There are six constants ($\kappa = 1, \dots, 6$) for each phase ($A = 1, \dots, 5$). They are shown in Table II. In each phase, C_1 is the trace of φ^2 and C_2 is the trace of φ^4 , for the φ given in Table I. The remaining constants pertain to the eigenvalues of the mass-matrix M_{ab}^2 . C_4 and C_6 are the nonzero values of $(\varphi_i - \varphi_j)^2$, and C_3 and C_6 are their respective degeneracies. For example, in the third row, the values of C_3 to C_6 tell you that the third phase has 12 "light" gauge

TABLE II
A Set of Six Integer Constants C_κ^A

C_κ^A	$\kappa = 1$	2	3	4	5	6
$A = 1$	30	210	12	25	0	-
2	20	260	8	25	0	-
3	2	2	12	1	2	4
4	4	4	8	1	8	4
5	0	0	0	-	0	-

Note. A denotes one of the five broken symmetry phases. C_1 and C_2 are trace φ^2 and trace φ^4 . The mass-matrix has C_3 degenerate eigenvalues C_4 and C_5 degenerate eigenvalues C_6 . The factor of $\frac{1}{2} g^2$ in the masses and the overall dimensions have been left out.

fields with $\text{mass}^2 = 1$, and 2 “heavy” gauge fields with $\text{mass}^2 = 4$. These constants will enable us to treat all the phases in a similar way.

Of the six constants, only four are independent. The relations (3.3) and (3.4) between trace φ^2 , trace φ^4 and the eigenvalues of the mass matrix imply similar relations among the C_k^A . For a given phase A , they are

$$10C_1^A = C_3^A C_4^A + C_5^A C_6^A \tag{3.5}$$

for the trace of φ^2 , and

$$10C_2^A + 6(C_1^A)^2 = C_3^A (C_4^A)^2 + C_5^A (C_6^A)^2 \tag{3.6}$$

for the trace of φ^4 . These relations will be used later.

The potential function has extrema in all five of the group space directions listed in Table I. At these extrema, the potential function can be a maximum, minimum or saddle point. In order for a phase to be stable, the extremum corresponding to it must be a local minimum. In general, some but not all of the phases will be stable.

Of the stable phases, some will lie lower on the potential hill than others. For convenience in discussing such matters, we are going to introduce “stability codes,” which provide a convenient notation for discussing such matters. A simple example shows how they work. Suppose that phases 1, 3 and 5 are minima, and that the other two phases are unstable saddle points. Also suppose that the potential hill is highest at phase 1 and lowest at phase 5.

$$\begin{aligned} V(\text{phase 5}) < V(\text{phase 3}) < V(\text{phase 1}) \\ \text{phases 2 and 4 not local minima.} \end{aligned} \tag{3.7}$$

Then the stability code corresponding to this configuration is 531. Each digit in the code corresponds to a stable phase: the leftmost digit is the lowest minimum, and the rightmost digit is the highest minimum. The order of the digits corresponds to an increasing value of the potential. In the example, phase 3 is called metastable because it can decay via tunnelling or barrier penetration to phase 5. The lowest extrema is called the stable phase, and the other local minima are called metastable phases.

4. THE EFFECTIVE POTENTIAL IN DESITTER SPACE

In this section, we will obtain an explicit form for the potential. In the second section, it was expressed symbolically, as a functional determinant. These functional determinants can be defined on S^4 by using generalized zeta functions. Our first paper [1] evaluated these determinants in terms of the psi function $\psi(z) = (d/dz) \log \Gamma(z)$. Those same results can be applied here.

The potential (2.21) is the sum of three terms. The first term is the tree potential, and the last terms are the one-loop contributions to the vacuum energy. The second

term comes from closed gauge-field loops, and the third term comes from closed scalar-field loops.

It is now well established that a viable inflationary universe model [2] can only result from a potential which is unnaturally flat at $\varphi = 0$. For this reason we will concentrate on the Coleman–Weinberg sector of $SU(5)$, where this is the case. In this sector, g^4 is of the same order as A_2 and A_4 , which means that the one-loop contribution of the scalar-field can be neglected in comparison to the one-loop contribution of the gauge-field [5].

To express the remaining determinant we introduce a special function $A(z) = \zeta'(1, \frac{1}{4} - z) - \zeta'(1, \frac{1}{4})$ defined in the notation of [1]. In terms of the psi function, it is

$$A(z) = \frac{1}{4}z^2 + \frac{1}{3}z - \int_2^{3/2 + (1/4 - z)^{1/2}} t(t - \frac{3}{2})(t - 3) \psi(t) dt - \int_1^{3/2 - (1/4 - z)^{1/2}} t(t - \frac{3}{2})(t - 3) \psi(t) dt. \tag{4.1}$$

We will also make use of the first two derivatives of $A(z)$

$$A'(z) = -\frac{1}{2}(z + 2)[\psi(\frac{3}{2} + (\frac{1}{4} - z)^{1/2}) + \psi(\frac{3}{2} - (\frac{1}{4} - z)^{1/2})] + \frac{1}{2}z + \frac{1}{3} \tag{4.2}$$

$$A''(z) = \frac{1}{2} - \frac{1}{2}[\psi(\frac{3}{2} + (\frac{1}{4} - z)^{1/2}) + \psi(\frac{3}{2} - (\frac{1}{4} - z)^{1/2})] + \frac{1}{4}(\frac{1}{4} - z)^{-1/2} (z + 2)[\psi'(\frac{3}{2} + (\frac{1}{4} - z)^{1/2}) - \psi'(\frac{3}{2} - (\frac{1}{4} - z)^{1/2})] \tag{4.3}$$

where $\psi'(z) = (d/dz) \psi(z)$ is the trigamma function. The following expressions are valid, for large z .

$$A(z) \cong -(\frac{1}{4}z^2 + z + \frac{19}{30}) \log z + \frac{3}{8}z^2 + z + \text{constant} + O(z^{-1}) \tag{4.4}$$

$$A'(z) \cong -\frac{1}{2}(z + 2) \log z + \frac{1}{2}z + O(z^{-1}) \tag{4.5}$$

$$A''(z) \cong -\frac{1}{2} \log z + O(z^{-1}). \tag{4.6}$$

The values at the origin will be useful later. They are $A(0) = 0$, $A'(0) = 2\gamma - \frac{2}{3}$ and $A''(0) = \gamma - 1$, where $\gamma = 0.5772$ is Euler's constant.

The symbolic determinant can be neatly expressed in terms of this special function $A(z)$ and the polynomial $P(z) = \frac{1}{4}z^2 + z$, as a sum over the 24 eigenvalues m_e^2 of the mass matrix M_{ab}^2 .

$$\log \text{Det } \mu^{-2} [\delta_{ab}(-g_{\mu\nu} \square + R_{\mu\nu}) + g_{\mu\nu} M_{ab}^2(\varphi)] = - \sum_{e=1}^{24} [A(a^2 m_e^2) + P(a^2 m_e^2) \log(\mu^2 a^2)]. \tag{4.7}$$

For use in the next section, when we study the renormalization of the potential, let us derive the flat-space form of this determinant. To obtain this limit, we send the

radius a of the four-sphere to ∞ . As the volume $\Omega = (8/3) \pi^2 a^4$ becomes infinite, the asymptotic form (4.4) can be used for $A(z)$, in (4.7)

$$\begin{aligned} \lim_{a \rightarrow \infty} \frac{1}{2\Omega} \log \text{Det } \mu^{-2} [\delta_{ab} (-g_{\mu\nu} \square + R_{\mu\nu}) + g_{\mu\nu} M_{ab}^2(\varphi)] \\ = \frac{3}{64\pi^2} \sum_e m_e^4 \left[\log \frac{m_e^2}{\mu^2} - \frac{3}{2} \right] \\ = \frac{3g^4}{256\pi^2} \sum_{i,j=1}^5 (\varphi_i - \varphi_j)^4 \left[\log \frac{g^2(\varphi_i - \varphi_j)^2}{2\mu^2} - \frac{3}{2} \right]. \end{aligned} \tag{4.8}$$

This result is of precisely the same form as the known [7] flat-space results for these determinants.

5. RENORMALIZATION, DIMENSIONAL TRANSMUTATION AND PHYSICAL PARAMETERS

We began with a theory containing no dimensional parameters in flat space. Yet at one-loop a mass μ had to be introduced in order to define a measure for the path integral. This mass is unavoidable—it is present even in flat space (4.7). We started without a dimensional parameter, yet, like a magician pulling a rabbit from his hat, it appeared “from thin air.” This remarkable surprise is at the heart of this section.

The phenomenon is known as dimensional transmutation [7]. It will be familiar to some readers. However, since it is so unusual, we are going to forego several conceptual “shortcuts” and treat it in detail. The most profitable way to begin is by examining the behavior of the potential in flat space.

First of all, regard μ as some definite, fixed, positive mass. We can express the flat-space potential in terms of the 24 eigenvalues m_e^2 of the mass-matrix $M_{ab}^2(\varphi)$.

$$\begin{aligned} V(m_e^2) = \frac{g^{-4}}{25} \left[A_2 - \frac{3}{5} A_4 \right] \left(\sum_e m_e^2 \right)^2 \\ + \frac{3}{64\pi^2} \sum_e m_e^4 \left[\log \frac{m_e^2}{\mu^2} - \frac{3}{2} + \frac{128\pi^2}{15} g^{-4} A_4 \right]. \end{aligned} \tag{5.1}$$

What properties does this potential have? Consider phase 1, the $SU(3) \times SU(2) \times U(1)$ phase. In this phase, the mass-matrix has 12 equal eigenvalues. Suppose the eigenvalues are $m_e^2 = m^2$ (for $e = 1, \dots, 12$) and $m_e^2 = 0$ (for $e = 13, \dots, 24$). Then the potential is

$$V(m^2) = \frac{9}{16\pi^2} m^4 \left[\log \frac{m^2}{\mu^2} + \frac{256}{25} \pi^2 g^{-4} \left(A_2 + \frac{7}{30} A_4 \right) - \frac{3}{2} \right] \tag{5.2}$$

and it has its minimum value when the eigenvalues are

$$m^2 = M_x^2 = \mu^2 \exp \left[1 - \frac{256}{25} \pi^2 g^{-4} \left(A_2 + \frac{7}{30} A_4 \right) \right]. \quad (5.3)$$

It is in this way that the physical masses in the model are “generated” by the regularization mass μ .

The 12 massive gauge fields in phase 1 are called the “X-bosons,” so we have denoted their mass by M_x . From now on, the nonphysical parameter μ^2 which appeared in the curved-space potential will be replaced by $\mu^2 = M_x^2 \exp[(256/25)\pi^2 g^{-4}(A_2 + (7/30)A_4) - 1]$. This expresses the potential in terms of a physical quantity: the mass M_x of the gauge fields in the $SU(3) \times SU(2) \times U(1)$ phase in flat space.

What about the extrema of the other phases in flat space? Although these extrema may be unstable, they allow us to formally define a flat-space gauge-field mass in each phase. We can carry out the same calculation as we have just done, with one small difference. Since phases 3 and 4 have two distinct eigenvalues m_e^2 , define M_A^2 to be the *smaller* eigenvalue in phase A . Then the flat-space gauge field mass in phase A is related to the flat-space gauge field mass in phase 1, by

$$\log \frac{M_A^2}{M_x^2} = \left(\frac{14(C_1^A)^2 - 60C_2^A}{15(C_1^A)^2 + 25C_2^A} \right) A - \frac{32C_3^A \log 2}{C_3^A + 16C_5^A}. \quad (5.4)$$

We have made use of the fact that C_6^A can always be set to 4. This formula must imply that $M_x = M_1$, and it does, since $14(C_1^1)^2 - 60C_2^1 = 0$ and $C_5^1 = 0$.

The quantity A is a dimensionless quartic coupling constant, which is a linear combination of the old quartic coupling constants A_2 and A_4 .

$$A = \frac{64\pi^2}{15} g^{-4} \left(\frac{3}{5} A_4 - A_2 \right). \quad (5.5)$$

It will soon become clear that the curved-space potential depends only upon this particular linear combination of A_2 and A_4 . In the classical theory, one has the two quartic couplings (A_2, A_4) but in the quantum theory they are replaced by (A, M_x). We have traded one dimensionless parameter for a dimensional one. This is the reason behind the name “dimensional transmutation.”

It will turn out in Section 6 that only phase 1 or phase 2 can be stable in flat space. The flat-space potential has the value $-3(C_3^A + 16C_5^A) M_A^4 / 128\pi^2$ at each extremum, which in phases 1 and 2 is

$$V_{SU(3) \times SU(2) \times U(1)} = \frac{-9M_x^4}{32\pi^2}, \quad (5.6)$$

$$V_{SU(4) \times U(1)} = \frac{-3M_x^4}{16\pi^2} \exp \left(\frac{-8}{5} A \right).$$

Consequently phase 1 is a global minimum in flat space if $A > -\frac{5}{8} \log \frac{3}{2}$, and phase 2 is a global minimum in flat space if $A < -\frac{5}{8} \log \frac{3}{2}$. If our universe is now in its true vacuum state, then we must restrict attention to $A > -\frac{5}{8} \log \frac{3}{2}$.

Let us turn our attention back to the curved-space potential, and replace μ^2 by its value (5.3) in terms of the physical mass M_x . We find that the potential is

$$V = \frac{3}{64\pi^2} \left[Q + \frac{1}{3} (1 - \log a^2 M_x^2) \right] R \sum_e m_e^2 - \frac{3}{320\pi^2} A \left(\sum_e m_e^2 \right)^2 + \frac{3}{64\pi^2} \left[\frac{12}{5} A + (1 - \log a^2 M_x^2) \right] \sum_e m_e^4 - \frac{3}{16\pi^2} a^{-4} \sum_e A(a^2 m_e^2) \quad (5.7)$$

where the parameter Q can be regarded as a renormalized value of ξ .

$$Q \equiv \frac{32\pi^2}{15g^2} \left[\xi - \frac{8}{5} g^{-2} \left(A_2 + \frac{7}{30} A_4 \right) \right]. \quad (5.8)$$

Because $A(z)$ is not a simple function, we studied this potential using numerical techniques. For a given direction in group space, i.e., in each phase, the problem of finding the extrema of the potential (5.7) is one-dimensional. In each phase A we introduce a dimensionless ratio

$$\tau_A(A, Q, a^2 M_x^2) \equiv \frac{M_A^2(A, Q, a^2 M_x^2)}{M_A^2} \quad (5.9)$$

of the gauge-field mass in curved space to its mass in flat space. It follows that the flat-space limit is $\lim_{a \rightarrow \infty} \tau_A(A, Q, a^2 M_x^2) \equiv 1$.

To write the equation satisfied by τ_A , it is useful to introduce some notation for the derivatives of the potential function. Because the potential (5.7) is invariant under any relabeling of the eigenvalues m_e^2 , we can use the notation $(\partial V / \partial M^2)$ to denote its partial derivative with respect to any eigenvalue, with the others held fixed. The eigenvalue with respect to which the derivative is taken is either $\tau_A M_A^2$ or $4\tau_A M_A^2$, since there are at most two different masses in the phases which we studied. The derivative can be easily evaluated from (5.7)

$$\frac{\partial V}{\partial M^2} = \frac{3}{64\pi^2} \left[Q + \frac{1}{3} (1 - \log a^2 M_x^2) \right] R - \frac{3A(C_3^4 + 4C_5^4) \tau_A M_A^2}{160\pi^2} + \frac{3}{32\pi^2} \left[\frac{12}{5} A + (1 - \log a^2 M_x^2) \right] M^2 - \frac{3}{16\pi^2} a^{-2} A'(a^2 M^2). \quad (5.10)$$

At an extremum, τ_A satisfies the equation

$$C_3 \frac{\partial V}{\partial M^2} \Big|_{M^2 = \tau_A M_A^2} + 4C_5 \frac{\partial V}{\partial M^2} \Big|_{M^2 = 4\tau_A M_A^2} = 0. \quad (5.11)$$

TABLE III

Coupling Constants and Parameters of the Theory before and after the One-Loop Corrections

	Before	After	
Curvature	R	$R = 12/a^2$	
Dimensionless coupling constants	$\begin{bmatrix} g \\ \xi \\ A_2 \\ A_4 \end{bmatrix}$	$\begin{bmatrix} g \\ Q(g, \xi, A_2, A_4) \\ A(g, A_2, A_4) \\ M_x \end{bmatrix}$	Dimensionless coupling constants Dimensional parameter

Note. One of the dimensionless coupling constants “disappears” and is replaced by a dimensional one.

So for a given set of parameters, this equation can be solved numerically to locate the extremum of each phase in group space.

We have seen how the flat-space limit can be used to express this theory’s curved space behavior in terms of well-defined physical quantities. The original set of parameters has been replaced by a new set. This is shown schematically in Table III. The table shows how each of the new parameters depends upon the old set.

The properties of the potential which we will go on to investigate depend only upon the three dimensionless parameters A , Q and R/M_x^2 . It will turn out that the theory’s stability properties are independent of the values of g or M_x . To determine the stability of each phase, we need to investigate the nature of the extrema at τ_A . One way to do this is to calculate the Higgs field’s masses in the different phases.

6. THE STABILITY OF THE EXTREMA AND THE HIGGS FIELD’S MASSES

In this section, we investigate the nature of the extrema in the five broken symmetry phases. We started with a potential which was a function of 24 variables, reduced it to four variables, and then to a single variable τ_A . This allowed us to find the extremum of the potential in each phase. However, this approach does not reveal if it is a minimum, maximum, saddle, or inflection point. To answer this question, we must return to the function of four variables.

Let us consider the shape of the potential near an extremum φ . If we perturb φ by a small traceless Hermitian matrix $\delta\varphi$ containing 24 independent infinitesimals

$$\varphi + \delta\varphi = \begin{bmatrix} \varphi_1 & & 0 \\ & \ddots & \\ 0 & & \varphi_5 \end{bmatrix} + \begin{bmatrix} \delta\varphi_{11} & \cdots & \delta\varphi_{15} \\ \vdots & & \vdots \\ \delta\varphi_{51} & \cdots & \delta\varphi_{55} \end{bmatrix} \tag{6.1}$$

then there exists a gauge transformation ρ which diagonalizes $\varphi + \delta\varphi$.

$$\rho(\varphi + \delta\varphi) = \text{diag}(\varphi_1 + \delta\bar{\varphi}_1, \dots, \varphi_5 + \delta\bar{\varphi}_5). \tag{6.2}$$

Here the diagonal perturbations $\delta\bar{\varphi}_i$ are linear combinations of the $\delta\varphi_{ij}$, which satisfy $\delta\bar{\varphi}_1 + \delta\bar{\varphi}_2 + \delta\bar{\varphi}_3 + \delta\bar{\varphi}_4 + \delta\bar{\varphi}_5 \equiv 0$. Near the extremum, the variation of the potential must be a quadratic form.

$$\delta V = \frac{1}{2}H_1(\delta\varphi_1)^2 + \frac{1}{2}H_2(\delta\varphi_2)^2 + \frac{1}{2}H_3(\delta\varphi_3)^2 + \frac{1}{2}H_4(\delta\varphi_4)^2 \quad (6.3)$$

where the $\delta\varphi_i$ are linear combinations of the $\delta\bar{\varphi}_i$. The four real numbers H_i are the Higgs fields (masses)².

In the vicinity of an extremum, the signs of the H_i tell you how many directions curve up, and how many curve down. Four positive H 's make the extremum a local minimum, and four negative H 's make it a local maximum. If some are positive, and some negative, it is a saddle, and if any H_i vanishes, then it is an inflection point.

The H_i are the four eigenvalues of the symmetric 4×4 matrix of second derivatives $[\partial^2 V / \partial\varphi_i \partial\varphi_j]$. If all four are positive, then they are the squares of the oscillation frequencies about the minimum, i.e., the Higgs field's masses². Since the number of massive Higgs fields is equal to the number of unbroken symmetries, some of the masses² are degenerate. An extremum is stable if, and only if, all of its Higgs masses² are positive.

To determine the Higgs masses in each group space direction, one needs to diagonalize a 4×4 matrix of second derivatives. This has been done in [13–15]. As well as the first derivative (5.10) of the potential, we need two second derivatives. The first is taken with respect to the same element of the mass matrix, and the second is taken with respect to two distinct elements of the mass matrix (which may be equal).

$$\frac{\partial^2 V}{(\partial M^2)^2} = \frac{33}{160\pi^2} A + \frac{3}{32\pi^2} (1 - \log a^2 M_x^2) - \frac{3}{16\pi^2} A''(a^2 M^2) \quad (6.4)$$

$$\frac{\partial^2 V}{\partial M^2 \partial N^2} = \frac{-3}{160\pi^2} A. \quad (6.5)$$

Now, following [13–15], we define the following derivatives in each phase A .

$$d_A \equiv 2 \left. \frac{\partial V}{\partial M^2} \right|_{M^2=0} \quad (6.6)$$

$$d'_A \equiv 2 \left. \frac{\partial V}{\partial M^2} \right|_{M^2=\tau_A M_A^2} \quad (6.7)$$

$$d''_A \equiv 2 \left. \frac{\partial V}{\partial M^2} \right|_{M^2=4\tau_A M_A^2} \quad (6.8)$$

$$e_A \equiv 4M^2 \left. \frac{\partial^2 V}{(\partial M^2)^2} \right|_{M^2=\tau_A M_A^2} \quad (6.9)$$

$$e'_A \equiv 4M^2 \frac{\partial^2 V}{(\partial M^2)^2} \Big|_{M^2 = 4\tau_A M_A^2} \quad (6.10)$$

$$f_A \equiv 4MN \frac{\partial^2 V}{\partial M^2 \partial N^2} \Big|_{M^2 = \tau_A M_A^2, N^2 = \tau_A M_A^2} \quad (6.11)$$

$$f'_A \equiv 4MN \frac{\partial^2 V}{\partial M^2 \partial N^2} \Big|_{M^2 = 4\tau_A M_A^2, N^2 = 4\tau_A M_A^2} \quad (6.12)$$

$$f''_A \equiv 4MN \frac{\partial^2 V}{\partial M^2 \partial N^2} \Big|_{M^2 = \tau_A M_A^2, N^2 = 4\tau_A M_A^2} \quad (6.13)$$

The Higgs masses H_1 to H_4 are linear combinations of these d 's, e 's and f 's. However, since each one is obtained by diagonalizing different matrices, they have to be treated phase by phase.

In phase 1, the $SU(3) \times SU(3) \times U(1)$ phase, the Higgs field's masses are

$$H_1 = H_2 = g^2(3d_1 + 2e_1 - 2f_1) \quad (6.14)$$

$$H_3 = 5g^2(e_1 + 11f_1) \quad (6.15)$$

$$H_4 = g^2(2d_1 + 3e_1 - 3f_1). \quad (6.16)$$

In phase 2, the $SU(4) \times U(1)$ phase, the Higgs field's masses are

$$H_1 = H_2 = H_3 = g^2(4d_2 + e_2 - f_2) \quad (6.17)$$

$$H_4 = 5g^2(e_2 + 7f_2). \quad (6.18)$$

In phase 3, the $SU(3) \times U(1) \times U(1)$ phase, the Higgs field's masses are

$$H_1 = H_2 = g^2(3d_3 + 2e_3 - 2f_3 + 2d'_3) \quad (6.19)$$

$$H_3 = 5g^2(e_3 - f_3 + d'_3) \quad (6.20)$$

$$H_4 = g^2(24f''_3 + 2f'_3 + 33f_3 + 2e'_3 + 3e_3) \quad (6.21)$$

In phase 4, the $SU(2) \times SU(2) \times U(1) \times U(1)$ phase, the Higgs field's masses are

$$H_1 = H_2 = g^2(2d'_4 - 2d''_4 + e_4 - f_4 + 2e'_4 - 2f'_4) \quad (6.22)$$

$$H_3 = 5g^2(e_4 - f_4 + d'_4) \quad (6.23)$$

$$H_4 = g^2(32f''_4 + 28f'_4 + 7f_4 + 4e'_4 + e_4). \quad (6.24)$$

In phase 5, the unbroken symmetry $SU(5)$ phase, the Higgs field's masses are all equal.

$$H_1 = H_2 = H_3 = H_4 = 5g^2 d_5. \quad (6.25)$$

Fortunately phase 5 is very interesting to us, because this expression can be easily calculated from (6.6) and (5.10) and the value of $A(0)$.

$$H_1 = H_2 = H_3 = H_4 = \frac{45}{8\pi^2} g^2 a^{-2} \left[Q + \frac{5}{9} - \frac{2}{3} \gamma - \frac{1}{3} \log a^2 M_x^2 \right]. \quad (6.26)$$

In the flat-space limit, as $a \rightarrow \infty$, these masses² approach zero, from the negative direction. Hence phase 5 is stable if and only if

$$a^2 M_x^2 < \exp(3Q + \frac{5}{3} - 2\gamma). \quad (6.27)$$

If $a^2 M_x^2 = \exp(3Q + \frac{5}{3} - 2\gamma)$ there is a second-order phase transition from the $SU(5)$ symmetric phase.

The critical curvature $R = 12M_x^2 \exp(-3Q - \frac{5}{3} + 2\gamma)$ is important, because the $SU(5)$ phase becomes unstable when the curvature reaches this value. For larger curvatures, the Higgs field is trapped at $\phi = 0$ by a barrier, and can only exit the $SU(5)$ phase via a first-order transition.

It is instructive to examine the Higgs masses in flat space. The d 's, e 's and f 's in (6.6) \rightarrow (6.13) can be easily evaluated in this limit by letting $a \rightarrow \infty$ and $\tau_A \rightarrow 1$, and using the asymptotic behavior (4.5), (4.6) of A' and A'' . One obtains the following *flat-space* masses, which are linear functions of A .

Phase 1:

$$H_1 = H_2 = \frac{9}{20\pi^2} g^2 M_x^2 \left(A + \frac{5}{3} \right) \quad (6.28)$$

$$H_3 = \frac{15}{8\pi^2} g^2 M_x^2 \quad (6.29)$$

$$H_4 = \frac{9}{5\pi^2} g^2 M_x^2 \left(A + \frac{5}{8} \right). \quad (6.30)$$

Phase 2:

$$H_1 = H_2 = H_3 = \frac{3}{5\pi^2} g^2 M_2^2 \left(-A + \frac{5}{8} \right) \quad (6.31)$$

$$H_4 = \frac{15}{8\pi^2} g^2 M_2^2. \quad (6.32)$$

Phase 3:

$$H_1 = H_2 = \frac{75}{44\pi^2} g^2 M_3^2 \left(-A + \frac{11}{25} - \frac{24}{25} \log 2 \right) \quad (6.33)$$

$$H_3 = \frac{15}{11\pi^2} g^2 M_3^2 \left(A + \frac{11}{8} - 3 \log 2 \right) \quad (6.34)$$

$$H_4 = \frac{33}{8\pi^2} g^2 M_3^2. \quad (6.35)$$

Phase 4:

$$H_1 = H_2 = \frac{90}{17\pi^2} g^2 M_4^2 \left(A - \frac{51}{80} - \frac{7}{30} \log 2 \right) \tag{6.35}$$

$$H_3 = \frac{15}{17\pi^2} g^2 M_4^2 \left(-A + \frac{17}{8} - 6 \log 2 \right) \tag{6.37}$$

$$H_4 = \frac{51}{8\pi^2} g^2 M_4^2. \tag{6.38}$$

Phase 5:

$$H_1 = H_2 = H_3 = H_4 = 0. \tag{6.39}$$

From these formulae, one can see that phases 3, 4, and 5 are unstable in flat space, for any value of A . Only phases 1 and 2 can be stable in flat space. If $A > -\frac{5}{8}$ then phase 1 is stable, and if $A < \frac{5}{8}$ then phase 2 is stable. From (5.6) we know that the line $A = -\frac{5}{8} \log \frac{3}{2}$ separates the regions in which phases 1 and 2 are the global minima. The resulting flat-space phase diagram is shown in Fig. 1. The regions of the phase diagram are labeled using the stability code notation introduced in Section 3.

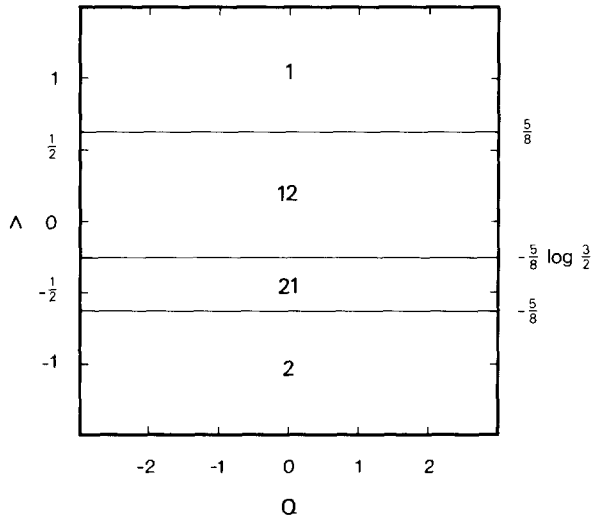


FIG. 1. The phase diagram for a flat-space ($a \rightarrow \infty$) Coleman-Weinberg potential. Since Q is a renormalized ξ parameter, the diagram is independent of it. Shown are the four possible stability code regions. For example, when $-\frac{5}{8} \log \frac{3}{2} < A < \frac{5}{8}$, phase 1 is stable and phase 2 is metastable.

7. CURVED-SPACE PHASE DIAGRAMS

For given values of Λ , Q , and aM_x , one can solve (5.11) for τ_A to find the location of the extrema in each phase. The Higgs masses in each phase can then be determined. Finally, the values of the potential at the stable extrema can be compared, to obtain a stability code for the given values of Λ , Q and aM_x . In this way we obtained some phase diagrams for this theory.

We solved the equations on a computer with a Fortran program composed of several subroutines. The lowest-level subroutines evaluate $A(z)$, $A'(z)$ and $A''(z)$. Another subroutine finds the location of the extremum in each phase, by solving (5.11) for τ_A . Another one evaluates the Higgs masses H_1 to H_4 at each extremum, and another calculates the potential. The last subroutine examines the Higgs masses, and compares the values of the potential in each phase, to obtain a stability code.

For any value of Λ and Q , there is a critical curvature with the property that if the curvature is greater than that value, then the only stable or metastable phase is the $SU(5)$ symmetric phase. This curvature is shown in Fig. 2. The contours show the critical values of $\log a^2 M_x^2$. The phase diagram also shows what happens at this

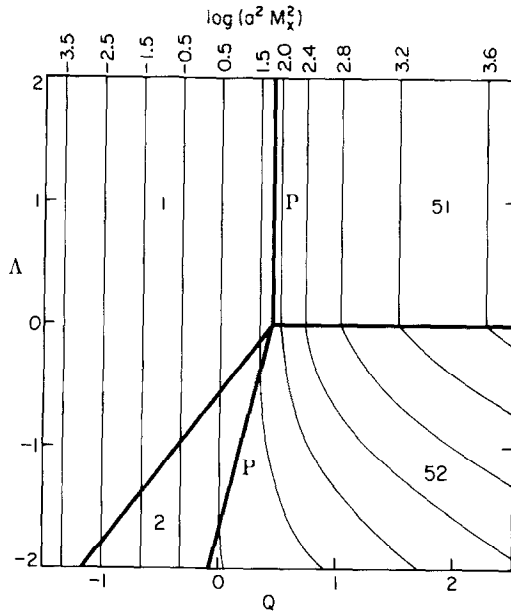


FIG. 2. The contours on this phase diagram show a critical value of $\log(a^2 M_x^2)$. When the curvature is greater than this value, phase 5 is the only stable or metastable phase. At the critical curvature, there are four regions of parameter space with different stability codes. To the left of the boundary P , the phase transition is second order, and phase 5 becomes unstable. To the right of P , phase 5 remains stable, and phase 1 or 2 becomes metastable.

critical curvature, when phase 5 ceases to be the only stable or metastable phase. The parameter space (A, Q) breaks up into four disjoint stability code regions. For example, in the top left corner of the diagram, at the critical curvature shown by the contour, the $SU(3) \times SU(2) \times U(1)$ phase becomes stable, and the $SU(5)$ phase becomes unstable. However, in the top right corner, the $SU(3) \times SU(2) \times U(1)$ phase becomes metastable, and the $SU(5)$ phase remains the true stable minimum.

It is not very interesting if another phase becomes metastable while the $SU(5)$ phase remains stable. This is what happens to the right of the boundary P , shown in Fig. 2. To the left of P , there is a second-order phase transition when the $SU(5)$ minimum disappears, at $\log(a^2 M_x^2) = 3Q + \frac{5}{3} - 2\gamma$. However, to the right of P , the universe would remain in phase 5, since the new local minimum is "higher up the hill."

Hence the next critical curvature: the one at which some minimum *other* than the $SU(5)$ symmetric one becomes lower than the $SU(5)$ minimum. The contours of this critical curvature are shown in Fig. 3. This critical curvature is slightly less than the one shown in Fig. 2, so the contours actually show the difference of their logarithms. It is clear that the metastable minima, which first develop at the curvature shown in Fig. 2, simply move down, and become lower than the $SU(5)$ minimum in Fig. 3.

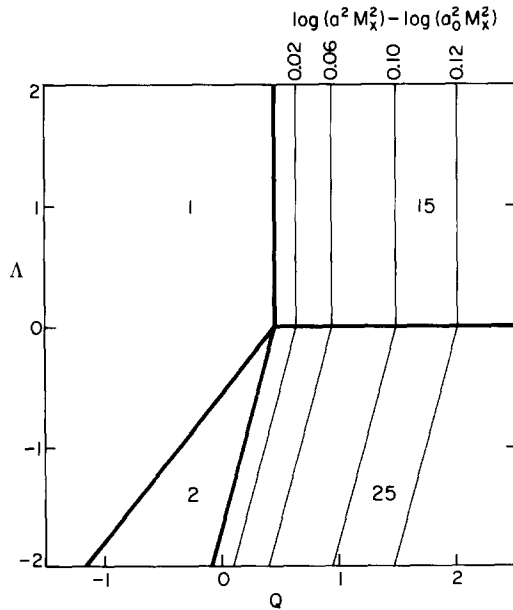


FIG. 3. Phase diagram showing the critical curvature at which phase 5 ceases to be the minimum of lowest energy. The critical curvature is shown relative to the critical curvature in Fig. 2, denoted as $\log(a_0^2 M_x^2)$. By comparing the two diagrams, one can see that the metastable minima simply "trade places" with phase 5, and become lower than it.

There is still a barrier between these new stable phases, and the now metastable $SU(5)$ phase. What happens when the $SU(5)$ minimum finally becomes unstable, and the barrier disappears? This interesting question is amenable to exact treatment.

When the $SU(5)$ phase becomes unstable, at a critical curvature $\log a^2 M_x^2 = 3Q + \frac{5}{3} - 2\gamma$, the Higgs field begins to slowly move away from $\phi = 0$. We want to know in what group-space direction it goes. Since ϕ is small, all of the eigenvalues m_e^2 of the mass matrix are small. Consequently, we can expand the potential (5.7) in a Taylor series. For small values of $a^2 m_e^2$, $A(a^2 m_e^2) = A(0) + A'(0) a^2 m_e^2 + \frac{1}{2} A''(0) a^4 m_e^4 + \dots$ where the values of $A'(0)$ and $A''(0)$ are given in Section 4. The potential becomes

$$V = \frac{3}{64\pi^2} \left[Q + \frac{1}{3} (1 - \log a^2 M_x^2 - A'(0)) \right] R \sum_e m_e^2 - \frac{3}{320\pi^2} A \left(\sum_e m_e^2 \right)^2 + \frac{3}{64\pi^2} \left[\frac{12}{5} A + 1 - \log a^2 M_x^2 - 2A''(0) \right] \sum_e m_e^4 + O(a^2 m_e^6). \tag{7.1}$$

At the critical curvature, the coefficient of the $\sum_e m_e^2$ term vanishes. We can use (3.3) and (3.4) to express the potential in terms of the Higgs field φ ,

$$V = \frac{3}{320\pi^2} g^4 [X(\text{trace } \varphi^2)^2 + Y \text{trace } \varphi^4] \tag{7.2}$$

where the coefficients X and Y are

$$X = -7 \left(A + \frac{45}{14} Q - \frac{10}{7} \right) \tag{7.3}$$

$$Y = 30 \left(A - \frac{5}{4} Q + \frac{5}{9} \right). \tag{7.4}$$

The polynomial potential (7.2) is a good approximation near $\varphi = 0$, when the $SU(5)$ phase becomes unstable.

These polynomial potentials have been analysed by Li [16]. He found that if $X > 0$ and $Y > 0$, then the minimum moves in the direction of phase 1, and if $X > 0$ and $Y < 0$, then it moves in the direction of phase 2. These conditions define two lines in the parameter space (A, Q) .

$$X > 0 \Leftrightarrow A < \frac{-45}{14} Q + \frac{10}{7} \tag{7.5}$$

$$Y > 0 \Leftrightarrow A > \frac{5}{4} Q - \frac{5}{9}. \tag{7.6}$$

The lines intersect at $A = 0, Q = \frac{4}{9}$. They are shown as dotted lines on the phase diagram in Fig. 4.

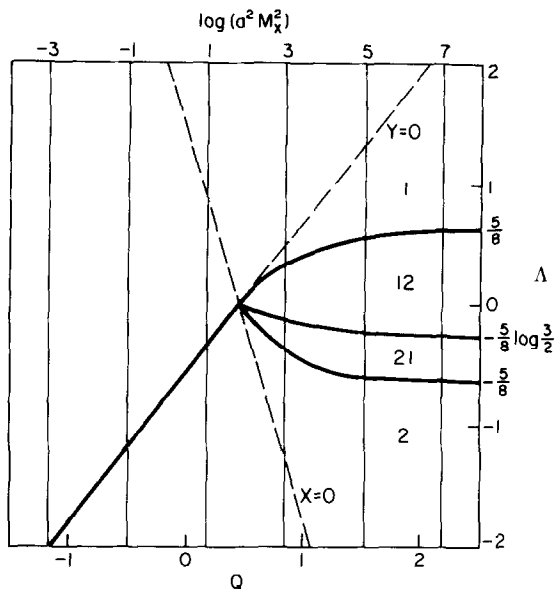


FIG. 4. Phase diagram showing the critical curvature at which the $SU(5)$ phase becomes unstable, and the corresponding stability codes. As the critical curvature $\rightarrow \infty$, the stability code pattern approaches the flat-space one, previously shown in Fig. 1. The two lines $X=0$ and $Y=0$ are the linear relations of equations (7.3) and (7.4). It is only when $\log(a^2 M_x^2) \lesssim 2$ that the phase diagram is significantly different than the flat-space one.

This last diagram shows the stability codes when the $SU(5)$ minimum disappears. It clearly shows that, as the critical curvature becomes very small, the shape of the potential approaches the flat-space pattern shown in Fig. 1. The “trifurcation point” is located by the intersection of the two boundary lines. This diagram was obtained entirely from numerical computation, and its agreement with the analysis above is convincing proof of the computer program’s reliability.

8. CONCLUSION

In this paper we have demonstrated how the one-loop effective potential can be calculated for non-Abelian gauge theories in DeSitter space. Furthermore, this analysis shows that curved-space effects do not modify the conclusions of Breit et al. [3] that in the $SU(5)$ theory, the universe is likely to end up in the $SU(4) \times U(1)$ phase. The reason is that the curvature effects only make significant modifications to the flat-space potential when $R \gtrsim M_x^2$. In the inflationary universe model, the curvature is smaller than this, by a factor $(M_x/M_p)^2 \sim 10^{-8}$. Consequently, the flat-space studies remain valid.

For this reason, we see no need to undertake a dynamical study, using the curved-space potential. Such a study would require the introduction of two

additional parameters, one to measure the amount of radiative damping as the Higgs field evolved, and another to fix the ratio of M_x/M_p .

A surprising result of our computer study is that phases 3 and 4, the $SU(2) \times SU(2) \times U(1) \times U(1)$ and $SU(3) \times U(1) \times U(1)$ phases, are unstable for any values of the parameters A and Q . We searched through a large region of parameter space, but were unable to find any values of A , Q and R/M_x^2 for which phase 3 or 4 was either stable or metastable.

ACKNOWLEDGMENTS

I would like to thank I. T. Drummond and I. G. Moss for a number of useful discussions. This work was supported in part by the Marshall Aid Commemoration Commission (U.K.) and NSF Grant PHY-81-07384.

REFERENCES

1. B. ALLEN, *Nucl. Phys. B* **226** (1983), 228.
2. See the articles about inflation in "The Very Early Universe" (Hawking, Gibbons, and Siklos, Eds.), Cambridge Univ. Press, Cambridge, 1983.
3. J. D. BREIT, S. GUPTA, AND A. ZAKS, *Phys. Rev. Lett.* **51** (1983), 1007.
4. P. H. FRAMPTON AND S. L. GLASHOW, *Phys. Lett. B* **131** (1983), 340.
5. S. COLEMAN AND E. WEINBERG, *Phys. Rev. D* **7** (1973), 1888.
6. H. GEORGI AND S. L. GLASHOW, *Phys. Rev. Lett.* **32** (1974), 438.
7. H. GEORGI, H. R. QUINN, AND S. WEINBERG, *Phys. Rev. Lett.* **33** (1974), 451.
8. L. D. FADDEEV AND V. N. POPOV, Perturbation theory for gauge invariant fields, in "Gauge Theory of Weak and Electromagnetic Interactions" (C. H. Lai, Ed.), World Scientific, Singapore, 1981.
9. J. S. KIM, *Nucl. Phys. B* **196** (1982), 285.
10. J. FRAUTSCHI AND J. S. KIM, *Nucl. Phys. B* **196** (1982), 301.
11. J. S. KIM, *Nucl. Phys. B* **197** (1982), 174.
12. J. S. KIM, *Nucl. Phys. B* **207** (1982), 374.
13. A. I. BOCHAREV, V. A. KUZMIN, M. E. SHAPOSHNIKOV, AND I. I. TKACHEV, *Z. Phys. C* **15** (1982), 33.
14. V. A. KUZMIN, M. E. SHAPOSHNIKOV, AND I. I. TKACHEV, *Z. Phys. C* **12** (1982), 83.
15. V. A. KUZMIN, M. E. SHAPOSHNIKOV, AND I. I. TKACHEV, *Phys. Lett. B* **102** (1981), 397.
16. L. F. LI, *Phys. Rev. D* **9** (1974), 1723.