PHYSICAL REVIEW D VOLUME 32, NUMBER 6 15 SEPTEMBER 1985

Gaussian effective potential. II. $\lambda\phi^4$ field theory

P. M. Stevenson
Physics Department, University of Wisconsin—Madison, Madison, Wisconsin 53706
*and T. W. Bonner Laboratories, Physics Department, Rice University, Houston, Texas 77251**
(Received 18 December 1984)

The Gaussian-effective-potential approach is used to explore the physics of $\lambda\phi^4$ field theory in 1, 2, 3, and 4 spacetime dimensions. A simple and systematic approach to the renormalization, without explicit regularization, is employed. In four dimensions we find a viable, nontrivial theory arising from a bare coupling constant of a particular negative, infinitesimal form. The theory is "precarious": it is stable in the absence of an ultraviolet cutoff, but unstable when regularized. Perturbation theory is related to this form of the theory, though not straightforwardly. We also discuss particle masses, and the absence of two-particle bound states in ϕ^4 theory.

I. INTRODUCTION

The Gaussian effective potential[1–10] (GEP) is a simple, nonperturbative approach to quantum field theory, based on intuitive ideas familiar in quantum mechanics. In the first paper of this series[1] we motivated the GEP, stressing its advantages over the conventional one-loop effective potential, which we illustrated in a variety of quantum-mechanical examples. In this paper we move on to consider scalar field theories with a $\lambda\phi^4$ interaction in $1+1$, $2+1$, and $3+1$ dimensions.

The basic ideas behind the GEP have a long history[2–9] and some of our material overlaps with earlier authors. The aim is to present a simple and self-contained overview of the basic physics of $\lambda\phi^4$ theory, and how it evolves from $0+1$ dimensions (the quantum-mechanical anharmonic oscillator) to $3+1$ dimensions. We shall try to be very clear and systematic about the renormalization process, often a source of confusion in the earlier papers. (Indeed, whatever its other merits, the GEP is a wonderful pedagogical tool for explaining renormalization, as it is free of the complications and ambiguities that arise when renormalization is applied in perturbation theory.)

Our most interesting results (outlined briefly in Ref. 10) are for the $(3+1)$-dimensional theory. Current dogma, based on the works of Refs. 11–13, holds that $(\phi^4)_{3+1}$ is a "trivial" theory. However, our results indicate a nontrivial version of the theory exists if the bare coupling constant is *negative* (and infinitesimal), a possibility not considered in Refs. 11–13. The stability of this theory is a subtle matter, and we are led to define the concept of a "precarious" field theory. This class of theories cannot be realized as the continuum limit of a lattice theory, at least not in the ordinary way. Nevertheless, we shall argue that such theories are not to be summarily dismissed as of no physical interest.

The plan of the paper is as follows. Section II explains the calculation of the GEP, in its bare form. Our approach to the renormalization process is discussed in detail in Sec. III. Section IV presents the results, working up from $0+1$ to $3+1$ dimensions. The nontrivial version of $(\phi^4)_{3+1}$ is analyzed and discussed in Sec. V. A brief discussion of one- and two-particle states, following Refs. 2 and 4, is given in Sec. VI, where we explain our attitude to the "factor of 2" puzzle which plagued these earlier works. Section VII summarizes our main conclusions, particularly regarding the nontriviality of $(\phi^4)_{3+1}$ theory. (An Appendix notes some curious results which arise if dimensional regularization is used.)

Familiarity with Ref. 1 (hereafter referred to as I) is not essential for understanding what we are doing here. However, it may help to explain why we are doing it.

II. CALCULATION OF THE GEP FOR $\lambda\phi^4$ THEORY

The Lagrangian we consider is

$$\mathscr{L}=\tfrac{1}{2}\partial_\mu\phi\partial^\mu\phi-\tfrac{1}{2}m_B{}^2\phi^2-\lambda_B\phi^4\,, \tag{2.1}$$

which corresponds to a Hamiltonian density

$$\mathscr{H}=\tfrac{1}{2}\dot{\phi}^2+\tfrac{1}{2}(\nabla\phi)^2+\tfrac{1}{2}m_B{}^2\phi^2+\lambda_B\phi^4\,. \tag{2.2}$$

The GEP is defined as

$$\begin{aligned}\bar{V}_G(\phi_0)&=\min_{\Omega} V_G(\phi_0,\Omega)\\&=\min_{\Omega}\;{}_{\phi_0,\Omega}\langle 0|\mathscr{H}|0\rangle_{\Omega,\phi_0}\,,\end{aligned} \tag{2.3}$$

where $|0\rangle_{\Omega,\phi_0}$ is a normalized Gaussian wave functional, centered on $\phi=\phi_0$:

$${}_{\phi_0,\Omega}\langle 0|0\rangle_{\Omega,\phi_0}=1,\quad {}_{\phi_0,\Omega}\langle 0|\phi|0\rangle_{\Omega,\phi_0}=\phi_0\,. \tag{2.4}$$

The mass parameter Ω must be positive definite in order for the wave functional to be normalizable, just as in quantum mechanics (cf. I). The calculation can be performed in a Schrödinger wave-functional formalism, as shown in Ref. 4, but here we use a simpler, but entirely equivalent, method.

We write the field ϕ as $\phi_0+\hat{\phi}$, where ϕ_0 is a constant classical field, and $\hat{\phi}$ is a quantum free field of mass Ω: The state $|0\rangle_{\Omega,\phi_0}$ is then the vacuum state of this free field. The required formalism is familiar from any field-theory textbook.[14] We write

$$\phi=\phi_0+\int(dk)_\Omega[a_\Omega(\mathbf{k})e^{-ik\cdot x}+a_\Omega^\dagger(\mathbf{k})e^{ik\cdot x}]\,, \qquad (2.5)$$

and hence,

$$\partial_\mu\phi=\int(dk)_\Omega(-ik_\mu)[a_\Omega(\mathbf{k})e^{-ik\cdot x}-a_\Omega^\dagger(\mathbf{k})e^{ik\cdot x}]\,, \qquad (2.6)$$

where the energy component of the four-vector k_μ is

$$k^0=\omega_{\underline{k}}(\Omega)\equiv(\mathbf{k}^2+\Omega^2)^{1/2}\,. \qquad (2.7)$$

The integration measure in ν spatial dimensions is

$$(dk)_\Omega\equiv\frac{d^\nu k}{(2\pi)^\nu 2\omega_{\underline{k}}(\Omega)}\,. \qquad (2.8)$$

The creation and annihilation operators obey the usual commutation relation

$$[a_\Omega(\mathbf{k}),a_\Omega^\dagger(\mathbf{k}')]=\delta_{kk'}\equiv 2\omega_{\underline{k}}(\Omega)(2\pi)^\nu\delta^{(\nu)}(\mathbf{k}-\mathbf{k}')\,, \qquad (2.9)$$

and the state $|0\rangle_\Omega$ has the defining property

$$a_\Omega(\mathbf{k})\,|0\rangle_\Omega=0\,. \qquad (2.10)$$

The evaluation of $V_G(\phi_0,\Omega)$ for the Hamiltonian (2.2) is a straightforward exercise. Term by term we have

$$\begin{aligned}
{}_{\phi_0,\Omega}\langle 0|\tfrac{1}{2}[\dot\phi^2+(\nabla\phi)^2]|0\rangle_{\Omega,\phi_0} &= \int(dk)_\Omega[\omega_{\underline{k}}^2(\Omega)-\tfrac{1}{2}\Omega^2]\,,\\
{}_{\phi_0,\Omega}\langle 0|\tfrac{1}{2}m_B^2\phi^2|0\rangle_{\Omega,\phi_0} &= \tfrac{1}{2}m_B^2\left[\phi_0^2+\int(dk)_\Omega\right],\\
{}_{\phi_0,\Omega}\langle 0|\lambda_B\phi^4|0\rangle_{\Omega,\phi_0} &= \lambda_B\left[\phi_0^4+6\phi_0^2\int(dk)_\Omega+3\int(dk)_\Omega\int(dk')_\Omega\right].
\end{aligned} \qquad (2.11)$$

Introducing the notation

$$I_N(\Omega)\equiv\int(dk)_\Omega[\omega_{\underline{k}}^2(\Omega)]^N \qquad (2.12)$$

we can write the result as

$$V_G(\phi_0,\Omega)=I_1+\tfrac{1}{2}(m_B^2-\Omega^2)I_0+\tfrac{1}{2}m_B^2\phi_0^2+\lambda_B\phi_0^4+6\lambda_B I_0\phi_0^2+3\lambda_B I_0^2\,. \qquad (2.13)$$

The GEP itself, $\overline{V}_G(\phi_0)$, is obtained by minimizing this expression with respect to the variational parameter Ω, in the range $0<\Omega<\infty$ (Ref. 15). We denote the optimum value of Ω by $\overline{\Omega}$. Normally, $\overline{\Omega}$ will be given by the equation $dV_G/d\Omega\,|_{\Omega=\overline{\Omega}}=0$, which, using the formal result that

$$dI_N/d\Omega=(2N-1)\Omega I_{N-1}\,, \qquad (2.14)$$

leads to the "$\overline{\Omega}$ equation"

$$\overline{\Omega}^2=m_B^2+12\lambda_B[I_0(\overline{\Omega})+\phi_0^2]\,. \qquad (2.15)$$

However, there are two subtleties to beware of: first, Eq. (2.15) may have more than one solution, and one must take care to select the right one. In particular, the solution must be a minimum, not a maximum of V_G. Second, the *global* minimum of $V_G(\phi_0,\Omega)$ may not be a solution of (2.15) at all, but may occur at one or another end point of the range $0<\Omega<\infty$. Except in the special cases where the latter caveat applies, one may use the $\overline{\Omega}$ equation to simplify the expression for $\overline{V}_G$, whenever it is convenient. For example, one can use (2.15) to write $\overline{V}_G(\phi_0)$ in the form

$$\overline{V}_G(\phi_0)=I_1-3\lambda_B I_0^2+\tfrac{1}{2}m_B^2\phi_0^2+\lambda_B\phi_0^4\,. \qquad (2.16)$$

Note that, implicitly, the argument of the I_N integrals here is $\overline{\Omega}$, which is ϕ_0 dependent by virtue of (2.15).

III. RENORMALIZATION

A. Outline

The above expressions for the GEP are full of divergent integrals. Our goal in this section is to explain how to reexpress $V_G(\phi_0)$ as a manifestly finite function of ϕ_0. This is true except for a divergent, but ϕ_0-independent constant D,

$$D\equiv\overline{V}_G(\phi_0=0)=I_1(\overline{\Omega}_0)-3\lambda_B[I_0(\overline{\Omega}_0)]^2\,, \qquad (3.1)$$

where $\overline{\Omega}_0$ is the solution to the $\overline{\Omega}$ equation at $\phi_0=0$. The constant D represents the vacuum energy density of the $\phi_0=0$ vacuum. The presence of this divergent constant has no physical consequences, since only energy differences, not absolute energies, are measurable. Later on, we shall follow the usual practice of redefining the zero of the energy scale so that $\overline{V}_G(\phi_0=0)=0$.

Except for D, the GEP contains divergences only because it is parametrized in terms of the bare parameters m_B,λ_B which are themselves not finite. The finiteness of $\overline{V}_G(\phi_0)$ becomes manifest once one reexpresses it in terms of finite parameters. This *reparametrization* of the theory, misnamed "renormalization," does not change the physical content of the theory, but simply reexpresses it in more transparent form.

A convenient choice for the two new parameters is to define[4]

$$\begin{aligned}
m_R^2&\equiv d^2\overline{V}_G/d\phi_0^2\,|_{\phi_0=0}\,,\\
\lambda_R&\equiv\frac{1}{4!}d^4\overline{V}_G/d\phi_0^4\,|_{\phi_0=0}\,.
\end{aligned} \qquad (3.2)$$

The claim of renormalizability is that by reparametrizing $\overline{V}_G(\phi_0)$ in terms of m_R,λ_R (i.e., by changing variables, eliminating m_B,λ_B in favor of m_R,λ_R), one achieves a manifestly finite result.

We stress that the physical content of the results is the same, no matter how one chooses to parametrize them. The use of other renormalized parameters m_R',λ_R' would lead to different looking but equivalent results. In other words, the GEP is exactly renormalization-group[16] (RG) invariant. This is not true of the one-loop effective potential,[17] because whether a term is of "one-loop" or "two-loop" order, etc., depends on precisely how the renormalized mass and coupling constant are defined. That is, RG invariance is spoiled by using the one-loop approximation. As with perturbation theory, this leads to a "renormalization-scheme-dependence problem,"[18] which is the question of how "best" to define the renormalized parameters; i.e., how to "RG-improve" the results.[17] For the GEP, exact RG invariance means that all ways of defining the renormalized parameters are equivalent, and it is merely a matter of finding the most convenient.

The parameter m_R, as defined above, is very convenient because it turns out to be the mass of a one-particle excitation in the $\phi_0=0$ vacuum.[2,4] See Sec. VI. In 1 + 1 and 2 + 1 dimensions λ_B is finite, so that there is no need to eliminate it in favor of λ_R; indeed this would only serve to complicate the expressions. The use of λ_R will, however, be essential in the (3 + 1)-dimensional case.

B. Handling the divergent integrals

To actually perform the change of variables from m_B (and λ_B) to m_R (and λ_R) is a somewhat delicate business, since it involves handling the divergent integrals I_N. It would be usual at this stage to introduce a regularization device. One might introduce an ultraviolet cutoff $|\mathbf{k}|<M_{\rm UV}$ to restrict the range of integration in the I_N integrals, or one might set up the theory on a spatial lattice, which effectively introduces an ultraviolet cutoff of order of the inverse lattice spacing. Alternatively, one could analytically continue in ν, the number of spatial dimensions.[19] Any of these modifications to the theory will serve to make all the algebra finite. One can then happily proceed to change variables from m_B (and λ_B) to m_R (and λ_R) and, *at the end*, remove the regulator (e.g., take $M_{\rm UV}\to\infty$) to recover the original theory. The bare parameters will have to have a certain regulator dependence—and may well tend to infinity, or to zero, as the regulator is removed—in order for the renormalized parameters to remain finite.

TABLE I. A summary of the leading divergent behavior of the I_N integrals, in terms of an ultraviolet cutoff $M_{\rm UV}$.

	I_{-1}	I_0	I_1
1 + 1		$\ln M_{\rm UV}$	$M_{\rm UV}{}^2$
2 + 1		$M_{\rm UV}$	$M_{\rm UV}{}^3$
3 + 1	$\ln M_{\rm UV}$	$M_{\rm UV}{}^2$	$M_{\rm UV}{}^4$

We follow this program in spirit, although we do not explicitly introduce a regularization device. The reader may regard, say, an ultraviolet cutoff as being implicitly present. (For the reader's convenience, the leading divergent behavior of the I_N integrals is summarized in Table I.) We may dispense with an explicit regularization because our calculation will involve only differences $I_N(\Omega)-I_N(m)$, for which we can show, formally, that

$$\begin{aligned}
I_N(\Omega)-I_N(m)&=\frac{1}{2(2\pi)^\nu}\int d^\nu k[(\mathbf{k}^2+\Omega^2)^{N-1/2}-(\mathbf{k}^2+m^2)^{N-1/2}]\\
&=\int\frac{d^\nu k}{(2\pi)^\nu}\frac{1}{2(\mathbf{k}^2+m^2)^{1/2}}(\mathbf{k}^2+m^2)^N\left[\left[1-\frac{(m^2-\Omega^2)}{(\mathbf{k}^2+m^2)}\right]^{N-1/2}-1\right]\\
&=\int (dk)_m\sum_{r=1}^{\infty}\begin{bmatrix}-N-\frac{1}{2}+r\\ r\end{bmatrix}\frac{(m^2-\Omega^2)^r}{(\mathbf{k}^2+m^2)^{r-N}}\\
&=\sum_{r=1}^{\infty}\begin{bmatrix}-N-\frac{1}{2}+r\\ r\end{bmatrix}(m^2-\Omega^2)^r I_{N-r}(m)\ .
\end{aligned}\tag{3.3}$$

This formula expresses the difference of two I_N integrals as a sum of less-divergent integrals. Indeed, from $r=N+k$ onward, where

$$k\equiv\begin{cases}(\nu+1)/2, & \nu=\text{odd}\\ \nu/2, & \nu=\text{even}\ ,\end{cases}\tag{3.4}$$

the terms involve convergent integrals, for which we may use

$$I_N(\Omega)=\frac{1}{2(4\pi)^{\nu/2}}\frac{\Gamma[-(2N+\nu-1)/2]}{\Gamma[-(2N-1)/2]}\Omega^{2N+\nu-1}\quad\text{for } 2N+\nu-1<0\ .\tag{3.5}$$

It is then possible to resum the bulk of the series, so that we obtain the following formula:

$$\begin{aligned}
I_N(\Omega)-I_N(m)&=\sum_{r=1}^{N+k-1}\begin{bmatrix}-N-\frac{1}{2}+r\\ r\end{bmatrix}(m^2-\Omega^2)^r I_{N-r}(m)\\
&\quad+\frac{\Gamma(N+\frac{1}{2})}{\Gamma(\frac{1}{2})}m^{2N+\nu-1}\frac{(-1)^k}{(4\pi)^k}L_{N+k}^{(\nu)}(x)\ ,
\end{aligned}\tag{3.6}$$

where the functions $L_i^{(\nu)}$ are defined recursively as

$$L_{i+1}^{(\nu)}(x)=\int_1^x dx' L_i^{(\nu)}(x')\ ,\tag{3.7}$$

with

$$L_0^{(\nu)}\equiv\begin{cases}1/x, & \nu=\text{odd}\\ 1/(2\sqrt{x}\,), & \nu=\text{even}\end{cases}\tag{3.8}$$

and

TABLE II. Useful formulas for the differences of I_N integrals in $\nu+1$ dimensions [from Eq. (3.6)]: $x \equiv \Omega^2/m^2$.

$\nu=1$ or 2
$I_1(\Omega)-I_1(m)=\frac{1}{2}(\Omega^2-m^2)I_0(m)-m^{\nu+1}L_2(x)/(8\pi)$
$I_0(\Omega)-I_0(m)=-m^{\nu-1}L_1(x)/(4\pi)$
$I_{-1}(\Omega)=\begin{cases}1/(2\pi\Omega^2), & \nu=1\\ 1/(4\pi\Omega), & \nu=2\end{cases}$
$\nu=3$ (or 4)
$I_1(\Omega)-I_1(m)=\frac{1}{2}(\Omega^2-m^2)I_0(m)-\frac{1}{8}(\Omega^2-m^2)^2I_{-1}(m)+m^{\nu+1}L_3(x)/(32\pi^2)$
$I_0(\Omega)-I_0(m)=-\frac{1}{2}(\Omega^2-m^2)I_{-1}(m)+m^{\nu-1}L_2(x)/(16\pi^2)$
$I_{-1}(\Omega)-I_{-1}(m)=-m^{\nu-3}L_1(x)/(8\pi^2)$

$$x \equiv \Omega^2/m^2 \,. \tag{3.9}$$

The special cases of the master formula (3.6) which we shall need are set out in Table II, and the first few L_i functions are listed in Table III. Note that the formal result for $dI_N/d\Omega$, used in Eq. (2.14) earlier, can be regarded as a special case of (3.6).

We also note here that I_N can be reexpressed as manifestly covariant integrals over the $(\nu+1)$-dimensional energy-momentum space; e.g.,

$$I_1(\Omega)=-\frac{i}{2}\frac{1}{(2\pi)^{\nu+1}}\int d^{\nu+1}k\,\ln(k^2-\Omega^2) + \text{constant}\,, \tag{3.10}$$

$$I_0(\Omega)=\frac{i}{(2\pi)^{\nu+1}}\int d^{\nu+1}k\frac{1}{k^2-\Omega^2}\,, \tag{3.11}$$

The others can be obtained by use of (2.14). Wick rotation, with the usual $i\epsilon$ prescription, shows the equivalence to the original expressions. (In the case of I_1, the integrand does not fall off as $|k^0|\to\infty$, so that there is a divergent constant contribution from the contour at infinity.)

Another remark concerns the infrared properties of the integrals, by which we mean the convergence or otherwise at the $|\mathbf{k}|\to 0$ end point when $\Omega=0$. Infrared convergence, in this sense, requires $2N+\nu-1>0$, which is just the opposite of the ultraviolet convergence condition. Thus infrared properties will change dramatically as one passes from an even spacetime to the next odd spacetime. For example, I_0, which is IR divergent in $0+1$ and $1+1$ dimensions, first becomes IR convergent in $2+1$ dimensions. This should be compared and contrasted with the pattern of UV divergences [see Eq. (3.6) and Table II], where the dramatic changes take place as one passes from an odd spacetime to the next even spacetime. The pattern can be summarized as follows:

UV: $0+1$ / $1+1$, $2+1$ / $3+1$,... (divergences getting worse),

IR: $0+1$, $1+1$ / $2+1$, $3+1$ / $\cdots$ (convergence getting better),

where the slash indicates a dramatic change in the UV or IR properties. This pattern has an important bearing on the physics of $\lambda\phi^4$ theories as a function of dimension, as we shall see.

C. Definition of the renormalized parameters

As explained in Sec. III A we find it convenient to define our renormalized parameters in terms of derivatives of the GEP at the origin. To find out how they are related to the bare parameters we need to examine the derivatives of $\overline{V}_G(\phi_0)$. It will be important to remember that $\overline{\Omega}$ is implicitly ϕ_0 dependent, since it obeys (2.15). This implies

$$\frac{d\overline{\Omega}}{d\phi_0}=\frac{\phi_0}{\overline{\Omega}}\frac{12\lambda_B}{(1+6\lambda_B I_{-1})}\,. \tag{3.12}$$

TABLE III. The first few $L_i(x)$ functions from Eqs. (3.6)–(3.9). Also shown are their expansions in $y\equiv(x-1)$ or $z\equiv\sqrt{x}-1$.

$\nu=$odd	
$L_1(x)=\ln x$	$=y(1-\frac{1}{2}y+\frac{1}{3}y^2-\cdots)$
$L_2(x)=x\ln x-(x-1)$	$=\frac{1}{2}y^2(1-\frac{1}{3}y+\cdots)$
$L_3(x)=\frac{1}{4}[2x^2\ln x-2(x-1)-3(x-1)^2]$	$=\frac{1}{6}y^3(1+\cdots)$
$\nu=$even	
$L_1(x)=(\sqrt{x}-1)$	$=z$
$L_2(x)=\frac{1}{3}(\sqrt{x}-1)^2(2\sqrt{x}+1)$	$=z^2(1+\frac{2}{3}z)$
$L_3(x)=\frac{1}{30}(\sqrt{x}-1)^3(8x+9\sqrt{x}+3)$	$=\frac{2}{3}z^3(1+\frac{5}{4}z+\frac{2}{5}z^2)$

Differentiating (2.16), one then finds

$$\frac{d\overline{V}_G}{d\phi_0}=\phi_0(m_B{}^2+4\lambda_B\phi_0{}^2+12\lambda_B I_0)\ . \tag{3.13}$$

[Alternatively, one could have noted that a *partial* differentiation of (2.13), treating Ω as constant, would have led to the correct result, since, by definition, the Ω dependence of $V_G(\phi_0,\Omega)$ vanishes when $\Omega=\overline{\Omega}$.]

In passing, we note that the condition for a stationary point of $\overline{V}_G(\phi_0)$ away from the origin is that the expression in parentheses in (3.13) should vanish. Using the $\overline{\Omega}$ equation this can be simplified to

$$\overline{\Omega}^2=8\lambda_B\phi_0{}^2\ . \tag{3.14}$$

Differentiating $\overline{V}_G$ again gives

$$\frac{d^2\overline{V}_G}{d\phi_0{}^2}=m_B{}^2+12\lambda_B(I_0+\phi_0{}^2)-\frac{I_{-1}(12\lambda_B\phi_0)^2}{(1+6\lambda_B I_{-1})}\ . \tag{3.15}$$

Evaluating this at the origin one has[4]

$$m_R{}^2\equiv\left.\frac{d^2\overline{V}_G}{d\phi_0{}^2}\right|_{\phi_0=0}=m_B{}^2+12\lambda_B I_0(\overline{\Omega}_0)\ , \tag{3.16}$$

where $\overline{\Omega}_0$ is the solution to the $\overline{\Omega}$ equation at $\phi_0=0$. But the $\overline{\Omega}$ equation, (2.15), shows that the right-hand side (RHS) of the above equation is just $\overline{\Omega}_0{}^2$, and so we see that

$$m_R{}^2=\overline{\Omega}_0{}^2\ . \tag{3.17}$$

This identification implies that $m_R{}^2$ is positive definite, so that the origin is always a minimum of the GEP. Later on, we shall see that, in the Gaussian approximation, m_R is indeed the physical particle mass.

Combining the last two equations allows us to express $m_B{}^2$ in terms of $m_R{}^2$ (Ref. 4):

$$m_B{}^2=m_R{}^2-12\lambda_B I_0(m_R)\ . \tag{3.18}$$

This equation will enable us to eliminate $m_B{}^2$ in favor of m_R.

Differentiating (3.15) twice more, keeping only terms which contribute at $\phi_0=0$ gives[4]

$$\lambda_R\equiv\frac{1}{4!}\left.\frac{d^4V_G}{d\phi_0{}^4}\right|_{\phi_0=0}=\lambda_B\frac{[1-12\lambda_B I_{-1}(m_R)]}{[1+6\lambda_B I_{-1}(m_R)]}\ . \tag{3.19}$$

In $1+1$ and $2+1$ dimensions I_{-1} is a convergent integral, so there is a finite relationship between λ_R and λ_B, confirming our expectation that, in these low dimensions, λ_B is finite. In $3+1$ dimensions I_{-1} is logarithmically divergent, and the situation is more subtle. We shall come to this in Sec. V.

It will be convenient later, especially when discussing numerical results, to work in units of m_R. To this end we define

$$\Phi_0{}^2=\phi_0{}^2/m_R{}^{\nu-1},\quad \hat{\lambda}_B=\lambda_B/m_R{}^{3-\nu}\ ,$$
$$\mathscr{V}_G=(V_G-D)/m_R{}^{\nu+1},\quad x=\Omega^2/m_R{}^2\ . \tag{3.20}$$

IV. THE PHYSICS OF $\lambda\phi^4$ THEORIES

A. $0+1$ dimensions

We are now ready to discuss the GEP results for $\lambda\phi^4$ theories. Our aim is to provide an overall picture of the basic physics, and how it evolves from $0+1$ to $3+1$ dimensions.

In $0+1$ dimensions the system is the familiar anharmonic oscillator, and the "integrals" I_1 and I_0 become

$$I_1=\tfrac{1}{2}\Omega,\quad I_0=1/(2\Omega)\ , \tag{4.1}$$

so that Eqs. (2.13) and (2.15) reproduce the previous results in I.

First, we dispose of the possibility that λ_B is negative. It is intuitively obvious that such a theory cannot be viable, but it is important to see just what does happen. For negative λ_B the expression for $V_G(\phi_0,\Omega)$, Eq. (2.13), has no finite minimum in Ω, since it tends to $-\infty$ as $\Omega\to 0$, the dominant term being $3\lambda_B I_0{}^2=-\frac{3}{4}\,|\,\lambda_B\,|\,\Omega^{-2}$. Thus $\overline{V}_G(\phi_0)$ is not just unbounded below for large ϕ_0, it is actually $-\infty$ everywhere. Note that the disaster is an infrared effect, being associated with the massless limit $\Omega\to 0$.

In $0+1$ dimensions the bare parameters $m_B{}^2$ and λ_B are, of course, finite, and they provide perhaps the most compact parametrization of the results. However, to facilitate comparison with the results in higher dimensions, we choose to swap m_B for m_R using

$$m_B{}^2=m_R{}^2-6\lambda_B/m_R,\qquad m_R>0\ . \tag{4.2}$$

The original parameter space $-\infty<m_B{}^2<\infty$, $0<\lambda_B<\infty$, maps 1-to-1 onto $0<m_R<\infty$, $0<\lambda_B<\infty$. Negative values of $m_B{}^2$ are associated with large values of $\hat{\lambda}_B\equiv\lambda_B/m_R{}^3$ ($>\frac{1}{6}$). Hence we expect to see a single-well effective potential for small $\hat{\lambda}_B$, evolving into a double-well shape as $\hat{\lambda}_B$ increases. This is indeed what we find in Fig. 1.

To plot the figure we have taken Eqs. (2.16) and (2.15) together with (4.1), and eliminated m_B using (4.2). Scaling the variables by m_R, as in Eq. (3.20) gives

$$\mathscr{V}_G(\Phi_0)=\tfrac{1}{2}(\sqrt{x}-1)-\tfrac{3}{4}\hat{\lambda}_B\left[\frac{1}{x}-1\right]$$
$$+\tfrac{1}{2}(1-6\hat{\lambda}_B)\Phi_0{}^2+\hat{\lambda}_B\Phi_0{}^4 \tag{4.3}$$

with

$$(\sqrt{x}\,)^3-\sqrt{x}\,(1-6\hat{\lambda}_B+12\hat{\lambda}_B\Phi_0{}^2)-6\hat{\lambda}_B=0\ . \tag{4.4}$$

Note that $\mathscr{V}_G$ is defined with the constant D subtracted, so that $\mathscr{V}_G(0)=0$.

It is convenient to define a "critical $\hat{\lambda}_B$" which marks the transition from single-well to double-well behavior. For $\hat{\lambda}_B<\hat{\lambda}_{B,\text{crit}}$ the global minimum of the GEP is at $\Phi_0=0$, while for $\hat{\lambda}_B>\hat{\lambda}_{\text{crit}}$ the global minima occur at $\Phi_0=\pm c$, $c\neq 0$. For $\hat{\lambda}_B=\hat{\lambda}_{\text{crit}}$ the GEP has three exactly degenerate minima. Numerically, we find

$$\hat{\lambda}_{B,\text{crit}}=1.149\,285\ . \tag{4.5}$$

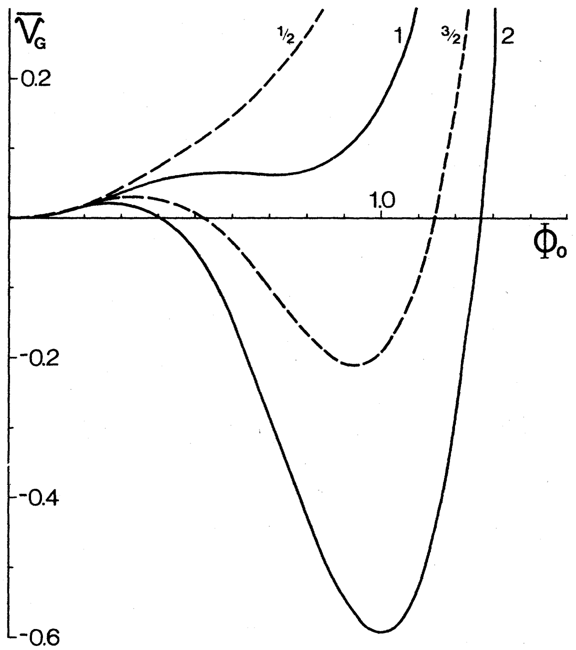

FIG. 1. The GEP for ϕ^4 in 0 + 1 dimensions. $\hat{\lambda}_B=\frac{1}{2},1,\frac{3}{2},2$. (Only one-half of the symmetrical potential is shown.)

In 0 + 1 dimensions the precise value of $\hat{\lambda}_{B,\mathrm{crit}}$ has no dramatic significance, because there is no phase transition. The physics evolves perfectly smoothly. This is related to the fact that the double-well regime does not have spontaneous symmetry breaking (SSB), strictly speaking. The two would-be vacua at $\Phi_0=+c$ and $\Phi_0=-c$ can communicate with each other by quantum-mechanical (QM) tunneling through the barrier separating them. Consequently, they mix with each other, and the true eigenstates are their symmetric and antisymmetric combinations, which are split by a tiny energy difference (exponentially small in $\hbar$).[20] The true vacuum state is the symmetric combination. There is no SSB because the vacuum expectation value of ϕ remains zero.

In semiclassical terms we could describe the vacuum as having a domain structure in time. For a long period of time T the system lives in one well, and the expectation value of ϕ averaged over a time scale $\lesssim T$ equals $+c$, say. Then quite abruptly, the system shifts itself to the other well, where the local average of ϕ is $-c$. Over very long time scales, $\gg T$, the average of ϕ is zero.[20] (Naively it would seem that this switching back and forth between "vacua" would only raise the energy. This is a fault of the semiclassical description, which loses sight of the quantum coherence of the process.) Note that, if the tunneling time T is very large—compared to the time scale of some experiment—then the system behaves as though it did have SSB, effectively.

These are familiar facts about quantum mechanics. The moral for interpreting the GEP is that double-well behavior is not necessarily synonymous with SSB. The GEP is quite correct in giving a double-well shape, for appropriate values of the parameters, because that is the natural way to describe the physics. The question of whether the two wells mix, and, if so, what effect this has on the physics, is not answered by the GEP: That requires a separate calculation (see Ref. 20). In a sense this question is only relevant in the transitional region (where we know the GEP is quantitatively least accurate[1]), because in the extreme double-well case, as we remarked above, the physics is "effectively" characterized by SSB. (These subtleties are characteristic of *low* dimensions, and will not bother us in 2 + 1 and 3 + 1 dimensions.)

B. 1 + 1 dimensions

In 1 + 1 dimensions the integrals I_1, I_0 are divergent, and we shall now see explicitly how the mass reparametrization leads to a manifestly finite result. Substituting (3.18) into (2.13) leads to

$$V_G(\phi_0,\Omega)=\tfrac{1}{2}m_R^2\phi_0^2+\lambda_B\phi_0^4+I_1+\tfrac{1}{2}(m_R^2-\Omega^2)I_0 +6\lambda_B[I_0-I_0(m_R)]\phi_0^2+3\lambda_B I_0^2 -6\lambda_B I_0 I_0(m_R)\ . \tag{4.6}$$

Using the formula in Table II for $I_1-I_1(m_R)$ gives

$$V_G(\phi_0,\Omega)=I_1(m_R)+\tfrac{1}{2}(m_R^2-\Omega^2)[I_0-I_0(m_R)] -L_2(x)m_R^2/(8\pi)+\tfrac{1}{2}m_R^2\phi_0^2+\lambda_B\phi_0^4 +6\lambda_B[I_0-I_0(m_R)]\phi_0^2 +3\lambda_B[I_0-I_0(m_R)]^2-3\lambda_B[I_0(m_R)]^2\ . \tag{4.7}$$

The first and last terms comprise the divergent constant D, Eq. (3.1), and the remaining terms can be simplified using the formula for $I_0-I_0(m_R)$. Hence,

$$V_G(\phi_0,\Omega)=D+\tfrac{1}{2}m_R^2\phi_0^2+\lambda_B\phi_0^4-L_2(x)m_R^2/(8\pi) -\frac{L_1(x)}{4\pi}[\tfrac{1}{2}m_R^2(1-x)+6\lambda_B\phi_0^2 -3\lambda_B L_1(x)/(4\pi)]\ . \tag{4.8}$$

Similarly, we can rewrite the $\overline{\Omega}$ equation, (2.15), by eliminating m_B^2 and using the $[I_0-I_0(m_R)]$ formula. This gives

$$\tfrac{1}{2}m_R^2(x-1)=6\lambda_B[\phi_0^2-L_1(x)/(4\pi)]\ . \tag{4.9}$$

From Table III, the functions L_1, L_2 are

$$L_1(x)=\ln x,\quad L_2(x)=x\ln x-(x-1)\ . \tag{4.10}$$

We first check the end-point values $\Omega=0$ and ∞ ($x=0$ and ∞), in case they should override the $\overline{\Omega}$ equation. In fact they do not, for $\Omega\to\infty$ gives $V_G\to\Omega^2/8\pi\to+\infty$, and $\Omega\to 0$ gives $V_G\to 3\lambda_B(\ln x/4\pi)^2$, which tends to $+\infty$ provided $\lambda_B>0$. Note that, just as in 0 + 1 dimensions, a negative λ_B would give $\overline{V}_G=-\infty$ everywhere.

We can tidy up Eqs. (4.8) and (4.9) by scaling the variables with powers of m_R, as specified by (3.20). Also we can use (4.9) to simplify (4.8) by eliminating the explicit logarithms, leaving us with

$$\overline{\mathscr{V}}_G(\Phi_0)=-2\hat{\lambda}_B\Phi_0^4+\frac{(x-1)}{24\hat{\lambda}_B}[1+3\hat{\lambda}_B/\pi+\tfrac{1}{2}(x-1)]\ , \tag{4.11}$$

where x is obtained by solving

$$(x-1)+(3\hat{\lambda}_B/\pi)\ln x=12\hat{\lambda}_B\Phi_0^2\ . \qquad (4.12)$$

The results obtained from these formulas are illustrated in Fig. 2, and they reproduce the earlier results of Chang.[6]

The qualitative behavior is similar to the (0 + 1)-dimensional case: The GEP evolves from a single-well to a double-well shape as $\hat{\lambda}_B$ increases, with the critical $\hat{\lambda}_B$ being

$$\hat{\lambda}_{B,\mathrm{crit}}=2.552\,7045\ . \qquad (4.13)$$

Recalling the discussion in the preceding section, we should not jump to the conclusion that there is a first-order phase transition to an SSB phase at this value of $\hat{\lambda}_B$. In fact, as discussed by Chang,[7] this would violate certain rigorous theorems.[21]

By analogy with the QM case, one might imagine that the vacuum at large $\hat{\lambda}_B$, described semiclassically, has a domain structure—in space as well as time—with each domain having $\langle\phi\rangle_{\mathrm{local}}=+c$ or $-c$, so that the overall ϕ is still zero. The point is that in 1 + 1 dimensions, as in 0 + 1, the system can still pass from one vacuum to another through configurations of finite energy: a soliton-antisoliton pair moving slowly apart, creating a domain of the other vacuum, will do the trick. If this picture is correct then the vacuum would be a coherent "gas" of solitons.

To discuss this question further would take us outside the scope of this paper.[22] The GEP definitely indicates a fairly dramatic change in the physics at $\hat{\lambda}_B$ around 2.5, but does not say if this is a true phase transition. Possibly the GEP could be used as a starting point for a semiclassical tunneling calculation of the intervacuum mixing.

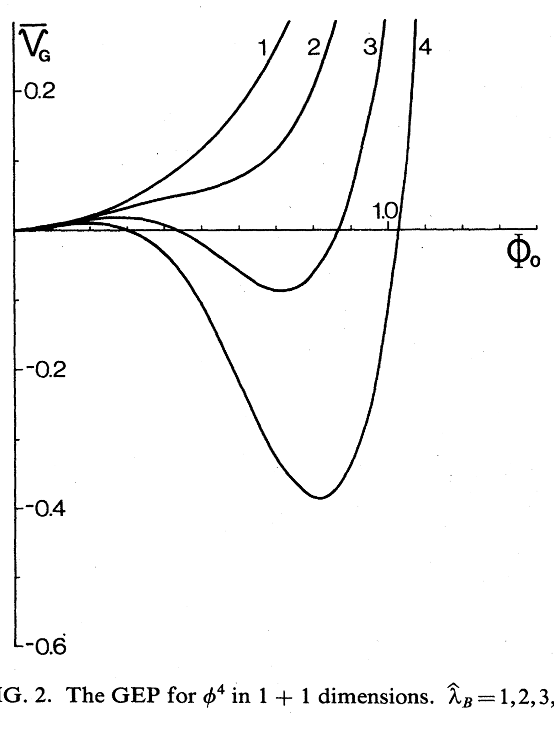

FIG. 2. The GEP for ϕ^4 in 1 + 1 dimensions. $\hat{\lambda}_B=1,2,3,4$.

C. 2 + 1 dimensions

Because the algebra of the divergent integrals in 2 + 1 dimensions is so similar to that in 1 + 1 dimensions (see Table II), we may simply take over the results in (4.8) and (4.9) except that now

$$L_1(x)=(\sqrt{x}-1)\ ,$$
$$L_2(x)=\tfrac{1}{3}(\sqrt{x}-1)^2(2\sqrt{x}+1) \qquad (4.14)$$

(and each occurrence of an L_i function will need a trivial extra factor of m_R).

We first check the end-point values $\Omega=0$ and ∞. As $\Omega\to\infty$, $V_G(\phi_0,\Omega)\to\Omega^3/(24\pi)\to+\infty$, so this is never relevant. However, $\Omega\to 0$ is more complicated, since, as remarked in Sec. III B a dramatic change in the infrared behavior of the integrals takes place between 1 + 1 and 2 + 1 dimensions. Thus $\Omega=0$ no longer leads to an infinite V_G, but gives

$$V_G(\phi_0,\Omega=0)=D+\tfrac{1}{2}m_R{}^2\phi_0{}^2+\lambda_B\phi_0{}^4 + \frac{m_R{}^3}{12\pi}\left[1+\frac{9\lambda_B}{4\pi m_R}\right]+\frac{3\lambda_B}{2\pi}m_R\phi_0{}^2\ . \qquad (4.15)$$

We observe that a negative λ_B is still unacceptable, because the GEP would be unbounded below at large ϕ_0. [However, at least we are no longer finding that λ_B negative implies $\overline{V}_G(\phi_0)=-\infty$ everywhere.] We therefore continue to restrict λ_B to positive values.

The $\overline{\Omega}$ equation, (4.9), now becomes [employing the scaled variables of (3.20) again]

$$\tfrac{1}{2}(x-1)=6\hat{\lambda}_B[\Phi_0{}^2-(\sqrt{x}-1)/(4\pi)]\ . \qquad (4.16)$$

In this case we are favored with an analytically soluble equation, so we can write explicitly

$$\sqrt{x}=-\frac{3\hat{\lambda}_B}{2\pi}+\left[\left[1+\frac{3\hat{\lambda}_B}{2\pi}\right]^2+12\hat{\lambda}_B\Phi_0{}^2\right]^{1/2}\ . \qquad (4.17)$$

(The sign of the square root is determined by the requirement that $\sqrt{x}\equiv\Omega/m_R>0$.) If we use (4.9) to simplify (4.8) by removing the $6\lambda_B\phi_0{}^2$ term we can express $\overline{V}_G(\Phi_0)$ in the convenient form

$$\overline{\mathscr{V}}_G(\Phi_0)=\tfrac{1}{2}\Phi_0{}^2+\hat{\lambda}_B\Phi_0{}^4 - \frac{(\sqrt{x}-1)^2}{24\pi}\left[1+\frac{9\hat{\lambda}_B}{2\pi}+2\sqrt{x}\right]\ . \qquad (4.18)$$

[It is easy to see that (4.18) is less than (4.15), which assures us that the $\overline{\Omega}$ equation, and not the $\Omega=0$ end point, is giving the global minimum of $V_G(\phi_0,\Omega)$.]

By computing (4.18) with $\sqrt{x}$ given by (4.17) we have obtained the results shown in Fig. 3. Again, we see a transition from single-well to double-well behavior, with the critical $\hat{\lambda}_B$ being

$$\hat{\lambda}_{B,\text{crit}}=3.078\,404 \ . \tag{4.19}$$

In this case, we do believe that this indicates a phase transition to an SSB phase. [Of course (4.19) is only an approximation to the exact critical coupling where the phase transition occurs.] Unlike the previous, lower-dimensional cases, we do not expect mixing between the $\Phi_0=c$ and $\Phi_0=-c$ vacua, because the system cannot pass from one to the other through configurations of finite energy. For instance, if one creates a circular bubble of one vacuum inside the other, the energy in the wall will become larger and larger as the bubble is expanded. This is quite different from $1+1$ dimensions where the energy in the domain walls (the soliton and antisoliton) is independent of the domain size. Again, one can correlate this difference with the dramatic change in infrared properties between $1+1$ and $2+1$ dimensions.

D. 3 + 1 dimensions

Ultraviolet behavior changes drastically when we pass to $3+1$ dimensions, because now I_{-1} is divergent. Proceeding as before; eliminating $m_B{}^2$ and using the formulas of Table II, we are left with

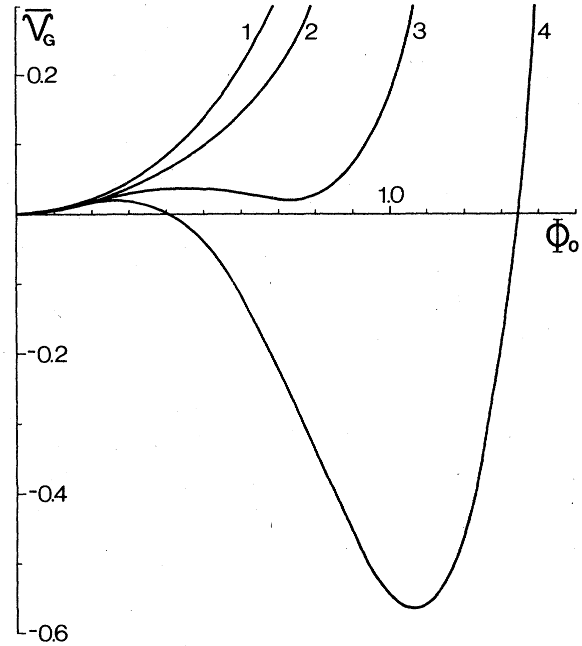

FIG. 3. The GEP for ϕ^4 in $2+1$ dimensions. $\hat{\lambda}_B=1,2,3,4$.

$$\begin{aligned}V_G(\phi_0,\Omega)=&D+\tfrac{1}{2}m_R{}^2\phi_0{}^2+\lambda_B\phi_0{}^4+L_3(x)m_R{}^4/(32\pi^2)+\tfrac{3}{4}\lambda_B m_R{}^4(x-1)^2[I_{-1}(m_R)]^2\\&+\tfrac{1}{8}I_{-1}(m_R)m_R{}^4(x-1)\left[(x-1)-\frac{3\lambda_B}{2\pi^2}[L_2(x)+16\pi^2\phi_0{}^2/m_R{}^2]\right]\\&-L_2(x)\frac{m_R{}^4}{32\pi^2}\left[(x-1)-\frac{3\lambda_B}{8\pi^2}[L_2(x)+32\pi^2\phi_0{}^2/m_R{}^2]\right] .\end{aligned} \tag{4.20}$$

Similarly, the $\overline{\Omega}$ equation becomes

$$(x-1)[1+6\lambda_B I_{-1}(m_R)]=\frac{3\lambda_B}{4\pi^2}[L_2(x)+16\pi^2\phi_0{}^2/m_R{}^2] \ . \tag{4.21}$$

Both these equations contain the divergent integral I_{-1}. If we had introduced a cutoff, I_{-1} would be of the form $\ln(M_{\text{UV}}{}^2/m_R{}^2)/(8\pi^2)$ as $M_{\text{UV}}\to\infty$. We shall therefore treat I_{-1} as arbitrarily large and positive, taking the limit $I_{-1}\to\infty$ *before* any other limit we might wish to investigate. (This last point is important, and will be discussed further in Secs. V C and V D.)

The relation between λ_B and λ_R, Eq. (3.19), allows the possibility that λ_B is still a finite parameter. We shall now show that this does not lead to a viable theory.

First, we dispose of the case that λ_B is a finite, negative number. In this case, as in lower dimensions, the $\overline{\Omega}$ equation gives a *maximum* of $V_G(\phi_0,\Omega)$, and one must look instead at the $\Omega=0$ end point. For $x=0$ the $L_i(x)$ functions are just constants, and so the result is

$$V_G(\phi_0,\Omega=0)=\text{constant}+\left[\tfrac{1}{2}+3\lambda_B I_{-1}(m_R)+\frac{3\lambda_B}{8\pi^2}\right]m_R{}^2\phi_0{}^2+\lambda_B\phi_0{}^4 \ . \tag{4.22}$$

This is clearly unbounded below: It has an infinitely negative mass term, as well as a negative $\phi_0{}^4$ term.

Now we consider λ_B finite and positive. In this case the $\overline{\Omega}$ equation does give the minimum of V_G, and it implies that x must be infinitesimally close to unity:

$$(x-1)=\frac{2\phi_0{}^2}{m_R{}^2}\frac{1}{I_{-1}}+O\left[\frac{1}{I_{-1}{}^2}\right] . \tag{4.23}$$

[Note that the $L_k(x)$ functions vanish like $(x-1)^k$ at $x=1$.] Substituting this in (4.20) yields a finite result, because each I_{-1} is accompanied by an $(x-1)$ factor. The result is simply

$$\overline{V}_G(\phi_0)=D+\tfrac{1}{2}m_R{}^2\phi_0{}^2-2\lambda_B\phi_0{}^4 , \tag{4.24}$$

since all other terms vanish as $I_{-1}\to\infty$. It should be no surprise that the $\phi_0{}^4$ coefficient is $-2\lambda_B$, since this follows from Eq. (3.19) for λ_R, taking λ_B finite. What we have now learned is that there are no ϕ_0^6, or higher, terms which might cause the potential to turn back upwards at large ϕ_0. Thus we conclude that a positive-λ_B theory is unstable. Since the GEP, being a variational approximation, should be an upper bound on the true effective potential, this amounts to a proof that there is no vacuum state in a continuum ϕ^4 theory with finite, positive λ_B.

To understand this rather surprising result, it is instructive to consider the situation in a cutoff version of the theory; for example, a lattice-regulated theory. If the lattice spacing is sufficiently small we can expect the GEP of the lattice system to closely resemble (4.24), *except when ϕ_0^2 is so large that it becomes greater than $I_{-1}^{\rm lat}$*, which is now finite. We can no longer discard terms suppressed by powers of $1/I_{-1}^{\rm lat}$ if these contain high powers of ϕ_0^2. Presumably, the presence of these terms in the lattice case must cause the GEP, which had been going down like $-2\lambda_B\phi_0^4$, to turn back up again as $\phi_0\to\infty$. (See Fig. 4.) This means that, while the GEP has a local minimum at $\phi_0=0$, its true minimum is much, much deeper, and lies out at some enormous value of ϕ_0^2, $\gtrsim I_{-1}^{\rm lat}$. That is, the lattice system will be in a SSB ("ordered") phase. As the inverse lattice spacing $M_{\rm UV}$ goes to infinity, the true vacuum, and hence all the physics, goes with it. (See Fig. 4.) This explains why we find nothing sensible when we look directly at the continuum theory, as in Eq. (4.24).[23]

We have now disposed of all finite values of λ_B, but there remains the possibility that λ_B is infinitesimal, of the form

$$\lambda_B = A/I_{-1}(m_R) \tag{4.25}$$

(i.e., λ_B tends to zero in a particular way as the cutoff

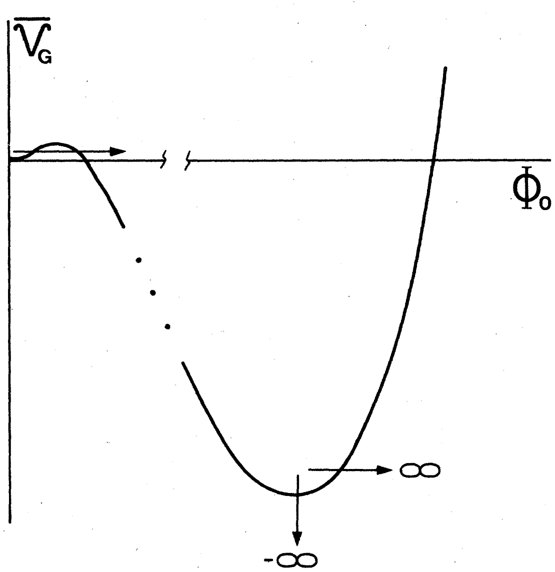

FIG. 4. A sketch of the GEP for a lattice-regularized ϕ^4 theory with finite, positive λ_B in 3 + 1 dimensions. The true vacuum of the system moves out to infinity as the continuum limit is approached, leaving the continuum theory with an unbounded-below effective potential.

tends to infinity). There are three cases: (i) $A < -\frac{1}{6}$, (ii) $A > -\frac{1}{6}$, and (iii) $A = -\frac{1}{6}+O(1/I_{-1})$. [The reason for this classification is that the condition for the $\overline{\Omega}$ equation to give a minimum, not a maximum, of $V_G(\phi_0,\Omega)$ is

$$1+6\lambda_B I_{-1}(m_R)>0 \;, \tag{4.26}$$

as one can easily derive from differentiating (2.13) twice.] In case (i) the $\overline{\Omega}$ equation does not give a minimum, and one should consider the end point $\Omega=0$. This case is essentially the same as the finite, negative λ_B case we analyzed earlier. One again obtains (4.22), which then reduces to

$$V_G(\phi_0,\Omega=0) = \text{constant} + \tfrac{1}{2}m_R^2(1+6A)\phi_0^2 + O\left[\frac{1}{I_{-1}}\right] . \tag{4.27}$$

Because $1+6A$ is negative, by assumption, this potential is unbounded below.

In case (ii), $A > -\frac{1}{6}$, the $\overline{\Omega}$ equation does provide the minimum of $V_G(\phi_0,\Omega)$. The analysis now proceeds just as for finite, positive λ_B. The $\overline{\Omega}$ equation again requires x to be infinitesimally close to unity:

$$(x-1) = \frac{12A}{(1+6A)}\frac{\phi_0^2}{m_R^2}\frac{1}{I_{-1}(m_R)} + O\left[\frac{1}{I_{-1}^2}\right] \tag{4.28}$$

[cf. (4.23)]. When we substitute into (4.20), all that remains is

$$\overline{V}_G(\phi_0) = D + \tfrac{1}{2}m_R^2\phi_0^2 \; . \tag{4.29}$$

All the other terms—including now the ϕ_0^4 terms—vanish as $I_{-1}\to\infty$. This should be no surprise since the coefficient of ϕ_0^4, namely, λ_R, as given in (3.19), becomes in this case

$$\lambda_R = \frac{A}{I_{-1}}\frac{(1-12A)}{(1+6A)} = O\left[\frac{1}{I_{-1}}\right] \to 0 \; . \tag{4.30}$$

This case corresponds to the "triviality" scenario,[11–13] since the resulting theory is merely a free field theory.

Our conclusions are thus in perfect accord with Refs. 11–13: If we start with a *positive*-λ_B lattice theory, we have two equally unpalatable alternatives. Either we keep λ_B finite, in which case the vacuum is driven to infinity, and we end up with an unstable continuum theory [Eq. (4.24)]: Or we let λ_B tend to zero like A/I_{-1}, in which case we end up with a free field theory [Eq. (4.29)].

However, one case remains, namely, $A = -\frac{1}{6} + O(1/I_{-1})$, and this is where it becomes interesting.

V. "PRECARIOUS ϕ^4"

A. The efficacious form of λ_B

We could have guessed that $\lambda_B \simeq -1/(6I_{-1})$ would be the important case, just from a perusal of the coupling constant renormalization [Eq. (3.19)]:

$$\lambda_R = \lambda_B\frac{[1-12\lambda_B I_{-1}(m_R)]}{[1+6\lambda_B I_{-1}(m_R)]} \; . \tag{5.1}$$

Taking the inverse of this relationship one has

$$\lambda_B=\lambda_R\frac{(1-6\lambda_R I_{-1})}{24\lambda_R I_{-1}}\left[1\pm\left[1-\frac{48\lambda_R I_{-1}}{(1-6\lambda_R I_{-1})^2}\right]^{1/2}\right], \tag{5.2}$$

which is double valued, in general. (See Fig. 5.) Taking I_{-1} arbitrarily large there are two possibilities that give a finite λ_R:

$$\lambda_B=-\tfrac{1}{2}\lambda_R+O(1/I_{-1}), \tag{5.3}$$

which we have already dismissed, and

$$\lambda_B=-\frac{1}{6I_{-1}(m_R)}\left[1+\frac{1}{2\lambda_R I_{-1}(m_R)}+\cdots\right], \tag{5.4}$$

which we now discuss.

It is worthwhile noting that (5.4) can be written simply as

$$\lambda_B=-\frac{1}{6I_{-1}(\Lambda)}, \tag{5.5}$$

where

$$\ln(m_R{}^2/\Lambda^2)=-\frac{4\pi^2}{\lambda_R}\equiv\kappa \tag{5.6}$$

by virtue of the formula for $I_{-1}(\Lambda)-I_{-1}(m_R)$ in Table II. The parameter Λ is a finite "characteristic scale" parameter, analogous to the Λ parameter of QCD. (It should not be confused with a UV cutoff.) Equation (5.5) is a pure manifestation of the "dimensional transmutation" phenomenon.[17,24] The parameter κ introduced in (5.6) will be very important in what follows.

If we evaluate $V_G(\phi_0,\Omega)$, Eq. (4.20) with λ_B given by (5.4), we observe several cancellations, which leave

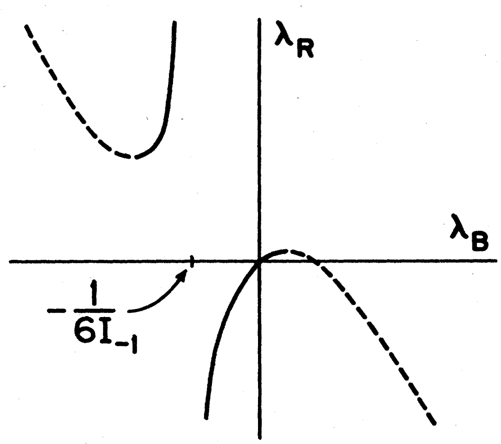

FIG. 5. A sketch (not to scale) of the λ_R,λ_B relationship, (5.1), with I_{-1} regularized. The dashed and solid curves correspond to the plus and minus sign, respectively, in Eq. (5.2).

$$V_G(\phi_0,\Omega)=D+\tfrac{1}{2}xm_R{}^2\phi_0{}^2-\frac{m_R{}^4}{16\lambda_R}(x-1)^2+\frac{L_3(x)m_R{}^4}{32\pi^2}, \tag{5.7}$$

dropping terms of order $1/I_{-1}$ or smaller. Observe that V_G is now finite for any Ω. The $\overline{\Omega}$ equation is also manifestly finite, since (5.4) inserted into (4.21) gives

$$(x-1)=\frac{\lambda_R}{4\pi^2}[L_2(x)+16\pi^2\phi_0{}^2/m_R{}^2]. \tag{5.8}$$

We are now in a position to calculate and plot some numerical results. First we tidy up the formulas in three ways: We scale the variables with respect to m_R [see Eq. (3.20)]; we use the parameter $\kappa=-4\pi^2/\lambda_R$, introduced above, in place of λ_R; and finally we take $L_2(x),L_3(x)$ from Table III. This gives us

$$\mathscr{V}_G(\Phi_0,x)=\tfrac{1}{2}x\Phi_0{}^2+\frac{1}{128\pi^2}[2x^2\ln x-2(x-1)-(3-2\kappa)(x-1)^2] \tag{5.9}$$

with x given by

$$(x-1)(1-\kappa)-16\pi^2\Phi_0{}^2=x\ln x, \tag{5.10}$$

or by the end point $x=0$, whichever gives the smaller result in $\overline{\mathscr{V}}_G$. We note that $x=0$ in (5.9) gives

$$\mathscr{V}_G(\Phi_0,0)=(2\kappa-1)/(128\pi^2), \tag{5.11}$$

which is a constant, independent of Φ_0. If, however, the appropriate x is given by (5.10), then we may use this equation to eliminate the logarithm in (5.9):

$$\overline{\mathscr{V}}_G(\Phi_0)=\tfrac{1}{4}x\Phi_0{}^2-\frac{(x-1)}{128\pi^2}(x-1+2\kappa). \tag{5.12}$$

Next, we need to understand the solutions to the transcendental equation (5.10), which we can do by plotting each side as a function of x. (See Fig. 6.) For the present we shall assume that κ is positive, which implies a negative λ_R. (The next section is devoted to showing that this assumption involves no loss of generality.) For $0<\kappa<1$ Eq. (5.10) has two solutions when $\Phi_0{}^2=0$. As $\Phi_0{}^2$ increases, one solution starts at $x=1$ and decreases, while the other starts at smaller x and increases. The first solution gives a minimum, the second a maximum, of $\mathscr{V}_G$. At some critical value of $\Phi_0{}^2$,

$$\Phi_{0,\text{crit}}{}^2=(e^{-\kappa}+\kappa-1)/(16\pi^2), \tag{5.13}$$

the two solutions coalesce, and thereafter disappear. For $\kappa>1$ the situation is similar, except that the second solution is not present at $\Phi_0=0$: It appears later at $x=0$, when $\Phi_0{}^2$ has reached $(\kappa-1)/(16\pi^2)$, and increases until it coalesces with the first solution.

When $\Phi_0{}^2$ exceeds $\Phi_{0,\text{crit}}{}^2$ there is no solution to the $\overline{\Omega}$ equation, and $\mathscr{V}_G(\Phi_0,x)$ is minimized at the end point $x=0$. In fact the end point will take over as soon as the result for $\overline{\mathscr{V}}_G$, obtained using the $\overline{\Omega}$ equation, exceeds the value in Eq. (5.11). This occurs at

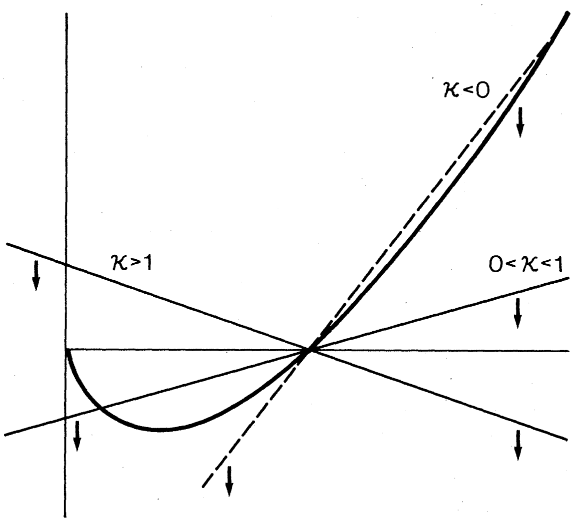

FIG. 6. Graphical solution of the $\overline{\Omega}$ equation, (5.10). The curve is the function $x\ln x$, while the straight lines represent the left-hand side of (5.10) with $\Phi_0{}^2=0$, for various ranges of κ. The arrows indicate how the lines move as $\Phi_0{}^2$ increases.

$$\Phi_{0,\text{break}}{}^2=[e^{1/2-\kappa}-2(1-\kappa)]/(32\pi^2)\ , \qquad (5.14)$$

which is just before $\Phi_0{}^2$ reaches $\Phi_{0,\text{crit}}{}^2$.

Thus the results have the form shown in Fig. 7. For weak coupling (large κ) the GEP is a slightly flattened parabola, which is "chopped off" at the top. As the coupling strength is increased there is little change to the almost-parabolic part, but the "chopping off" becomes much more drastic. Indeed, for couplings stronger than $-\lambda_R=8\pi^2$ ($\kappa<\frac{1}{2}$), the GEP is entirely given by the horizontal line that results from the $x=0$ end point. (See Fig. 8.)

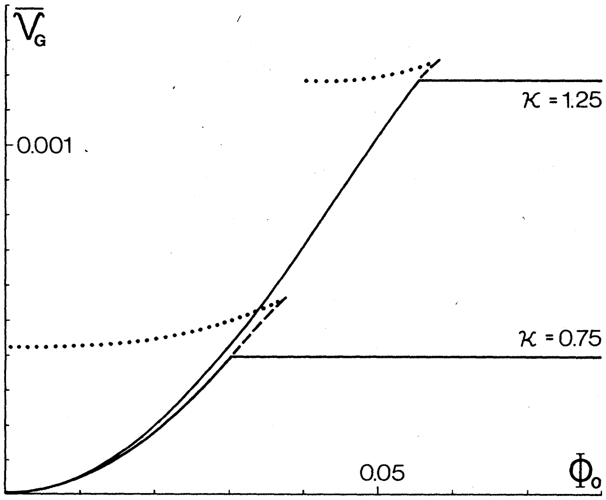

FIG. 7. The GEP for "precarious ϕ^4" in 3 + 1 dimensions, with $\kappa=-4\pi^2/\lambda_R=0.75$, 1.25. The dashed and dotted curves correspond to solutions of the $\overline{\Omega}$ equation which are local-but-not-global minima, and maxima of $V_G(\phi_0,\Omega)$, respectively. (See also Fig. 2 of Ref. 10, which is for $\kappa=1$.)

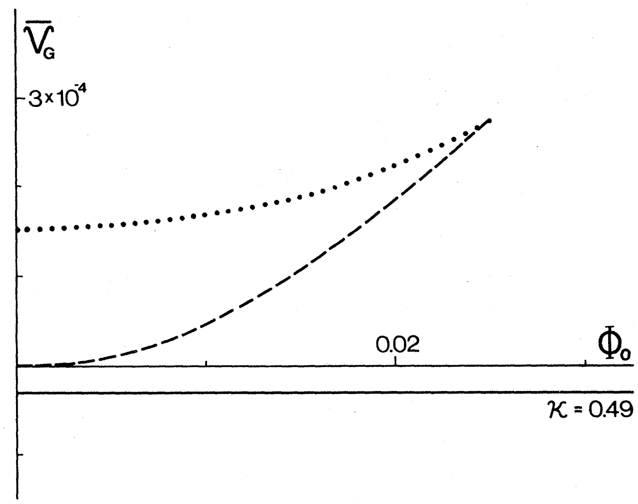

FIG. 8. As Fig. 7, but for $\kappa=0.49$. For $\kappa<\frac{1}{2}$ the GEP itself is a horizontal line. However, the dashed curve, arising from a local minimum of $V_G(\phi_0,\Omega)$, can be interpreted as a short-lived, metastable survival of the weak coupling phase.

What do these results mean? We first note the strong resemblance to the results found in the QM example of the "crater potential," studied in I. This analogy gives us confidence in the following straightforward interpretation. The theory appears to have two phases. One phase, at weak coupling (and low temperature) is a perfectly ordinary field theory, with massive particles interacting through a weak, attractive (because λ_R is negative) force. The other phase, at strong coupling (or high temperature), is characterized by a completely flat potential, which is the potential of a massless free field theory. We shall argue in Sec. V D, that this rather mysterious phase has a natural interpretation from an "effective field theory" viewpoint.

We should explain our allusions to finite temperatures in the above. $\overline{V}_G$ represents an energy density, and any system with a finite number of particles has an energy density infinitesimally close to that of the vacuum, since the particles' energy is diluted by an infinite volume factor. Therefore, all particle masses, scattering properties, etc., are associated with the behavior of the GEP in the immediate neighborhood of the origin. Only if we consider systems with finite energy density, such as finite-temperature systems, will the behavior of the GEP away from the origin become important. The "chopping off" of the potential implies that there is a critical energy density $m_R{}^4(2\kappa-1)/(128\pi^2)$, and hence a critical temperature, at which a transition to the "free, massless" (or whatever) phase occurs. These considerations imply the phase diagram of Fig. 9. We can interpret the dashed curve in Fig. 8 as a metastable survival of the weak coupling phase into the strong coupling regime (cf. the crater-potential example in I).

Many of the results here have counterparts in the earlier analysis of Bardeen and Moshe.[9] The analyses are not

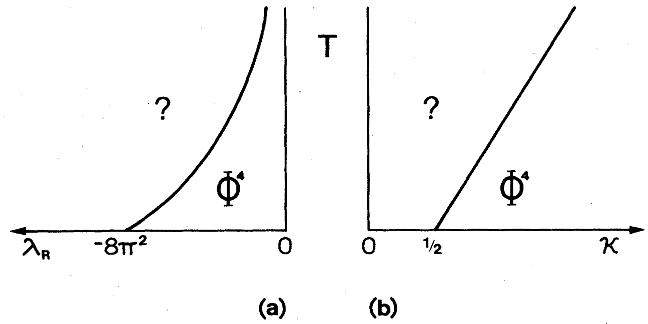

FIG. 9. The phase diagram suggested by the GEP results. The weak coupling phase, labeled "ϕ^4," is an ordinary, massive field theory. It becomes unstable to decay to a more mysterious phase, characterized by a flat GEP, at a temperature proportional to $(2\kappa-1)$. The diagram is shown in two equivalent forms: (a) temperature vs coupling constant λ_R, and (b) temperature vs κ, where $\kappa=-4\pi^2/\lambda_R$.

directly comparable since Ref. 9 studies the O(N)-symmetric generalization of ϕ^4 theory and employs the large-N approximation in addition to the GEP method.[25] They discuss the finite-temperature case in more detail, and show how to perform a finite-temperature generalization of the GEP calculation.

In the next section we dispose of the technical matter of why $\kappa<0$ gives nothing new. We then discuss the weak coupling limit, and the relation to perturbation theory and the one-loop effective potential. Finally, in Sec. V D we discuss the "paradox" of how a negative λ_B leads to a stable theory.

B. Why κ is positive

In the preceding section it was assumed that $\kappa\equiv-4\pi^2/\lambda_R$ was positive. If κ is negative then there are two solutions to the $\overline{\Omega}$ equation (5.10) at $\Phi_0=0$: One is at $x=1$, while the other is at some larger value of x (see Fig. 6). It turns out that $x=1$ is not a minimum, but a *maximum,* of V_G, and it is the other root that we should be using.

However, from the definition of $m_R{}^2$ as $d^2\overline{V}_G/d\phi_0{}^2\,|_{\phi_0=0}=\overline{\Omega}_0{}^2$, where $\overline{\Omega}_0$ is the solution to the $\overline{\Omega}$ equation at $\phi_0=0$, we ought to have $x=1$ at $\Phi_0=0$. The problem is that for negative κ we have inadvertently misdefined m_R by allowing the algebra to work with the wrong root for $\overline{\Omega}_0$. We can repair the damage by identifying the correct m_R and rescaling all the variables accordingly. From (5.10) at $\Phi_0=0$ it follows that the real $m_R{}^2$, denoted $m_R'{}^2$, is a factor x_0 bigger than the fake $m_R{}^2$, where

$$(x_0-1)(1-\kappa)=x_0\ln x_0 \quad (x_0>1)\ . \tag{5.15}$$

Thus we need to rescale the variables as follows:

$$\begin{aligned} x' &= x/x_0\ ,\\ \Phi_0'^2 &= \Phi_0{}^2/x_0\ ,\\ \kappa' &= \ln(m_R'{}^2/\Lambda^2)=\kappa+\ln x_0\ ,\\ \mathscr{V}'_G &= \mathscr{V}_G/x_0{}^2-\text{constant}\ . \end{aligned} \tag{5.16}$$

We must expect to subtract a constant from $\mathscr{V}_G$ because the algebra as it stands defines the zero of energy such that $\mathscr{V}_G(\Phi_0=0,\Omega=\overline{\Omega}_0)=0$ for the wrong $\overline{\Omega}_0$. Substituting into Eqs. (5.10) and (5.12) and using (5.15) to simplify the results, one discovers that these equations regain exactly their original forms in terms of the primed variables [except for the anticipated constant term in the case of (5.12)]. One also observes that κ' is positive: In fact, $0>\kappa>-\infty$ maps onto $0<\kappa'<1$. Hence the results obtained with negative κ are just a repeat of results obtained previously with a positive κ parameter. Thus there is no loss of generality involved in taking κ to be positive.

In terms of Fig. 7 what is happening is that the algebra with negative κ is trying to put us off with the dotted curve of the corresponding κ'. (Note that the dotted curve extends to the origin only for $\kappa'<1$.) By means of our repair work we have recovered the lower, solid curve, which is the true GEP.

This story is strongly reminiscent of the history of $1/N$-expansion investigations of O(N)-symmetric ϕ^4 theory. The original analysis[26] found the equivalent of the dotted curve, and observed that this phase contained tachyons. Later, it was shown that the one-loop effective potential actually had two branches,[27] and that the phase corresponding to the lower curve was free of tachyonic instabilities. The curves obtained show a remarkable similarity to Fig. 7.

Another way of understanding the situation is the following. Take the mass renormalization equation (3.18) and substitute $\lambda_B=-1/[6I_{-1}(\Lambda)]$. After some rearrangement, and use of the $I_0(m_R)-I_0(\Lambda)$ formula from Table II, this can be expressed as

$$\frac{16\pi^2}{\Lambda^2}\left[\tfrac{1}{2}(m_B{}^2-\Lambda^2)I_{-1}(\Lambda)-I_0(\Lambda)\right]=L_2(m_R{}^2/\Lambda^2)\ . \tag{5.17}$$

The left-hand side is a function only of the bare parameters; Λ being directly related to λ_B. The form of the right-hand side (see Fig. 10), reveals that two values of $m_R{}^2$ may correspond to the same bare parameters. Therefore, both these values of $m_R{}^2$ should correspond to the same physical theory. This is, indeed, just what we have found above. A theory with $m_R{}^2<\Lambda^2$ (negative κ) is just a reparametrization of an equivalent theory with $m_R{}^2>\Lambda^2$ (positive κ): It is a less convenient parametrization, in that the m_R parameter, for negative κ, does not represent the particle mass.

In conclusion, we emphasize the important consequences of the result that, effectively, $\kappa>0$. It means that the renormalized coupling λ_R is negative, corresponding to an attractive interaction. It implies that $m_R{}^2>\Lambda^2$ (and if we require $\kappa>\frac{1}{2}$, so that the weak coupling phase is

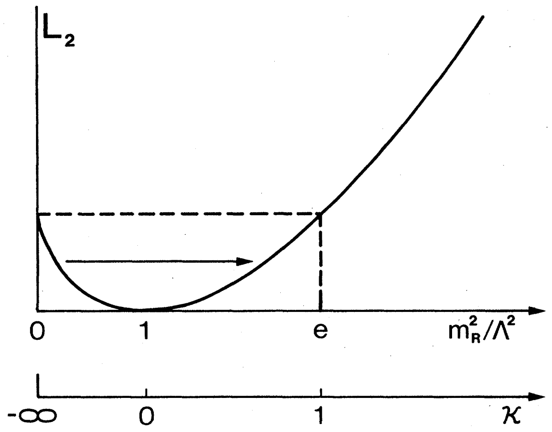

FIG. 10. Equation (5.17) in graphical form, showing how two values of $m_R{}^2$ can lead to the same value of $L_2(m_R{}^2/\Lambda^2)$, and hence must correspond to the same bare parameters. Theories with $m_R{}^2<\Lambda^2$ [i.e., $\kappa\equiv\ln(m_R{}^2/\Lambda^2)<0$] map 1-to-1 onto equivalent theories with $\Lambda^2<m_R{}^2<e\Lambda^2$ (i.e., $0<\kappa<1$). Only in the latter form of the theory does the parameter m_R have its usual interpretation as the particle mass.

stable, then we have the stronger condition, $m_R{}^2>e^{1/2}\Lambda^2$). Consequently, *there is no such thing as a weakly coupled, massless* ϕ^4 *theory*. Such an object can be formally studied in perturbation theory,[17] but it appears to be a mirage. (One could say that the theory has dynamical mass generation, because, in a sense, the *bare* mass is zero. See the Appendix.)

C. The weak coupling limit

In this section we undertake the question of how our results are related to perturbation theory and the one-loop potential. We begin by making a weak coupling expansion of the GEP results. That is, we consider the $\kappa\to\infty$ limit of Eqs. (5.10) and (5.12). Note that the chopping off of the potential by (5.11) can be ignored in this limit. Expanding $x\ln x$ about $x=1$ one can solve Eq. (5.10) as a power series in λ_R. The result is

$$(x-1)=H\left[\frac{\lambda_R}{4\pi^2}\right]\left[1+A_1\left[\frac{\lambda_R}{4\pi^2}\right]+A_2\left[\frac{\lambda_R}{4\pi^2}\right]^2+\cdots\right], \tag{5.18}$$

with

$$A_1=0,\quad A_2=\tfrac{1}{2}H,\quad A_3=-\tfrac{1}{6}H^2\,,$$
$$A_4=\tfrac{1}{2}H^2(1+\tfrac{1}{6}H),\quad A_5=-\tfrac{5}{12}H^3(1+\tfrac{3}{25}H)\,, \tag{5.19}$$

and

$$H\equiv 16\pi^2\Phi_0{}^2\,. \tag{5.20}$$

Substituting for $(x-1)$ in (5.12) then yields the weak coupling expansion of the GEP,

$$\overline{\mathscr{V}}_G(\Phi_0)=\tfrac{1}{2}\Phi_0{}^2+\lambda_R\Phi_0{}^4+\frac{1}{128\pi^2}\left[B_1\left[\frac{\lambda_R}{4\pi^2}\right]^2+B_2\left[\frac{\lambda_R}{4\pi^2}\right]^3+\cdots\right], \tag{5.21}$$

where

$$B_1=0,\quad B_2=\tfrac{2}{3}H^3,\quad B_3=-\tfrac{1}{6}H^4,\quad B_4=\tfrac{1}{2}H^4(1+\tfrac{2}{15}H)\,, \tag{5.22}$$

Note the absence of a $\lambda_R{}^2$ term.

Next we consider the one-loop effective potential. As explained in Sec. II of I, the one-loop result is contained in the GEP, and can be found by discarding the $3\lambda_B I_0{}^2$ term in (2.13), and correspondingly the $12\lambda_B I_0$ term in (2.15). That is,

$$V_{\text{1-loop}}=I_1+\tfrac{1}{2}(m_B{}^2-\Omega^2)I_0+\tfrac{1}{2}m_B{}^2\phi_0{}^2+\lambda_B\phi_0{}^4+6\lambda_B I_0\phi_0{}^2 \tag{5.23}$$

with

$$\Omega^2=m_B{}^2+12\lambda_B\phi_0{}^2\,. \tag{5.24}$$

Proceeding, as before, to define renormalized parameters through the derivatives at the origin one finds

$$m_R{}^2=m_B{}^2+12\lambda_B I_0(m_B{}^2)\,,$$
$$\lambda_R=\lambda_B[1-18\lambda_B I_{-1}(m_B{}^2)]\,. \tag{5.25}$$

The mass renormalization is subtly different from before because here $\overline{\Omega}_0{}^2=m_B{}^2$, so that $m_B{}^2$ not $m_R{}^2$ appears as the argument of I_0. Not only is the integral divergent, but so is its mass-parameter argument. However, if one treats the equation iteratively then it agrees with the previous result to first order in λ_B. Similarly, the result for λ_B reproduces the first nontrivial term in a perturbation expansion of (5.1).

Substituting the mass renormalization, and using (5.24), we can rewrite the one-loop result as

$$V_{\text{1-loop}}=I_1(m_B{}^2)+\tfrac{1}{2}m_R{}^2\phi_0{}^2+\lambda_R\phi_0{}^4+\frac{m_B{}^4}{32\pi^2}L_3(\Omega^2/m_B{}^2)\,. \tag{5.26}$$

The first term is the divergent vacuum-energy constant, which we are free to subtract [though note that it differs from D in (3.1)]. The remaining terms, using (5.24), are

$$V_{\text{1-loop}}=\tfrac{1}{2}m_R{}^2\phi_0{}^2+\lambda_R\phi_0{}^4+\frac{m_B{}^4}{64\pi^2}\left[\left[1+\frac{12\lambda_B\phi_0{}^2}{m_B{}^2}\right]^2\ln\left[1+\frac{12\lambda_B\phi_0{}^2}{m_B{}^2}\right]-\frac{12\lambda_B\phi_0{}^2}{m_B{}^2}-\frac{216\lambda_B{}^2\phi_0{}^4}{m_B{}^4}\right], \tag{5.27}$$

which coincides with the result of Coleman and Weinberg [Eq. (B5), after correction of an obvious misprint], *except* that the correction term involves bare and not renormalized quantities. This reflects the fact that the one-loop effective potential is not RG invariant: It does not renormalize properly by substitution. The result is only finite if we neglect "higher-order terms" in λ_R in Eq. (5.25) and replace m_B,λ_B by m_R,λ_R. This means that, strictly, we are only entitled to keep the first nontrivial term in the expansion of the L_3 term in powers of λ_R: i.e. (scaling the variables by $m_R{}^2$),

$$V_{\text{1-loop}}=\tfrac{1}{2}\Phi_0{}^2+\lambda_R\Phi_0{}^4+\frac{9}{\pi^2}\lambda_R{}^3\Phi_0{}^6+O(\lambda_R{}^4)\ . \qquad (5.28)$$

This makes the point that the loop expansion—even though each order involves summing an infinite number of Feynman diagrams—is still basically a *perturbative* expansion of the effective potential. It suffers from all the usual troubles of perturbation theory.

Comparing (5.28) with (5.21) and (5.22), we see that they do *not* agree: There is a factor of 27 discrepancy in the coefficient of the $\lambda_R{}^3\Phi_0{}^6$ term. We now want to explain the origin of this factor.

One of the diseases of perturbation theory is that it makes an unjustified interchange of limits. It assumes that one may introduce a UV cutoff to regulate I_{-1}, then make an expansion about $\lambda_R=0$, and finally send the cutoff to infinity. However in the real theory a cutoff is never present. If one is introduced, it ought to be sent back to infinity *before* the $\lambda_R\to 0$ limit is investigated. This is the order we used in deriving (5.21). The perturbative result (5.28) is different because it reverses the order of these limits. We can show this in the following manner.

Let us go back to the GEP in the form (4.20) and (4.21), where the mass renormalization has been performed, but not the coupling-constant renormalization. Instead of performing the coupling constant renormalization properly, as in Sec. V A, let us proceed perturbatively: That is, we expand (5.1) as

$$\lambda_R=\lambda_B(1-18\lambda_B I_{-1}+108\lambda_B{}^2 I_{-1}{}^2+\cdots)\ , \qquad (5.29)$$

or, inversely,

$$\lambda_B=\lambda_R(1+18\lambda_R I_{-1}+540\lambda_R{}^2 I_{-1}{}^2+\cdots)\ . \qquad (5.30)$$

Such expansions implicitly assume that I_{-1} is regulated (UV cutoff held finite) while the $\lambda_R\to 0$ limit, needed to generate the series expansion, is taken. With the proper ordering of the limits ($I_{-1}\to\infty$, while λ_R held finite), these expansions are pure nonsense. However, let us see where the perturbative route takes us.

Before beginning this exercise, we first put Eqs. (4.20) and (4.21) in more convenient form. Judicious use of (4.21) in (4.20), and scaling the variables by $m_R{}^2$, gives

$$\overline{\mathscr{V}}_G(\Phi_0)=\tfrac{1}{2}\Phi_0{}^2-2\lambda_B\Phi_0{}^4+\frac{1}{64\pi^2}\{(x-1)[H-L_2(x)]+2L_3(x)\} \qquad (5.31)$$

with $H\equiv 16\pi^2\Phi_0{}^2$, as above. Also, using (5.1), we can write (4.21) in the "mixed," but convenient form

$$(x-1)=\frac{3\lambda_R}{4\pi^2}\frac{1}{[1-12\lambda_B I_{-1}(m_R)]}[L_2(x)+H]\ . \qquad (5.32)$$

Substituting perturbatively for λ_B using (5.30) we can solve (5.32) for $(x-1)$ as a power series in λ_R:

$$(x-1)=3H\left[\frac{\lambda_R}{4\pi^2}\right]\left[1+A_1'\left[\frac{\lambda_R}{4\pi^2}\right]+A_2'\left[\frac{\lambda_R}{4\pi^2}\right]^2+\cdots\right]\ , \qquad (5.33)$$

where

$$A_1'=12(4\pi^2 I_{-1})\ ,$$

$$A_2'=360(4\pi^2 I_{-1})^2+\tfrac{9}{2}H\ ,$$

$$A_3'=13\,392(4\pi^2 I_{-1})^3+162H(4\pi^2 I_{-1})-\tfrac{9}{2}H^2\ ,$$

etc. This is quite different from (5.18), as it contains lots of I_{-1} terms. However, if we ignore the I_{-1} terms, we see that the coefficients agree except for systematic factors of 3. The reason is obvious when we compare (5.32) with (5.8): In one case $1/(1-12\lambda_B I_{-1})$ is a factor of $\frac{1}{3}$; in the other it is $1+(I_{-1})$ terms).

Substituting (5.33) into (5.31) gives

$$\overline{\mathscr{V}}_B^{\text{pert}}(\Phi_0)=\tfrac{1}{2}\Phi_0{}^2+\lambda_R\Phi_0{}^4+\frac{9}{\pi^2}\lambda_R{}^3\Phi_0{}^6+\frac{27}{\pi^2}\lambda_R{}^4\Phi_0{}^6(12I_{-1}-\Phi_0{}^2)+\cdots\ . \qquad (5.34)$$

Several cancellations of I_{-1}'s are encountered in deriving this result, but the cancellations fail in the $\lambda_R{}^4$ and higher terms. It is easy to see that, if we simply ignore the I_{-1} terms, the result reproduces (5.21) with systematic factors of 3 associated with each λ_R—except in the $\lambda_R\Phi_0{}^4$ term. The latter is a special case because the $-2\lambda_B\Phi_0{}^4$ term in (5.31) contributes $-2\lambda_R\Phi_0{}^4+(I_{-1}$ terms) in the perturbative case, while it contributes nothing in the proper procedure, being $O(1/I_{-1})$.

To summarize: The one-loop result (5.28) agrees perfectly with the GEP result (5.34) obtained using the illegal perturbative procedure, but differs from the true GEP result (5.21). The crucial point is that the $I_{-1}\to\infty$ and $\lambda_R\to 0$ limits do not commute, so that perturbation theory mistreats $(1-12\lambda_B I_{-1})$ as $1+O(\lambda_R)$, when it should properly be 3.

We conjecture that this noncommutativity of limits is a general feature of any theory which is not asymptotically free (AF) in perturbation theory. The reasoning is simple. In a non-AF theory, perturbation theory is valid, if at all, only at low energies, $Q^2\ll\Lambda^2$, where Λ is the characteristic scale of the theory. The perturbative limit $\lambda_R\to 0$ is equivalent to sending Λ^2 to infinity. However, there is also the cutoff $M_{\text{UV}}{}^2$ to be sent to infinity. It is only to be expected that the results will depend on which mass scale is sent to infinity first. In an AF theory, on the other hand, perturbation theory is valid at high energies, $Q^2\gg\Lambda^2$. The perturbative limit $\lambda_R\to 0$ then corresponds to sending Λ^2 to zero. There is then no reason to suppose that this interferes with taking $M_{\text{UV}}{}^2$ in the opposite direction, towards infinity.

D. Stability and precariousness

The analysis above reveals that ϕ^4 in 3 + 1 dimensions does have a nontrivial form, if the bare coupling constant is negative and infinitesimal. Classical intuition and perturbation theory would both indicate that a negative-λ theory is unstable, but we have nevertheless found that the GEP is bounded below. In this section we discuss why, and in what sense, the theory is stable.

Some insight can be gained by considering IR behavior, and how it improves with increasing dimension. In 0+1 and 1+1 dimensions a negative λ_B was unthinkable: it led to $\overline{V}_G(\phi_0)=-\infty$ everywhere, due to the IR divergences associated with the $\Omega=0$ end point. In 2+1 dimensions a negative λ_B was unacceptable, but not unthinkable: it gave a finite $\overline{V}_G(\phi_0)$, but one which became unbounded below at large ϕ_0. In 3 + 1 dimensions the IR behavior has improved so much that the $\Omega=0$ end point gives a GEP which is both finite and bounded below as $\phi_0\to\infty$ (being in fact constant for $\phi_0>\phi_{0,\text{break}}$), for $\lambda_B=-1/(6I_{-1})$. Thus in 3 + 1 dimensions a "sufficiently infinitesimally small," negative λ_B becomes *just* viable.

[The above paragraph suggests an explanation for the analogy with the crater potential QM example of I. The shape of the crater potential resembles $\frac{1}{2}m^2\phi^2+\lambda\phi^4$ with negative λ, except at large ϕ, where it becomes asymptotically constant. One can regard this difference as a way of forcing the QM system to show the same good IR ($\Omega=0$) behavior as the (3 + 1)-dimensional system, thereby completing the analogy.]

Perhaps a better way of understanding the situation is to realize that, in a sense, the theory is really "infinitely metastable" rather than truly stable. To explain this fine distinction we need to consider the situation in a cutoff version of the theory, such as a lattice-regularized theory. In a lattice version of the theory the equivalent of I_{-1} is large, but finite. Terms of order $1/I_{-1}^{\text{lat}}$ can no longer be neglected, and for sufficiently large ${\phi_0}^2$ they may even come to dominate. Indeed, if we go back to (4.20) and investigate the $\Omega=0$ end point, which governs the large-ϕ_0 behavior, keeping the $1/I_{-1}^{\text{lat}}$ terms we find

$$\mathscr{V}_G(\Phi_0,\Omega=0)=\frac{(2\kappa-1)}{128\pi^2}+(A+B{\Phi_0}^2-\tfrac{1}{6}{\Phi_0}^4)\frac{1}{I_{-1}^{\text{lat}}}+\cdots\,. \tag{5.35}$$

The precise form of A and B is unimportant. What matters is the presence of the negative-${\Phi_0}^4$ term—which is, of course, just the $\lambda_B{\Phi_0}^4$ term of the classical potential. This term will dominate for ${\Phi_0}^2\gg(I_{-1}^{\text{latt}})^{1/2}$, causing the potential to be unbounded below. Thus the lattice system, which has a *finite*, negative λ_B, $=-1/(6I_{-1}^{\text{lat}})$, is indeed unstable. Intuition is satisfied.

However, as the regularization is removed, and $I_{-1}^{\text{lat}}\to\infty$, the value of ${\Phi_0}^2$ at which the instability occurs also tends to infinity. (See Fig. 11.) In the limit $I_{-1}^{\text{lat}}\to\infty$ we recover the situation in Fig. 7, where the potential *is* bounded below at all finite Φ_0's. Although the lattice system does not have a ground state, it *does* have a metastable state at $\Phi_0=0$. If the system is placed initially in this state, its decay will be hindered by the very wide potential barrier, extending out to ${\Phi_0}^2\sim(I_{-1}^{\text{lat}})^{1/2}$. (See Fig. 11.) As $I_{-1}^{\text{lat}}\to\infty$, quantum tunneling through the barrier will be more and more suppressed, and the lifetime of the metastable vacuum will tend to infinity. Thus in the continuum limit we recover our previous conclusion that the theory has a stable ("infinitely metastable") vacuum.

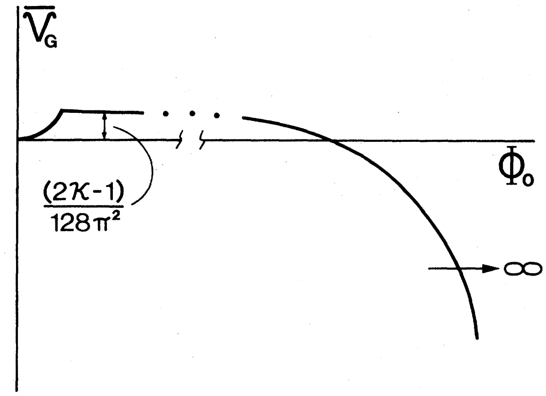

FIG. 11. A sketch of the GEP for a lattice-regularized ϕ^4 theory with $\lambda_B=-1/(6I_{-1}^{\text{lat}})$. The potential is unbounded below, so the system has no ground state. However, the instability moves out to infinity as the continuum limit is approached, leaving the continuum theory with a potential which *is* bounded below [and equal to the constant value $(2\kappa-1)/(128\pi^2)$ for all ϕ_0's above $\phi_{0,\text{break}}$].

To describe this situation we are led to introduce the following terminology: *A "precarious" theory is a theory which is unstable for any finite UV cutoff, but which becomes stable (infinitely metastable) when the cutoff is removed.* We believe that such theories may have physical interest, and should not be dismissed out of hand. We can argue this from two different viewpoints.

Our "hard-line" viewpoint is that, after all, we are studying continuum quantum field theory. Such theories have no UV cutoff. If one is temporarily introduced for technical reasons, it must be removed again before any physical properties of the theory—such as vacuum stability—are investigated. Any limit we may wish to consider, such as $\lambda_R\to 0$, or $\phi_0\to\infty$, must be taken *after* the cutoff has been taken back to infinity. Otherwise we are not studying quantum field theory, but something else.

It is a very important feature of quantum field theory that it contains field modes with arbitrarily short wavelengths. If a mutilated version of the theory with these short-wavelength modes removed is unstable, why should we care? The mutilated version is not Lorentz invariant, even.

The alternative viewpoint is to adopt the language of "effective field theories." The philosophy here is based on the recognition that we are, and always will be, ignorant of physics at arbitrarily small distances. Therefore, we ought to regard any currently successful theory as merely a low-energy "effective-theory," valid only for experiments involving momentum scales much less than some very large-scale M_{UV}. That is, we should imagine our theory as being embedded in some unknown, underlying theory, which may have a quite different high-energy/short-distance behavior from our current theory. We may say that our current theory is "truly renormalizable" if, for sufficiently large $M_{\rm UV}$, all its physical predictions are insensitive to the exact value of $M_{\rm UV}$, and to the nature of the physics above $M_{\rm UV}$. In some ways, $M_{\rm UV}$ acts as a UV cutoff on the theory—but there is more to it than that.

Let us suppose that our ϕ^4 theory arises as an "effective low-energy theory": to be specific, suppose the underlying theory has fermionic "preons" with some Lagrangian

$$\mathscr{L}=\mathscr{L}(\psi,\bar{\psi};M_{\rm UV})\ , \tag{5.36}$$

where $M_{\rm UV}$ is the characteristic scale of this preon theory. Imagine that the fermions and antifermions pair up to form very tightly bound spin-0 bound states, which can be described by an effective (pseudo-)scalar field, ϕ. At low energies the theory looks like a theory of scalar particles, and the original Lagrangian is approximately equivalent to an "effective" Lagrangian

$$\mathscr{L}_{\rm eff}=\tfrac{1}{2}\partial_\mu\phi\partial^\mu\phi+\tfrac{1}{2}m_B{}^2\phi^2+\lambda_B\phi^4+A_1\phi^6+A_2\phi^2\partial_\mu\phi\partial^\mu\phi+B_1\phi^8+B_2\phi^4\partial_\mu\phi\partial^\mu\phi+B_3(\partial_\mu\phi\partial^\mu\phi)^2+\cdots\ . \tag{5.37}$$

(We are assuming that $\mathscr{L}$ had some symmetry which implies a $\phi\to-\phi$ symmetry in $\mathscr{L}_{\rm eff}$.) On dimensional grounds one expects, up to logarithms, that $m_B{}^2\sim M_{\rm UV}{}^2$, $\lambda_B\sim 1$, $A_i\sim M_{\rm UV}{}^{-2}$, $B_i\sim M_{\rm UV}{}^{-4}$, etc., in 3 + 1 dimensions. If we neglect all terms suppressed by inverse powers of $M_{\rm UV}$, we find exactly the ϕ^4 Lagrangian we have been studying. However, it is clearly not valid to neglect the higher-dimension terms if we attempt to address questions involving either large momentum $|\mathbf{k}|\sim M_{\rm UV}$, or large field values $|\phi|\sim M_{\rm UV}$.

It is this last point which is important for us. The effective-field-theory framework is not the same as merely introducing a UV cutoff. It also introduces a cutoff on large field values. (Indeed, it is probably not meaningful to speak of very large ϕ's, $\sim M_{\rm UV}$, because the notion of an "effective field" presumably breaks down at these scales, and one has to describe the system in terms of the elementary preons, or whatever.) In the naive UV-cutoff approach there is no way to obtain "precarious ϕ^4" as the limit of a finite, well-defined cutoff theory with a bounded-below Hamiltonian. However, in the effective-field-theory framework one can envisage "precarious ϕ^4" arising as an effective low-energy theory from a well-defined, bounded-below, underlying theory. (See Fig. 12.) It would only be a metastable phase of the underlying theory, whose true vacuum would be at large ϕ_0. (The physics in the true vacuum could only be adequately described in terms of the preons.) Although metastable, the $\phi_0=0$ vacuum would be very long-lived, by our previous arguments. For sufficiently large $M_{\rm UV}$ its instability would be physically irrelevant. Thus, precarious ϕ^4 would satisfy our criterion for a truly renormalizable theory: all the physics is insensitive to $M_{\rm UV}$, for sufficiently large $M_{\rm UV}$.

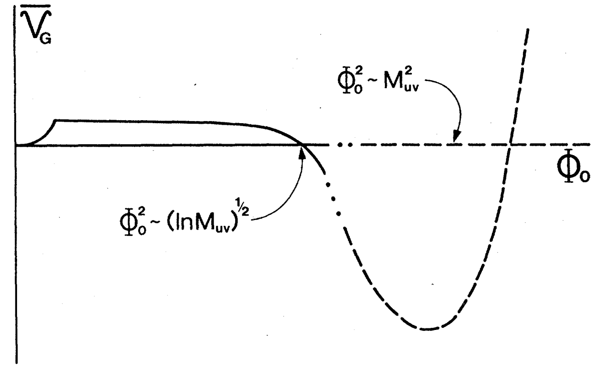

FIG. 12. "Precarious ϕ^4" in effective-field-theory language. The true vacuum of the underlying theory is, in some sense, at very large values of $\phi^2\sim M_{\rm UV}{}^2$ (where, however, the concept of an effective scalar field ϕ probably ceases to be meaningful). The $\phi_0=0$ vacuum of the precarious-ϕ^4 theory corresponds to a metastable state of the whole system. However, since its decay is hindered by a very wide barrier, extending out to $\phi_0{}^2\sim(\ln M_{\rm UV})^{1/2}$, its decay rate becomes negligible for sufficiently large $M_{\rm UV}$. All the physics in the $\phi_0=0$ vacuum becomes insensitive to $M_{\rm UV}$ and to the details of the underlying theory, as $M_{\rm UV}\to\infty$. In this sense, precarious ϕ^4 is a "truly renormalizable" theory.

We think this framework is the most satisfactory one for interpreting our results. Indeed, it supplies a simple meaning to the mysterious behavior at strong coupling ($\kappa<\frac{1}{2}$) and/or high temperatures: Beyond the phase transition there is no longer any barrier to hinder the decay to the true vacuum of the underlying theory. The new phase will have SSB, but everything else about it is literally mysterious, because it is necessarily sensitive to $M_{\rm UV}$ and to the unknown, preonic dynamics that lies beyond $M_{\rm UV}$. Hence our designation of this phase by the question mark in Fig. 9 is indeed appropriate.

VI. EXCITED STATES

A. One-particle states

As in quantum mechanics (see I), we can investigate excited states of the theory by constructing excited-state generalizations of the GEP. We do this by computing the expectation value of the Hamiltonian $H=\int d^\nu x\,\mathscr{H}$ between n-quantum states built on the trial vacuum state $|0\rangle_{\Omega,\phi_0}$. (Note that we shall be working with the Hamiltonian H rather than the Hamiltonian density $\mathscr{H}$.) The analysis in this section follows Schiff[2] and Barnes and Ghandour.[4]

For the one-particle state we need to evaluate

$$E_1(\phi_0,\mathbf{p})\equiv\frac{{}_{\phi_0,\Omega}\langle\mathbf{p}\,|\,H\,|\,\mathbf{p}\rangle_{\Omega,\phi_0}}{{}_{\phi_0,\Omega}\langle\mathbf{p}\,|\,\mathbf{p}\rangle_{\Omega,\phi_0}}\ , \tag{6.1}$$

where

$$|\mathbf{p}\rangle_{\Omega,\phi_0} \equiv a_\Omega^\dagger(\mathbf{p})\,|\,0\rangle_{\Omega,\phi_0} . \tag{6.2}$$

The calculation proceeds just as in Eq. (2.11), except for the extra $a_\Omega(\mathbf{p})\cdots a_\Omega^\dagger(\mathbf{p})$ operators in the matrix elements. From (2.9) we have the normalization

$$_\Omega\langle \mathbf{p}\,|\,\mathbf{p}\rangle_\Omega = 2\omega_p(\Omega)(2\pi)^\nu \delta^{(\nu)}(0) = 2\omega_p(\Omega)\left[\int d^\nu x\right] \tag{6.3}$$

[see any textbook[14] for the justification for identifying $(2\pi)^\nu\delta^{(\nu)}(0)$ with the spatial volume $\int d^\nu x$], and so we obtain

$$E_1(\phi_0,\mathbf{p}) = \omega_p(\Omega) + [m_B{}^2 - \Omega^2 + 12\lambda_B(I_0+\phi_0{}^2)]/[2\omega_p(\Omega)] + \left[\int d^\nu x\right][I_1 + \tfrac{1}{2}(m_B{}^2-\Omega^2)I_0 + 3\lambda_B I_0{}^2 + 6\lambda_B I_0\phi_0{}^2 + \tfrac{1}{2}m_B{}^2\phi_0{}^2 + \lambda_B\phi_0{}^4] . \tag{6.4}$$

The last term is just the vacuum energy

$$E_0(\phi_0) \equiv {}_{\phi_0,\Omega}\langle 0\,|\,H\,|\,0\rangle_{\Omega,\phi_0} = \int d^\nu x\; V_G(\phi_0,\Omega) .$$

In QM, the optimum Ω for the first excited state was in general different from the optimum Ω for the GEP itself. In field theory proper ($\nu>0$) this is no longer true, because any change in Ω would lead to a volume-divergent term in E_1. Thus it is obvious that (6.4) is minimized by the same Ω value as for the GEP at the same ϕ_0. Since that Ω satisfies (2.15) we observe that the second term in (6.4) will vanish. Therefore, we find that the energy of a one-particle state, over and above the vacuum energy, is just

$$E_1(\phi_0,\mathbf{p}) - E_0(\phi_0) = \omega_p(\bar\Omega) \equiv (\mathbf{p}^2+\bar\Omega^2)^{1/2} . \tag{6.5}$$

Thus, for any candidate vacuum (i.e., minimum of the GEP), we have the identification

$$\text{particle mass} = \bar\Omega . \tag{6.6}$$

In particular, for the candidate vacuum at $\phi_0=0$ we see that the particle mass is m_R, which justifies, *a posteriori*, calling this parameter the "renormalized mass."

Note that for SSB vacua at $\phi_0{}^2=c^2,\neq 0$, the mass is no longer identical to the second derivative of the GEP, because the last term in (3.19) now contributes. However, from (3.14), $\bar\Omega^2\,|_c = 8\lambda_B c^2$.

B. Two-particle states

In a similar way one can look for bound states by considering a trial state made from two creation operators acting on the trial vacuum. The simplest ansatz for an s-wave state is[2,4]

$$|\,2\rangle_{\Omega,\phi_0} = \int (dp)_\Omega \sigma(\mathbf{p}) a_\Omega^\dagger(\mathbf{p}) a_\Omega^\dagger(-\mathbf{p})\,|\,0\rangle_{\Omega,\phi_0} , \tag{6.7}$$

where, for simplicity, we have chosen to work in the c.m. frame. The function $\sigma(\mathbf{p})$ is the Fourier transform of the spatial wave function of the bound state, and is to be determined variationally by minimizing the energy.[2,4] Proceeding as in the preceding section we obtain

$$m_2 \equiv E_2(\phi_0) - E_0(\phi_0) = \frac{{}_{\phi_0,\Omega}\langle 2\,|\,H\,|\,2\rangle_{\Omega,\phi_0}}{{}_{\phi_0,\Omega}\langle 2\,|\,2\rangle_{\Omega,\phi_0}} - {}_{\phi_0,\Omega}\langle 0\,|\,H\,|\,0\rangle_{\Omega,\phi_0} = \frac{\int (dp)\sigma^2(\mathbf{p})4\omega_p{}^2 + 12\lambda_B\left[\int (dp)\sigma(\mathbf{p})\right]^2}{\int (dp)2\omega_p\sigma^2(\mathbf{p})} , \tag{6.8}$$

where terms in the numerator which vanish by virtue of the $\bar\Omega$ equation have been eliminated. The two terms in the above expression can be interpreted as the kinetic energy of the two constituent particles, and their binding energy, respectively. To minimize the result with respect to $\sigma(\mathbf{k})$ we perform the functional differentiation, obtaining[2,4]

$$\sigma(\mathbf{k})\omega_k(2\omega_k - m_2) + 6\lambda_B \int (dp)\sigma(\mathbf{p}) = 0 . \tag{6.9}$$

As the last term is independent of $\mathbf{k}$, the solution has the form

$$\sigma(\mathbf{k}) = \frac{A}{\omega_k(2\omega_k - m_2)} , \tag{6.10}$$

where A is some normalization constant. Substituting back in (6.8), or more conveniently (6.9), gives the equation for m_2 in terms of λ_B and the single-particle mass, $\bar\Omega$. (It is implicit here that ϕ_0 coincides with a minimum of the GEP, but note that the results have the same form in an SSB vacuum as in the $\phi_0=0$ vacuum.)

In 1 + 1 and 2 + 1 dimensions the integral

$$\int (dp)\sigma(\mathbf{p}) = \int (dp)\frac{A}{\omega_p(2\omega_p - m_2)} \tag{6.11}$$

is UV convergent, and Eq. (6.9) is manifestly finite. However, it is immediately clear that there are no solutions with $m_2 < 2\bar\Omega$, which would correspond to bound states, because both terms in (6.9) are then positive.

However, for $m_2 > 2\bar\Omega$ the equation contains useful information about particle scattering.[2] This can be found by using a well-known trick in QM scattering theory.[28] One imagines the system enclosed in a spherical box and considers the ratio of δE_n, the displacement of the nth energy level due to the interaction, to ΔE_n, the spacing $E_{n+1}^0 - E_n^0$ of the energy levels of the free-particle system.

Both δE_n and ΔE_n tend to zero as the walls of the box tend to infinity, but their ratio stays finite and is related to the phase shift δ,

$$\delta E_n/\Delta E_n \rightarrow -\delta/\pi\ . \tag{6.12}$$

Schiff shows that this leads to a prescription for dealing with the pole in the integral (6.11) when $m_2 > 2\overline{\Omega}$: One replaces the integral by its principal part plus $\pi \cot\delta$ times its residue. One then has an equation giving the phase shift as a function of the c.m. energy m_2. From the phase shift one can obtain the scattering amplitude, and hence the cross section. We shall not attempt to carry through this program in the present paper.

In 3 + 1 dimensions, however, things are more interesting and problematical. The integral in (6.11) is now logarithmically divergent, and although this cancels against the logarithmic divergence in $\lambda_B = -1/(6I_{-1})$, there is a recalcitrant factor of 2 which prevents Eq. (6.9) from having any solutions. In other words, the UV divergences do not cancel in Eq. (6.8). We have checked carefully that there is no factor of 2 error in Eq. (6.8), which agrees with both Schiff[2] and Barnes and Ghandour.[4]

Of course, if λ_B were twice as big, i.e., $-1/(3I_{-1})$, then Eq. (6.8) would yield a finite, nontrivial bound-state equation. Schiff[2] followed this course, but at the end of the paper he admits that this choice of λ_B leads to an unstable theory [as is clear from Eqs. (4.25) and (4.27) of our own analysis]. This makes for a very lame conclusion to his pioneering paper. Barnes and Ghandour[4] also come up against this factor-of-2 puzzle. Apparently, it discouraged them from completing their analysis of the GEP in ϕ^4 theory, and their paper moves on immediately to fermion-scalar theories.

Our provisional attitude to the problem is as follows. The GEP analysis clearly rules out $\lambda_B = -1/(3I_{-1})$, and definitely indicates $\lambda_B = -1/(6I_{-1})$ as the only viable, nontrivial case. We must therefore concede that Eq. (6.9) has no solution: The expression (6.8) is minimized instead by the "end point" $\sigma(\mathbf{p}) \propto \delta(\mathbf{p})$, corresponding to two free particles. There is no bound state—because although λ_B is negative, it is not negative enough to produce a strong enough attraction to give binding. More seriously, we are also having to concede that there is no scattering, since, by the arguments mentioned above, the phase shift δ is zero, too. However, we do not take this as meaning that the theory is a free field theory; that would be inconsistent with the nontrivial form of the GEP we found earlier. Instead, we conclude that two-particle scattering is weak, and vanishes in the approximation we are using. However, we expect that a nonzero result will be obtained in the next order of the Gaussian approximation. To settle the question one needs a higher-order calculation along the lines sketched in Sec. V of I, but this we must defer to future work.

VII. DISCUSSION

While the results in 1 + 1 and 2 + 1 dimensions are relatively straightforward, the situation in 3 + 1 dimensions is quite subtle. We summarize our conclusions here. (i) The obvious choice of a finite, positive λ_B is not viable. It leads to a GEP of the form $\frac{1}{2}m_R{}^2\phi_0{}^2 - 2\lambda_B\phi_0{}^4$, which is unbounded below. (ii) A positive λ_B which vanishes like $1/I_{-1}$ leads only to a free field theory. (iii) The only nontrivial form of ϕ^4 arises from the case $\lambda_B = -1/(6I_{-1})$: The GEP is bounded below—and bounded above. It indicates a phase with massive particles, interacting through an attractive force. This unbroken-symmetry phase is stable below a critical coupling $(-\lambda_R) < 8\pi^2$, and below a critical temperature.

These conclusions are markedly different from other authors who have used a similar method of analysis.[9,29] The reader must appreciate that the conclusions depend on whether one is interested in ϕ^4 or in "cutoff ϕ^4" (by which we include lattice ϕ^4, and any other UV-regulated form of ϕ^4). The point is that *the stability properties of ϕ^4 and of cutoff ϕ^4 are the opposite of one another.* One can see this from Figs. 4 and 11. In cutoff ϕ^4, positive λ_B gives a bounded potential with a pair of minima at very large ϕ_0, corresponding to a (cutoff-sensitive) SSB phase (Fig. 4), while negative λ_B leads to an unbounded potential (Fig. 11). However, in the *absence* of a UV cutoff the situation is the reverse: positive λ_B gives an unbounded potential, $\frac{1}{2}m_R{}^2\phi_0{}^2 - 2\lambda_B\phi_0{}^4$ (Sec. IV D), while a negative λ_B of the form $-1/(6I_{-1})$ leads to a stable, nontrivial theory, with unbroken symmetry (Sec. V A).

Our concern here has been to study ϕ^4, which, being a quantum field theory, contains modes of arbitrarily high momentum. It has no cutoff. A modified form of the theory with a finite UV cutoff is not Lorentz invariant; is not a quantum field theory; and, in our view, is not interesting. In this sense we believe the conclusions we stated above are the correct ones. If, for technical reasons, a UV cutoff is temporarily introduced, explicitly or implicitly, it should be taken back to infinity *before* any conclusions are drawn about the physics of the theory. Physicists seldom worry about interchanging the order of two limits, but here it is crucial. We saw, in Sec. V C, that our results differ from perturbation theory because the latter takes the weak coupling, $\lambda_R \rightarrow 0$, limit *before* letting the UV cutoff tend to infinity. Similarly, our conclusions differ from other authors because they have investigated stability in the $\phi_0 \rightarrow \infty$ limit *before* taking the UV cutoff back to infinity.

We have also discussed our results in the language of "effective field theories" (Sec. V D), making the point that in this framework the UV scale $M_{\rm UV}$ acts more subtly than a simple momentum cutoff. In this language one can give a physical meaning to both the positive-λ_B, SSB phase of cutoff ϕ^4 (Ref. 29), and the $\lambda_B = -1/(6I_{-1})$ version of ϕ^4. The difference is that the former is intrinsically sensitive to the UV scale $M_{\rm UV}$, and as $M_{\rm UV} \rightarrow \infty$ the physics goes with it. Particle masses, and other physical properties, will depend on $M_{\rm UV}$, and hence are sensitive to the detailed dynamics of the underlying theory. The $\lambda_B = -1/(6I_{-1})$ theory, on the other hand, is a truly renormalizable theory, in that, for sufficiently large $M_{\rm UV}$, all the physics become insensitive to $M_{\rm UV}$. In particular, the decay rate of the vacuum—for this would be a metastable state of the underlying theory—vanishes as $M_{\rm UV} \rightarrow \infty$.

Our interest here has been purely theoretical. For readers interested in ϕ^4 only in relation to the Higgs sector of gauge theories, we have little to say. The $\lambda_B=-1/(6I_{-1})$ form of ϕ^4 clearly has nothing to do with the Higgs mechanism, since it has no SSB. As we noted in the last paragraph, the positive-λ_B form of cutoff ϕ^4 does exhibit SSB (Ref 29) but the properties of this theory depend intrinsically upon the context within which it is embedded. The moral is that the Higgs mechanism is not a one-way street. One does not have a quasiautonomous ϕ^4 sector with SSB, which then feeds the SSB through to the gauge sector. Instead, the nature of the symmetry-broken vacuum, the value of $\langle\phi\rangle$, etc., depend crucially on the way the scalar sector is embedded in the full theory. Several authors have already pursued these considerations, showing that they imply an upper limit on the Higgs-boson mass.[30]

We believe the GEP approach sheds a lot of light on the mystery of $(\phi^4)_{3+1}$. The results, for positive λ_B, are consistent with, and provide a simple interpretation of, the arguments for "triviality" of Refs. 11–13. However, we have also found that there exists a nontrivial, renormalizable, form of ϕ^4 hiding at negative values of λ_B. The nontrivial results found in perturbation theory are related to this theory—although not straightforwardly, because of the interchange-of-limits problem (Sec. V C). The nontrivial ϕ^4 is *stable*—but only in the *absence* of a UV cutoff, which perhaps explains why it has been overlooked in the past.

We hope that these new insights into a very old model will serve to illustrate our point that the GEP is a very simple and powerful way to investigate field theories nonperturbatively.

Note added. Since the completion of this work several additional references have come to our attention. (i) Bollini and Giambiagi[31] have recently shown, using dimensional regularization, that for any $\lambda_B\neq 0$ the GEP is unbounded below for ϕ^4 theories in more than four dimensions, in agreement with the known triviality of such theories.[12] (ii) The possibility of a negative λ in $(\lambda\phi^4)_{3+1}$ was discussed very carefully from the standpoint of perturbation theory in Ref. 32. It makes interesting reading in conjunction with our Sec. V C. (iii) Nonperturbative approaches different from, but clearly related to, the GEP were introduced in Refs. 33 and 34. Reference 33 is particularly interesting as it considers $(\phi^4)_{3+1}$ with a negative λ_R. The formalism is based on treating $\chi\equiv\phi^2$, thought of as a bound state, as a separate field. The authors found that a bound state exists for $\lambda_R<-8\pi^2$ (which is, however, the region where we find the normal vacuum is not stable; see Figs. 8 and 9). For $-8\pi^2<\lambda_R<0$ the authors found an apparently viable theory with no bound states, in agreement with our findings. The absence of ghosts (tachyons) was noted, but the puzzle of the theory's stability was not elucidated, as we have tried to do here. (iv) A quite different approach to the $(\phi^4)_{3+1}$ problem[35] via Brownian motion has also led to the conclusion that a negative-λ theory is viable. (v) Kurt Symanzik was apparently the first to realize that a $(\phi^4)_{3+1}$ theory with a negative, but infinitesimal, λ_B could be stable.[36] Our results can be seen as a vindication of his original arguments.

ACKNOWLEDGMENTS

This research was supported in part by the University of Wisconsin Research Committee with funds granted by the Wisconsin Alumni Research Foundation and in part by the U.S. Department of Energy under Contracts Nos. DE-AC02-76ER0081 (Wisconsin) and DE-AS05-76ER05096 (Rice University)

APPENDIX: SPECIAL RESULTS IN DIMENSIONAL REGULARIZATION

Dimensional regularization is based on the analytic continuation of Eq. (3.5) to noninteger values of ν. At the end of the calculation, the limit $\nu\to 3$ (or whatever) is taken. In other forms of regularization I_1 is incomparably bigger than I_0, which in turn is incomparably bigger than I_{-1}. However, in dimensional regularization the divergences in each case correspond to a simple pole in $1/(3-\nu)$. This property allows some strange cancellations to occur, giving rise to some intriguing results.

We concentrate on the (3 + 1)-dimensional theory with $\lambda_B=-1/[6I_{-1}(\Lambda)]$. Evaluating the bare mass from Eq. (3.18) we observe that the poles in $I_0(m_R)$ and $I_{-1}(\Lambda)$ cancel to give

$$m_B{}^2=m_R{}^2-\frac{2}{(\nu-1)}m_R{}^2\left[\frac{\Lambda}{m_R}\right]^{3-\nu} . \tag{A1}$$

Putting $\lambda=3-\epsilon$ and taking the limit $\epsilon\to 0$ we see that the two terms cancel, leaving

$$m_B{}^2=\tfrac{1}{2}\epsilon m_R{}^2\ln(m_R{}^2/\Lambda^2)+O(\epsilon^2) . \tag{A2}$$

Thus in dimensional regularization *the bare mass is zero* (or, rather, is infinitesimal). This result has been obtained independently by Bollini and Giambiagi[31] and we are grateful to them for discussions on these matters.

Another intriguing exercise is to compute the vacuum energy density

$$D\equiv I_1(m_R)-3\lambda_B[I_0(m_R)]^2 \tag{A3}$$

with $\lambda_B=-1[6I_{-1}(\Lambda)]$. Again there are two cancellations, one multiplicative between I_{-1} and one of the I_0's, and one subtractive between the first and second terms. The result is therefore finite:

$$D=\frac{m_R{}^4}{128\pi^2}[1-2\ln(m_R{}^2/\Lambda^2)] . \tag{A4}$$

Remembering that $\kappa\equiv\ln(m_R{}^2/\Lambda^2)$, we see that this is *negative* in the weak coupling phase (in spite of the fact that this a *boson* theory). In fact, D is just the negative of Eq. (5.11), so that the potential is "chopped off" by the $\Omega=0$ end point (see Fig. 7) precisely when it has risen to zero energy. [Actually, the easy way to derive (A4) is to note that for $\Omega=0$ dimensional regularization defines $I_1(0),I_0(0)$ to be zero. Since both bare parameters are $O(\epsilon)$, one sees immediately from the original expression for V_G Eq. (2.13), that $V_G(\phi_0,\Omega=0)=0$ as $\epsilon\to 0$.]

While the absolute energy density has no physical significance in ϕ^4 alone, it would have significance if the theory were coupled to gravity. Thus the above result may be telling us something about the cosmological constant generated by a scalar theory.

*Present address.
[1]P. M. Stevenson, Phys. Rev D **30**, 1712 (1984).
[2]L. I. Schiff, Phys. Rev. **130**, 458 (1963).
[3]G. Rosen, Phys. Rev. **173**, 1632 (1968); in *Path Integrals and Their Applications in Quantum, Statistical, and Solid-State Physics,* edited by G. J. Papadopoulos and J. T. Devreese (Plenum, New York, 1978).
[4]T. Barnes and G. I. Ghandour, Phys. Rev. D **22**, 924 (1980).
[5]J. Kuti (unpublished), as discussed in J. M. Cornwall, R. Jackiw, and E. Tomboulis, Phys. Rev. D **10**, 2428 (1974).
[6]S. J. Chang, Phys. Rev. D **12**, 1071 (1975); Phys. Rep. **23C**, 301 (1975).
[7]S. J. Chang, Phys. Rev. D **13**, 2778 (1976).
[8]M. Weinstein, S. Drell, and S. Yankielowicz, Phys. Rev. D **14**, 487 (1976); K. Huang and D. R. Stump, *ibid.* **14**, 223 (1976).
[9]W. A. Bardeen and M. Moshe, Phys. Rev. D **28**, 1372 (1983).
[10]P. M. Stevenson, Z. Phys. C **24**, 87 (1984).
[11]K. G. Wilson, Phys. Rev. D **6**, 419 (1972) (Ref. 8); K. G. Wilson and J. Kogut, Phys. Rep. **12C**, 75 (1974); G. A. Baker and J. M. Kincaid, Phys. Rev. Lett. **42**, 1431 (1979).
[12]M. Aizenmann, Phys. Rev. Lett. **47**, 1 (1981); Commun. Math. Phys. **86**, 1 (1982); J. Fröhlich, Nucl. Phys. **B200**, [FS4], 281 (1982); C. Aragão de Carvalho, C. S. Carracciolo, and J. Fröhlich, *ibid.* **B215** [FS7], 209 (1983), and references therein.
[13]B. Freedman, P. Smolensky, and D. Weingarten, Phys. Lett. **113B**, 481 (1982); I. G. Halliday, Rep. Prog. Phys. **47**, 987 (1984).
[14]See, for example, P. Roman, *Introduction to Quantum Field Theory* (Wiley, New York, 1969).
[15]An alternative starting point is to treat ω_k (as a function of **k**) as the variational parameter. (See Refs. 2, 4, and 9.) This leads to exactly the same results as in the text because, when the energy is functionally minimized with respect to ω_k, one finds that ω_k has the form $(\mathbf{k}^2+\text{constant})^{1/2}$. This approach provides a variational justification for our assumption that Ω is the same for all **k**.
[16]E. C. G. Stueckelberg and A. Peterman, Helv. Phys. Acta **26**, 449 (1953); M. Gell-Mann and F. Low, Phys. Rev. **95**, 1300 (1954); N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience, New York, 1959).
[17]S. Coleman and E. Weinberg, Phys. Rev. D **7**, 1888 (1973).
[18]P. M. Stevenson, Phys. Rev. D **23**, 2916 (1981); Nucl. Phys. **B203**, 472 (1982); in *Radiative Corrections in* $SU(2)_L\times U(1)$, edited by B. W. Lynn and J. F. Wheater (World Scientific, Singapore, 1984).
[19]C. G. Bollini and J. J. Giambiagi, Nuovo Cimento **128**, 20 (1972); G.'t Hooft and M. Veltman, Nucl. Phys. **B44**, 189 (1972).
[20]See almost any QM textbook. Two particularly interesting discussions are R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, Mass., 1963), Vol. III; S. Coleman, in The *Whys of Subnuclear Physics,* proceedings of the International School of Subnuclear Physics, Erice, 1977, edited by A. Zichichi (Plenum, New York, 1979).
[21]B. Simon and R. G. Griffiths, Commun. Math. Phys. **33**, 145 (1973); B. Simon, *The* $P\,(\phi)_2$ *Euclidean (Quantum) Field Theory* (Princeton University Press, Princeton, N.J., 1974).
[22]We are unsure of the status of this question in the literature. Some possibly relevant references are J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973); J. M. Kosterlitz, *ibid.* **7**, 1046 (1972); J. Tobochnik and G. V. Chester, Phys. Rev. B **20**, 3761 (1979); J. Glimm and A. Jaffe, *Quantum Physics: A Functional Integral Point of View* (Springer, New York, 1981).
[23]The continuum theory does have a metastable state at $\phi_0=0$ according to the GEP. However, unlike the theory we discuss in Sec. V, the lifetime of this state does not become arbitrarily long as $M_{\rm UV}\rightarrow\infty$, because here the barrier is governed solely by the parameters $m_R{}^2$ and λ_B. See Fig. 4.
[24]P. M. Stevenson, Ann. Phys. (N.Y.) **132**, 383 (1981).
[25]Although the analyses are quite similar, the conclusions are rather different. The main difference appears to lie in our different attitude to the UV cutoff, which we discuss at length in Sec. V D.
[26]S. Coleman, R. Jackiw, and H. D. Politzer, Phys. Rev. D **10**, 2491 (1974).
[27]M. Kobayashi and T. Kugo, Prog. Theor. Phys. **54**, 1537 (1975); R. W. Haymaker, Phys. Rev. D **13**, 968 (1976); L. F. Abbott, J. S. Kang, and H. J. Schnitzer, *ibid.* **13**, 2212 (1976).
[28]J. Schwinger, Phys. Rev. **94**, 1362 (1954); B. S. DeWitt, *ibid.* **103**, 1565 (1956); M. Baker, Ann. Phys. (N.Y.) **4**, 271 (1958); K. Gottfried, *Quantum Mechanics* (Benjamin, New York, 1966), Sec. 49.
[29]M. Consoli and A. Ciancitto, Nucl. Phys. **B254**, 653 (1985).
[30]R. Dashen and H. Neuberger, Phys. Rev. Lett. **50**, 1897 (1983); M. A. B. Bég, C. Panagiotakopoulos, and A. Sirlin, *ibid.* **52**, 883 (1984); D. J. E. Callaway, Nucl. Phys. **B233**, 189 (1984).
[31]C. G. Bollini and J. J. Giambiagi, Report No. CBPF-NF-046/84, 1984 (unpublished).
[32]J. Iliopoulos, C. Itzykson, and A. Martin, Rev. Mod. Phys. **47**, 165 (1975), Sec. V and Appendix C.
[33]F. Cooper, G. S. Guralnik, and S. H. Kasdan, Phys. Rev. D **14**, 1607 (1976).
[34]C. Bender, F. Cooper, and G. S. Guralnik, Ann. Phys. (N.Y.) **109**, 165 (1977); R. W. Haymaker and F. Cooper, Phys. Rev. D **19**, 562 (1979).
[35]S. Albeverio *et al.*, Phys. Lett. **104B**, 396 (1984).
[36]K. Symanzik, Lett. Nuovo Cimento **6**, 77 (1973); see also G.'t Hooft, Nucl. Phys. **B254**, 11 (1985).