ings of the Johns Hopkins Workshop on Current Problems in High Energy Particle Theory Physics, 1974 (unpublished).

[27]O. W. Greenberg and A. M. L. Messiah, Phys. Rev. 138, B1155 (1965); J. Hartle and J. R. Taylor, Phys. Rev. 178, 2043 (1968); T. Palev, CERN Report No. TH-1749, 1973 (unpublished). References to earlier works can be traced from these papers.

[28]Y. Nambu and M. Y. Han, Phys. Rev. D (to be published).

[29]See, for example, N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Wiley, New York, 1959).

[30]O. Klein, J. Phys. Rad. (7), 9, 1 (1938).

[31]P. A. M. Dirac, Rev. Mod. Phys. 26, 392 (1949).

[32]See, for example, O. W. Greenberg, in *Proceedings of the Conference on Mathematical Theory of Elementary Particles*, edited by R. Goodman and I. Segal (MIT Press, Cambridge, Mass., 1965).

[33]M. Kaku and K. Kikkawa, CCNY report, 1973 (unpublished); P. Ramond, in proceedings of the Johns Hopkins Workshop on Current Problems in High Energy Particle Theory Physics, 1974 (unpublished); M. Kaku and K. Kikkawa, Phys. Rev. D (to be published).

[34]K. Bardakci and M. B. Halpern, Phys. Rev. 176, 1686 (1968).

[35]P. Ramond, Phys. Rev. D 3, 2493 (1971); A. Neveu and J. H. Schwarz, Nucl. Phys. B31, 86 (1971); A. Neveu, J. H. Schwarz, and C. B. Thorn, Phys. Lett. 35B, 529 (1971); Y. Aharanov, A. Casher, and L. Susskind, *ibid.* 35B, 512 (1971); J. L. Gervais and B. Sakita, Nucl. Phys. B34, 633 (1971).

[36]Y. Iwasaki and K. Kikkawa, Phys. Rev. D 8, 440 (1973); S. Mandelstam, Phys. Lett. 46B, 447 (1973).

[37]P. Goddard, C. Rebbi, and C. B. Thorn, Nuovo Cimento 12A, 425 (1972).

PHYSICAL REVIEW D VOLUME 9, NUMBER 12 15 JUNE 1974

Gauge and global symmetries at high temperature*

Steven Weinberg
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 19 February 1974)

It is shown how finite-temperature effects in a renormalizable quantum field theory can restore a symmetry which is broken at zero temperature. In general, for both gauge symmetries and ordinary symmetries, such effects occur only through a temperature-dependent change in the effective bare mass of the scalar bosons. The change in the boson bare mass is calculated for general field theories, and the results are used to derive the critical temperatures for a few special cases, including gauge and nongauge theories. In one case, it is found that a symmetry which is unbroken at low temperature can be broken by raising the temperature above a critical value. An appendix presents a general operator formalism for dealing with higher-order effects, and it is observed that the one-loop diagrams of field theory simply represent the contribution of zero-point energies to the free energy density. The cosmological implications of this work are briefly discussed.

I. INTRODUCTION

The idea of broken symmetry was originally brought into elementary-particle physics on the basis of experience with many-body systems.[1] Just as a piece of iron, although described by a rotationally invariant Hamiltonian, may spontaneously develop a magnetic moment pointing in any given direction, so also a quantum field theory may imply physical states and S matrix elements which do not exhibit the symmetries of the Lagrangian.

It is natural then to ask whether the broken symmetries of elementary-particle physics would be restored by heating the system to a sufficiently high temperature, in the same way as the rotational invariance of a ferromagnet is restored by raising its temperature. A recent paper by Kirzhnits and Linde[2] suggests that this is indeed the case. However, although their title refers to a gauge theory, their analysis deals only with ordinary theories with broken global symmetries. Also, they estimate but do not actually calculate the critical temperature at which a broken symmetry is restored.

The purpose of this article is to extend the analysis of Kirzhnits and Linde to gauge theories,[3] and to show how to calculate the critical temperature for general renormalizable field theories, with either gauge or global symmetries. Our results completely confirm the more qualitative conclusions of Kirzhnits and Linde.[2]

The diagrammatic formalism[4] used here is described in Sec. II. Any finite-temperature Green's function is given by a sum of Feynman diagrams, just as in field theory, except that energy integrals are replaced with sums over a discrete imaginary energy. The justification of the use of this formalism in gauge theories is discussed briefly.

Section III lays the general foundation for calculations of the critical temperature. Our work here is based on the observation that a symmetry which is broken or unbroken in the lowest order of perturbation theory will remain broken or unbroken to all orders, unless there is some circumstance which invalidates the perturbation expansion. We assume that the theory is characterized by a small dimensionless coupling constant $e \ll 1$, so it might be thought that the symmetries of the theory are simply determined by the minima of the scalar-field polynomial $P(\phi)$ in the Lagrangian, and therefore could not be affected by an increase in temperature. However, at very high temperatures, powers of the temperature θ can compensate for powers of e, leading to a breakdown of the perturbation expansion. The leading effect of this sort arises from the $e^2\theta^2$ terms which accompany single-loop quadratic divergences. Since the theory is renormalizable, and we work in a renormalizable gauge, all such terms can be absorbed into a redefinition of the mass terms in $P(\phi)$. Once this is done, the validity of the perturbation expansion is restored. In a general renormalizable theory involving scalar fields ϕ_i, the change in the effective polynomial is calculated here as

$$\Delta P(\phi) = \tfrac{1}{48}\theta^2\{f_{ijkk} + 6(\theta_\alpha\theta_\alpha)_{ij} + \mathrm{Tr}[\gamma_4\Gamma_i\gamma_4\Gamma_j]\}\phi_i\phi_j \ , \tag{1.1}$$

where f_{ijkl} is the coefficient (of order e^2) of the quartic term in $P(\phi)$, θ_α is the matrix (of order e) representing the αth generator of the gauge group on the scalar fields, and Γ_i is the Yukawa coupling matrix (of order e) for the scalar-spinor interaction. (This notation is the same as used in Refs. 7 and 25, and is reviewed here in Sec. II.) In general a critical temperature is reached at values of θ for which the symmetries of the minimum of $P(\phi)+\Delta P(\phi)$ are gained or lost; usually this occurs when one of the eigenvalues of the bare mass matrix in $P+\Delta P$ vanishes.

This general formalism is used in Sec. IV to calculate critical temperatures in three special cases. The first case is a scalar field theory with an $O(n)$ global symmetry group; it is found that the spontaneous symmetry breakdown encountered at low temperature disappears at a finite temperature θ_c, given by

$$\theta_c = \left(\frac{6}{n+2}\right)^{1/2}\left(\frac{M(0)}{e}\right) , \tag{1.2}$$

where the quadrilinear self-coupling is taken as $\frac{1}{4}e^2(\phi_i\phi_i)^2$, and $M(0)$ is the mass of the single non-Goldstone boson at zero temperature. The second case is a gauge theory with a local $O(n)$ symmetry group; it is found that there is again a critical temperature θ_c above which the gauge symmetry is restored, now given by

$$\theta_c = [\tfrac{1}{6}(n+2)e^2 + \tfrac{1}{2}(n-1)e'^2]^{-1/2}M(0) , \tag{1.3}$$

where e' is the gauge coupling constant and e^2 is again the quadrilinear coupling constant. [In gauge theories of the weak and electromagnetic interactions,[5] we typically have $M(0)/e$ of the order of $G_F^{-1/2}$, so these examples indicate that θ_c will be of the order of 300 GeV.] The third case is a scalar field theory with a global $O(n)\times O(n)$ symmetry; it is found that for certain ranges of the parameters in the theory it is possible for one of the $O(n)$'s to be broken at low temperatures and restored at high temperatures, while the other $O(n)$ is unbroken at low temperatures and broken at high temperatures. This has the appearance of a violation of the second law of thermodynamics, but this is not the case: In fact, certain crystals, such as the ferroelectric known as Rochelle or Seignette salt, also have a smaller invariance group above some critical temperature than below it.[6]

Section V compares these results with those that would be found by calculation in a non-renormalizable "unitarity" gauge.[7] In general, in this gauge the introduction of an effective polynomial would not eliminate the θ^2 terms which accompany quadratic divergences, and therefore would not restore perturbation theory at high temperature. In certain simple cases the θ^2 terms which accompany tadpole graphs are eliminated by introduction of an effective polynomial, but the critical temperatures deduced in this way disagree with those calculated in renormalizable gauges, and are argued to be physically irrelevant.

The problem of calculating the critical temperature more accurately and of determining the nature of the phase transition is discussed briefly in Sec. VI. The difficulty here is that as we approach the critical temperature we encounter infrared divergences which invalidate perturbation theory, even after introducing an effective polynomial. It is estimated that the true critical temperature differs from the critical temperature calculated in Sec. IV by an amount at most of order $e^2\theta_c$.

Section VII deals with the question of the observability of phase transitions in gauge theories. It is concluded that spontaneous symmetry breaking can be detected by measurement of Green's functions for gauge-invariant operators carrying zero energies and moderate momenta. Also, although the pressure and energy and entropy densities are continuous at the critical temperature, the specific heat per unit volume has a discontinuity of order $e^2\theta_c{}^3$.

An appendix deals with the problem of defining and calculating a "potential" whose minimum will be at the precise thermodynamic mean value of the scalar field. The potential is defined, using operator rather than diagram methods, as the free energy per unit volume, and it is observed that the corresponding potential calculated earlier for field theories at zero temperature[8] simply represents the contribution of zero-point energies to the free energy. The calculations go through smoothly for scalar-field theories, with the same results as found in Sec. III and IV. However, for gauge theories this operator formalism requires canonical quantization in the unitarity gauge, and in consequence divergences appear in the potential which cannot be eliminated by renormalization of the scalar-field polynomial. Suggestions are offered for further progress along these lines.

This paper is mainly concerned with the study of the phase transition itself, but the existence of this phase transition has wider implications. These are not discussed in the body of this paper, but a word about them may be in order here.

One implication is philosophical. It has been suggested[9] that all the complicated properties of a theory that are usually derived from an assumed broken gauge symmetry may also be derived from the requirements either of perturbative unitarity or of renormalizability. If this is so, then perhaps the gauge symmetry is in some sense a fiction, not representing any truly fundamental invariance principle. It is not clear to me whether this is a question of words or of substance. However, if a gauge symmetry becomes unbroken for sufficiently high temperature, then it is difficult to doubt its reality.

Another implication is cosmological. In "big-bang" cosmologies the critical temperature was presumably reached at some time in the past (unless the richness of hadron states imposes some upper limit on the temperature).[10] In earlier epochs the weak interactions would have produced long-range forces similar to Coulomb forces, with the difference that while the universe appears to be electrically neutral, it may not be neutral with respect to the conserved quantities to which the intermediate vector bosons couple. Such long-range forces would have had profound effects on the evolution of the universe; in particular, as noted by Kirzhnits and Linde,[2] the universe could not have been isotropic and homogeneous if permeated by these lines of force. Long-range vector fields would also play an important role in determining the nature of the initial singularity[11] (if any). Finally, the analogy with ferromagnetism suggests a strange possibility that may occur as the universe cools below the critical temperature.[12] Field theorists are used to the idea that whenever a continuous or discrete symmetry is broken by the appearance of a nonvanishing vacuum expectation value $\langle\phi_i\rangle$ of a scalar-field multiplet, it can be broken in a variety of ways, represented by the different directions of $\langle\phi_i\rangle$. Usually we regard these different directions as entirely equivalent, and ignore the multiplicity of broken-symmetry solutions. However, when a ferromagnet cools below its critical temperature, it does not acquire a single magnetization in some arbitrary direction; rather it breaks up into domains, each with its own direction of magnetization. Does the universe consist of domains, in which symmetries are broken in equivalent but different directions? If so, what happens when a particle or an observer travels from one domain to another?

For reasons of simplicity, it is assumed in this paper that we are interested in states of thermodynamic equilibrium in which all conserved quantum numbers have mean value zero, so that all chemical potentials vanish. (For this reason, the phase transition found here is quite unrelated to the superfluid transition in liquid helium.) It would not be at all difficult to include a chemical potential μ for an absolutely conserved quantity like baryon number; in this case the baryonic part of the term $\mathrm{Tr}[\gamma_4\Gamma_i\gamma_4\Gamma_j]$ in Eq. (1.1) would simply be multiplied with a factor

$$\frac{6}{\pi^2}\int_0^\infty\left(\frac{1}{e^{x-\mu}+1}+\frac{1}{e^{x+\mu}+1}\right)xdx\,,$$

with no change in any other results. This is an increasing function of the absolute value of the chemical potential μ, so the presence of a net baryon number would lower the critical temperature. However, μ appears to have a very small cosmological value,[13] of order 10^{-9}, in which case such effects would be quite negligible.

A much more interesting and challenging problem is presented by the possibility of a nonvanishing net mean value for some quantum number carried by bosons as well as fermions, which is exactly conserved only above the critical temperature. In this case we would have to consider not only the effects of a chemical potential but also the possibility of a true superfluid condensate at sufficiently high densities. Work on this problem is being continued.

II. GENERAL FORMALISM

We will consider a general renormalizable quantum field theory, which can be either a simple scalar theory, or a theory of scalar and spinor fields, or a full-fledged gauge theory, with or without spinor fields. For a simple scalar-field theory we would take the Lagrangian to be

$$\mathcal{L} = -\tfrac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i - P(\phi) , \tag{2.1}$$

where ϕ_i is a set of Hermitian spin-zero fields and $P(\phi)$ is a quartic polynomial. In this case we will assume $\mathcal{L}$ and $P(\phi)$ to be invariant under a group of global transformations with generators θ_α:

$$\frac{\partial P(\phi)}{\partial\phi_i}(\theta_\alpha)_{ij}\phi_j = 0 , \tag{2.2}$$

$$\theta_\alpha^\dagger = \theta_\alpha . \tag{2.3}$$

For a gauge theory, we would take the Lagrangian as

$$\mathcal{L} = -\tfrac{1}{2}(D_\mu\phi)_i(D^\mu\phi)_i - \tfrac{1}{4}F_{\alpha\mu\nu}F_\alpha^{\mu\nu} - \bar\psi\gamma^\mu D_\mu\psi - \bar\psi m_0\psi - P(\phi) - \bar\psi\Gamma_i\psi\phi_i, \tag{2.4}$$

where ϕ_i is a set of Hermitian spin-zero fields, $(D_\mu\phi)_i$ is their gauge-covariant derivative, $A_{\alpha\mu}$ is a set of gauge fields, $F_{\alpha\mu\nu}$ is their gauge-covariant curl, ψ_n is a set of spin-$\frac{1}{2}$ fields, $(D_\mu\psi)_n$ is their gauge-covariant derivative, m_0 is a gauge-invariant bare mass matrix, Γ_i is a gauge-covariant Yukawa coupling matrix, and $P(\phi)$ is a gauge-invariant quartic polynomial. This notation is explained more fully in Refs. 7 and 25; for our present purposes it will suffice to note that if θ_a are the Hermitian matrices representing the gauge generators on the scalar multiplet, then

$$(D_\mu\phi)_i \equiv \partial_\mu\phi_i - (\theta_\alpha)_{ij}\phi_j A_{\alpha\mu} , \tag{2.5}$$

and Eq. (2.2) now furnishes the necessary gauge-invariance condition on $P(\phi)$. Almost all of our discussion will apply equally well to theories described by the Lagrangian (2.1) or (2.4) or anything in between.

We shall need to impose some sort of weak-coupling condition in order to justify the use of perturbation theory. For the sake of both simplicity and physical relevance, it will be assumed that the orders of magnitude of the various parameters in the Lagrangian are characterized by a mass parameter $\mathfrak{M}$ and a *small* dimensionless coupling parameter $e \ll 1$, with

$$\begin{aligned}&\text{coefficient of quartic term in } P(\phi) \approx e^2 ,\\ &\text{coefficient of cubic term in } P(\phi) \approx e\mathfrak{M},\\ &\text{coefficient of quadratic term in } P(\phi) \approx \mathfrak{M}^2 ,\\ &\text{gauge couplings } (\theta_\alpha) \approx e ,\\ &\text{Yukawa couplings } (\Gamma_i) \approx e ,\\ &\text{Fermion bare mass } (m_0) \approx \mathfrak{M} .\end{aligned} \tag{2.6}$$

[For simplicity we are assuming that the Lagrangian involves no gauge-invariant scalar fields, so there are no linear terms in $P(\phi)$. Of course, we do not rule out the possibility that some of the parameters in the theory may be anomalously small; in particular the symmetries of the theory may require m_0, Γ_i, and/or the cubic term in $P(\phi)$ to vanish.] With this form of the weak-coupling assumption, the expansion of any given S-matrix element at zero temperature in powers of e^2 is the same as an expansion in the number of loops appearing in Feynman diagrams.

It will further be assumed that the symmetries of the Lagrangian are spontaneously broken at zero temperature. This symmetry breakdown is manifested in the appearance of a nonvanishing lowest-order vacuum expectation value λ_i of the scalar fields ϕ_i, given by

$$\frac{\partial P(\phi)}{\partial\phi_i} = 0 \text{ at } \phi = \lambda . \tag{2.7}$$

The criterion for spontaneous symmetry breaking in lowest order is

$$(\theta_\alpha)_{ij}\lambda_j \neq 0 . \tag{2.8}$$

It follows from the weak-coupling assumptions (2.6) that λ is of order

$$\lambda \approx \mathfrak{M}/e . \tag{2.9}$$

At this point, the reader may wonder how raising the temperature can possibly restore a broken symmetry. The effects of a finite temperature appear only through diagrams of higher order in e^2, so it appears that at all temperatures the leading term in the mean value of ϕ_i will be the temperature-independent term λ_i given by (2.7). The answer, to be discussed in the next section, is that the symmetries of the theory can only be affected by a finite temperature when the temperature is so high that powers of temperature can compensate for powers of e. However, before we can discuss such matters, we need to review the formalism for perturbative calculations at general finite temperature.

In general, the physical quantities with which we will be concerned here are the partition function

$$Q \equiv \mathrm{Tr}[e^{-H/\theta}] \tag{2.10}$$

and its variational derivatives with respect to external perturbations, the temperature Green's functions[4]

$$Q\langle T_\tau\{A(\vec{x}_1,\tau_1)B(\vec{x}_2,\tau_2)\cdots\}\rangle \equiv \mathrm{Tr}[T_\tau\{A(\vec{x}_1,\tau_1)B(\vec{x}_2,\tau_2)\cdots\}e^{-H/\theta}]\ , \qquad (2.11)$$

where H is the Hamiltonian, θ is the temperature (times Boltzmann's constant), $A(\vec{x},\tau)$ is an operator defined in terms of the Schrödinger-representation operators $A(\vec{x})$ by

$$A(\vec{x},\tau)\equiv e^{H\tau}A(\vec{x})e^{-H\tau}\ , \qquad (2.12)$$

and T_τ denotes ordering according to the values of τ, with τ values decreasing from left to right, and with an extra minus sign for odd permutations of fermion operators. The perturbation in Q caused by the addition of terms proportional to the operators $A(\vec{x}_1)$, $B(\vec{x}_2)$, etc. is an integral involving the temperature Green's functions (2.11) at τ values in the range

$$0\leqslant\tau\leqslant 1/\theta\ . \qquad (2.13)$$

It is therefore convenient to express these Green's functions as a Fourier integral over momenta and a Fourier *sum* over discrete energies. However, the Green's functions satisfy a periodicity property of having the same (opposite) values when any one of the τ's for a boson (fermion) operator has the values 0 and $1/\theta$; for instance

$$\begin{aligned}\langle T_\tau\{A(\vec{x}_1,1/\theta)B(\vec{x}_2,\tau_2)\cdots\}\rangle &= Q^{-1}\mathrm{Tr}[A(\vec{x}_1,1/\theta)T_\tau\{B(\vec{x}_2,\tau_2)\cdots\}e^{-H/\theta}]\\ &= Q^{-1}\mathrm{Tr}[T_\tau\{B(\vec{x}_2,\tau_2)\cdots\}e^{-H/\theta}A(\vec{x}_1,1/\theta)]\\ &= Q^{-1}\mathrm{Tr}[T_\tau\{B(\vec{x}_2,\tau_2)\cdots\}A(\vec{x}_1,0)e^{-H/\theta}]\\ &= \pm\langle T_\tau\{A(\vec{x}_1,0)B(\vec{x}_2,\tau_2)\cdots\}\rangle \qquad (2.14)\end{aligned}$$

with a $+(-)$ sign when A is a boson (fermion) operator. The same periodicity property also applies to arbitrary τ derivatives of the Green's function, and therefore requires that the Fourier sums contain only even or only odd Fourier components.[14] We can therefore write

$$\langle T_\tau\{A(\vec{x}_1,\tau_1)B(\vec{x}_2,\tau_2)\cdots\}\rangle = \int d^3p_1d^3p_2\cdots\sum_{\omega_1}\sum_{\omega_2}\cdots G(\vec{p}_1,\omega_1,\vec{p}_2,\omega_2,\cdots)\times\exp[i\vec{p}_1\cdot\vec{x}_1-i\omega_1\tau_1+i\vec{p}_2\cdot\vec{x}_2-i\omega_2\tau_2+\cdots]\ , \qquad (2.15)$$

where

$$\omega=\pi\theta\times\begin{cases}\text{even integer (bosons)}\\ \text{odd integer (fermions)}\ .\end{cases} \qquad (2.16)$$

(These integers can, of course, be positive or negative.)

There is a well-known diagrammatic procedure[4] for calculating the G's: Simply draw all Feynman graphs (dropping vacuum fluctuations) with one external line for each operator $A, B, \ldots,$ and evaluate as usual in field theory, except that every internal energy p^0 is replaced with a quantity $i\omega$ satisfying the "quantization" conditions (2.16), and all energy integrals are replaced with ω sums:

$$\begin{aligned}&p^0\to i\omega\ ,\\ &\int d^4p\to 2i\pi\theta\int d^3p\sum_\omega\ , \qquad (2.17)\\ &\delta^4(p-p')\to(2i\pi\theta)^{-1}\delta_{\omega\omega'}\delta^3(\vec{p}-\vec{p}')\ .\end{aligned}$$

The same procedure gives $\ln Q$ as the sum of connected diagrams with no external lines.

In what follows this diagrammatic procedure will be used to calculate Green's functions of gauge-invariant operators using the renormalizable "ξ gauge" of Fujikawa, Lee, and Sanda.[15] This use of a "nonunitarity" gauge may be justified by a three-step argument:

(a) First quantize the theory in the unitarity gauge,[7] and use the Hamiltonian in this gauge to derive finite-temperature Feynman rules as indicated above.

(b) In the same manner as in field theory,[7] show that these Feynman rules are equivalent to the Feynman rules for a ξ gauge with $\xi=0$.

(c) Either directly or by functional methods, show that the results obtained for the partition function or for the Green's functions of gauge-invariant generators are ξ-independent, and therefore correctly given by renormalizable ξ gauges[15] with $\xi\neq 0$.

The last two steps go through just as in field theory,[16] because none of the algebraic manipulations depend on whether we integrate over real energies or sum over complex energies. The same result can also be obtained by a more direct functional approach.[17]

One important advantage of our use of a renormalizable perturbative formalism is that we can check that the counterterms which remove divergences in S matrix elements at zero temperature also remove the divergences in finite-temperature Green's functions. For the milder divergences this can be seen from the classic formula[18]

$$h\sum_{n=-N}^{+N}f(nh)-\int_{-(N+1/2)h}^{(N+1/2)h}f(\omega)d\omega=\frac{-h^3}{24}\sum_{n=-N}^{+N}f''(\xi_nh)\ , \qquad (2.18)$$

where $f(\omega)$ is an arbitrary twice-differentiable function, h is an arbitrary interval (in our case taken as $2\pi\theta$), and ξ_n is for each n some definite point in the range

$$n-\tfrac{1}{2}\leq\xi_n\leq n+\tfrac{1}{2}\ .$$

Even if the sum and the integral on the left-hand side diverge for $N\to\infty$, their difference is finite in this limit, as long as the divergence is mild enough so that the right-hand side converges. This will in particular be the case for the linear and logarithmic divergences encountered in physical theories, for which $f(\omega)$ behaves like 1 or $1/\omega$ times powers of $\ln\omega$ as $|\omega|\to\infty$, so that $f''(\omega)$ behaves like $1/\omega^2$ or $1/\omega^3$ times powers of $\ln\omega$. In these cases we can pass to the limit $N\to\infty$ in (2.18), and we see that the divergences in the temperature-*dependent* sum on the left are removed by whatever temperature-*independent* subtraction renders the integral convergent. A similar result is obtained for the quadratic divergences in the next section.

III. CALCULATION OF THE EFFECTIVE POLYNOMIAL

We now begin our calculation of the temperature at which a broken symmetry is restored. As already mentioned in the last section, this can only happen in a weak-coupling theory at a temperature so high that powers of the temperature can compensate for powers of the coupling. The number of factors of e in a given graph is simply given by the number of loops, increasing by two units for each additional loop. Hence we must ask how many powers of θ are contributed by each loop.

Consider a single loop, with superficial divergence D, determined by counting powers of momenta as usual, including +4 for each loop. We can rescale all internal momenta as well as energies by a factor θ, so that the whole loop takes the form

$$\theta^D I(p_{\text{ext}}/\theta,\ \omega_{\text{ext}}/\theta,\ m_{\text{int}}/\theta), \tag{3.1}$$

where p_{ext} and ω_{ext} represent the various external momenta and energies, and m_{int} represents the various internal masses. Thus for $\theta\to\infty$, the loop behaves like θ^D, unless there are infrared divergences when the arguments of the function I vanish. If $D<1$ there are such infrared-divergences, but they occur only for a finite number of terms in the energy sum, in which two or more of the internal lines of the loop represent a boson carrying zero energy. Such terms are convergent three-dimensional integrals, and therefore can increase no faster than θ as $\theta\to\infty$ because aside from the factor θ in (2.17), the integrands are decreasing functions of θ. (Note that it *is* possible to get a factor θ even from a convergent loop with $D<0$, in particular from the term in which all internal boson energies vanish.) On the other hand, for $D>1$, there are no infrared divergences in $I(0,0,0)$, so the loop simply contributes a factor θ^D. The leading terms for large θ therefore come from those loops with $D>1$ which are as divergent as possible.

Now, aside from an uninteresting quartic divergence in $\ln Q$, the worst divergences in any renormalizable field theory are quadratic. We therefore conclude that *the leading terms for e small and θ large are those in which all loops beyond the lowest order are quadratically divergent.* The convergent part of such a loop contributes a factor θ^2e^2, so we can anticipate that the critical temperature is reached when θ^2e^2 is of order $\mathfrak{M}^2$, i.e.,

$$\theta_c\approx\mathfrak{M}/e\ . \tag{3.2}$$

At temperatures of this order of magnitude, the contribution of other loops is suppressed either by a factor $e^2\theta/\mathfrak{M}\approx e$ or by a factor e^2.

Further, we know that in any renormalizable field theory, including renormalizable gauge theories, all quadratic divergences can be eliminated by a renormalization of the quadratic term in the polynomial $P(\phi)$. [We are assuming here that there are no gauge-invariant scalar fields in the Lagrangian, in which case the quadratic terms in $P(\phi)$ are the only terms in the Lagrangian with the correct dimensionality needed to cancel quadratic divergences.] We therefore expect that *at finite temperature, all leading terms contributed by multiloop graphs, which survive when $e\ll1$ with $\theta\approx\mathfrak{M}/e$, as well as all quadratic divergences, are canceled by a redefinition of the polynomial part of the Lagrangian,*

$$P_{\text{eff}}(\phi)=P(\phi)+\tfrac{1}{2}Q_{ij}(\theta)\phi_i\phi_j\ , \tag{3.3}$$

and the compensating introduction of a counterterm in the interaction

$$\delta\mathfrak{L}'=\tfrac{1}{2}Q_{ij}(\theta)\phi_i\phi_j\ , \tag{3.4}$$

where Q_{ij} is some gauge-invariant quadratically divergent matrix. Since the nonleading terms are suppressed by factors $e^2\theta/\mathfrak{M}\approx e$ or e^2 for each loop, we conclude that any *Green's function is given to a lowest approximation for $e\ll1$ and $\theta\lesssim\mathfrak{M}/e$ by just the lowest-order graphs, but calculated using $P_{\text{eff}}(\phi)$ in place of $P(\phi)$.* In particular, we must define the perturbation expansion by using a shifted field

$$\phi_i'=\phi_i-\lambda_i\ , \tag{3.5}$$

with λ_i a minimum of the new polynomial

$$\frac{\partial P_{\rm eff}(\phi)}{\partial \phi_i}\bigg|_{\phi=\lambda} = 0 \,. \tag{3.6}$$

Thus the presence or absence of spontaneous symmetry breaking at any given temperature can be determined by an examination of the minimum of $P_{\rm eff}(\phi)$.

In order to calculate Q_{ij}, we note that the only graphs that contain quadratic divergences in any renormalizable theory are the tadpole T_i and the boson self-energy Π_{ij}. After we perform the shift (3.5), the interaction term (3.4) provides counterterms for both of these:

$$\delta\mathcal{L}' = \tfrac{1}{2}Q_{ij}\phi_i'\phi_j' + Q_{ij}\lambda_i\phi_j' + \tfrac{1}{2}Q_{ij}\lambda_i\lambda_j \,. \tag{3.7}$$

Hence we can determine Q_{ij} by requiring that the divergences and $\theta^2 e^2$ terms in Π_{ij} and T_i are canceled by the divergences and $\theta^2 e^2$ terms in (3.7). (To the order in e that concerns us here, we do not need to worry about other divergences.)

The one-loop tadpole graphs were calculated in a renormalizable ξ gauge at zero temperature in Ref. 25 (see Fig. 1). The result was

$$\begin{aligned} T_i^{(\theta=0)} = &-\tfrac{1}{2}\int d^4k f_{ijk}(k^2+M^2)^{-1}{}_{jk} \\ &+\int d^4k\, \mathrm{Tr}[\Gamma_i(i\gamma_\lambda k^\lambda + m)^{-1}] \\ &-3(\theta_\beta\theta_\alpha\lambda)_i\int d^4k(k^2+\mu^2)^{-1}{}_{\alpha\beta} \\ &+\tfrac{1}{2}(M^2\theta_\alpha\theta_\beta\lambda)_i \\ &\times\int d^4k(k^2)^{-1}(\xi k^2+\mu^2)^{-1}{}_{\alpha\beta}\,. \end{aligned} \tag{3.8}$$

(A $-i\epsilon$ term is understood in all denominators.) Here f_{ijk} is the trilinear coupling

$$f_{ijk} = \frac{\partial^3 P(\phi)}{\partial\phi_i\partial\phi_j\partial\phi_k}\bigg|_{\phi=\lambda}, \tag{3.9}$$

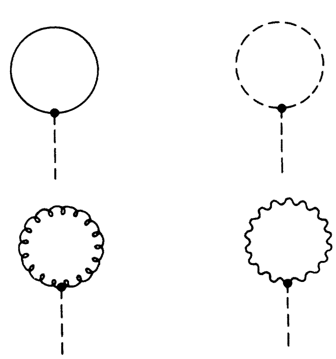


FIG. 1. Feynman graphs for the tadpole T. (Here dashed lines refer to scalar fields, solid lines refer to spinor fields, wavy lines refer to gauge fields, and looped lines refer to Faddeev-Popov "ghost" fields.)

and M, m, and μ are the lowest-order scalar, spinor, and vector mass matrices:

$$M^2{}_{ij} = \frac{\partial^2 P(\phi)}{\partial\phi_i\partial\phi_j}\bigg|_{\phi=\lambda}, \tag{3.10}$$

$$m = m_0 + \Gamma_i\lambda_i \,, \tag{3.11}$$

$$\mu^2{}_{\alpha\beta} = \lambda_i\lambda_j(\theta_\alpha\theta_\beta)_{ij} \,. \tag{3.12}$$

Also recall that θ_α is the Hermitian matrix (of order e) which appears in the gauge-covariant derivative of the scalar fields

$$(D_\mu\phi)_i \equiv \partial_\mu\phi_i - (\theta_\alpha)_{ij}\phi_j A_{\alpha\mu} \,. \tag{3.13}$$

We can easily extract the quadratically divergent part, and note that for $\xi\neq 0$ it is ξ-independent:

$$\begin{aligned} T_i^\infty = &\{-\tfrac{1}{2}f_{ikk} + \mathrm{Tr}[\Gamma_i\gamma_4 m\gamma_4] - 3(\theta_\alpha\theta_\alpha\lambda)_i\} \\ &\times\int d^4k(k^2)^{-1}\,. \end{aligned} \tag{3.14}$$

[Note that Γ_i may contain terms proportional to γ_5, which anticommute with both $(\gamma_\lambda k^\lambda)^{-1}$ and γ_4.] Under our assumption that the Lagrangian contains no gauge-invariant scalar fields, the first two terms are purely of first order in λ:

$$f_{ikk} = f_{ijkk}\lambda_j \,, \tag{3.15}$$

$$\mathrm{Tr}[\Gamma_i\gamma_4 m\gamma_4] = \mathrm{Tr}[\Gamma_i\gamma_4\Gamma_j\gamma_4]\lambda_j \,, \tag{3.16}$$

where f_{ijkl} is the coefficient of the quadrilinear term in $P(\phi)$

$$f_{ijkl} \equiv \frac{\partial^4 P(\phi)}{\partial\phi_i\partial\phi_j\partial\phi_k\partial\phi_l} \,. \tag{3.17}$$

Therefore, T_i^∞ is also of first order in λ:

$$\begin{aligned} T_i^\infty = &\{-\tfrac{1}{2}f_{ijkk} + \mathrm{Tr}[\Gamma_i\gamma_4\Gamma_j\gamma_4] - 3(\theta_\alpha\theta_\alpha)_{ij}\} \\ &\times\lambda_j\int d^4k(k^2)^{-1}\,. \end{aligned} \tag{3.18}$$

In accordance with the finite-temperature Feynman rules discussed in the last section, the leading terms in the tadpole are obtained by replacing the energy integral with a sum over the discrete energies (2.16):

$$\begin{aligned} T_i = &-i(2\pi)^4[\tfrac{1}{2}f_{ijkk} + 3(\theta_\alpha\theta_\alpha)_{ij}]\lambda_j I_B(\theta) \\ &+ i(2\pi)^4\mathrm{Tr}[\Gamma_i\gamma_4\Gamma_j\gamma_4]\lambda_j I_F(\theta)\,. \end{aligned} \tag{3.19}$$

where

$$I_B(\theta) \equiv (2\pi)^{-4}(2\pi\theta)\sum_{n=-\infty}^{\infty}\int d^3k[\vec{k}^2 + 4n^2\pi^2\theta^2]^{-1}\,, \tag{3.20}$$

$$I_F(\theta) \equiv (2\pi)^{-4}(2\pi\theta)\sum_{n=-\infty}^{\infty}\int d^3k[\vec{k}^2 + (2n+1)^2\pi^2\theta^2]^{-1}\,. \tag{3.21}$$

The counterterm (3.7) supplies an additional contribution

$$\delta T_i = i(2\pi)^4 Q_{ij}\lambda_j \ , \tag{3.22}$$

so in order to cancel the leading terms in the one-loop tadpole we must choose Q_{ij} as the matrix

$$Q_{ij}(\theta) = [\tfrac{1}{2}f_{ijkk} + 3(\theta_\alpha\theta_\alpha)_{ij}]I_B(\theta) - \mathrm{Tr}[\Gamma_i\gamma_4\Gamma_j\gamma_4]I_F(\theta) \ . \tag{3.23}$$

Before discussing the calculation of I_B and I_F, let us check that the counterterms in (3.7) now also cancel the leading terms in the scalar-boson self-energy. The one-loop self-energy graphs at zero boson momentum and zero temperature were also calculated in Ref. 25 (see Fig. 2). The result was

$$\begin{aligned}
\Pi_{ij}^{(\theta=0)}(0) = {} & \frac{-3i}{2(2\pi)^4}(\{\theta_\alpha, \theta_\beta\}\lambda)_i(\theta_\gamma, \theta_\delta\}\lambda)_i \int d^4k(k^2+\mu^2)^{-1}{}_{\alpha\gamma}(k^2+\mu^2)^{-1}{}_{\beta\delta} \\
& - \frac{i}{2(2\pi)^4} f_{ikl}f_{jpq} \int d^4k(k^2+M^2)^{-1}{}_{kp}(k^2+M^2)^{-1}{}_{lq} \\
& - \frac{i}{(2\pi)^4} \int d^4k(k^2)^{-1}(\xi k^2+\mu^2)^{-1}{}_{\alpha\beta}[M^2\theta_\alpha(k^2+M^2)^{-1}\theta_\beta M^2 + M^2\theta_\alpha(k^2+M^2)^{-1}M^2\theta_\beta \\
& \qquad + \theta_\alpha M^2(k^2+M^2)^{-1}\theta_\beta M^2]_{ij} \\
& - \frac{i}{(2\pi)^4} \int d^4k\, \mathrm{Tr}[\Gamma_i(i\gamma_\lambda k^\lambda + m)^{-1}\Gamma_j(i\gamma_\lambda k^\lambda + m)^{-1}] \\
& - \frac{i}{2(2\pi)^4}(M^2\theta_\gamma\theta_\alpha\lambda)_i(M^2\theta_\beta\theta_\delta\lambda)_j \int d^4k(k^2)^{-1}(\xi k^2+\mu^2)^{-1}{}_{\alpha\beta}(\xi k^2+\mu^2)^{-1}{}_{\gamma\delta} \\
& + \frac{3i}{2(2\pi)^4}(\{\theta_\beta, \theta_\alpha\})_{ij} \int d^4k(k^2+\mu^2)^{-1}{}_{\alpha\beta} + \frac{i}{2(2\pi)^4} f_{ijkl} \int d^4k(k^2+M^2)^{-1}{}_{kl} \\
& - \frac{i}{2(2\pi)^4} f_{ijk}M^{-2}{}_{kl}f_{lpq} \int d^4k(k^2+M^2)^{-1}{}_{pq} + \frac{i}{(2\pi)^4} f_{ijk}M^{-2}{}_{kl} \int d^4k\, \mathrm{Tr}[\Gamma_l(i\gamma_\lambda k^\lambda + m)^{-1}] \\
& - \frac{3i}{2(2\pi)^4} f_{ijk}M^{-2}{}_{kl}(\{\theta_\alpha, \theta_\beta\}\lambda)_l \int d^4k(k^2+\mu^2)^{-1}{}_{\alpha\beta} \ .
\end{aligned}$$

The quadratically divergent part for $\xi \neq 0$ is ξ-independent:

$$\begin{aligned}
\Pi_{ij}^{\infty} = \frac{i}{(2\pi)^4}\{ & -\mathrm{Tr}[\Gamma_i\gamma_4\Gamma_j\gamma_4] + 3(\theta_\alpha\theta_\alpha)_{ij} + \tfrac{1}{2}f_{ijkk} \\
& - \tfrac{1}{2}f_{ijk}M^{-2}{}_{kl}f_{lpp} + f_{ijk}M^{-2}{}_{kl}\mathrm{Tr}[\Gamma_l\gamma_4 m\gamma_4] - 3f_{ijk}M^{-2}{}_{kl}(\theta_\alpha\theta_\alpha\lambda)_l\} \int d^4k(k^2)^{-1} \ .
\end{aligned}$$

Replacing energy integrals by energy sums and using (3.15), (3.16), and (3.23), the leading term at finite temperature may be written

$$\Pi_{ij} = -Q_{ij} + f_{ijk}M^{-2}{}_{kl}Q_{lm}\lambda_m \ .$$

The first term here is immediately canceled by the first term in (3.7), while the second is canceled by the tadpole (3.22) produced by the second term in (3.7).

Returning now to the functions (3.20) and (3.21), we note that the sums may be turned back into integrals,

$$I_B(\theta) = \frac{-i}{2(2\pi)^4} \int d^3k \oint_C d\omega(\vec{k}^2+\omega^2)^{-1}\cot\left(\frac{\omega}{2\theta}\right) \ ,$$

$$I_F(\theta) = \frac{+i}{2(2\pi)^4} \int d^3k \oint_C d\omega(\vec{k}^2+\omega^2)^{-1}\tan\left(\frac{\omega}{2\theta}\right) .$$

The contour C runs from $+\infty$ to $-\infty$ just above the real axis and then back from $-\infty$ to $+\infty$ just below the real axis. By closing the two halves of this contour with large semicircles in the upper and

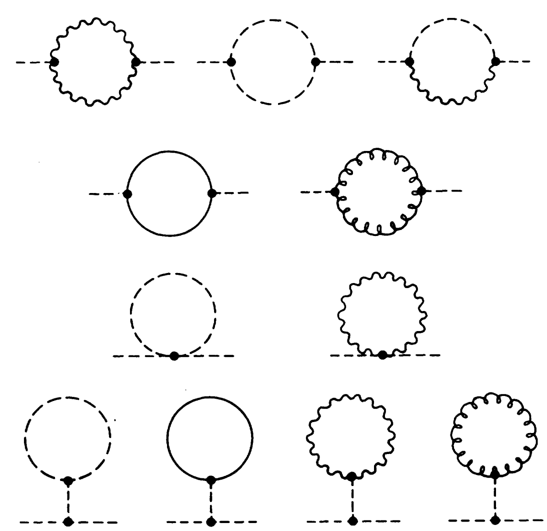


FIG. 2. Feynman graphs for the scalar self-energy Π. (Conventions same as in Fig. 1.)

lower half-planes, we pick up the poles at $\omega = \pm|\vec{k}|$,

$$I_B(\theta) = (2\pi)^{-3} \int \frac{d^3k}{2|\vec{k}|} \coth\frac{|\vec{k}|}{2\theta} \quad ,$$

$$I_F(\theta) = (2\pi)^{-3} \int \frac{d^3k}{2|\vec{k}|} \tanh\frac{|\vec{k}|}{2\theta} \quad .$$

These integrals are of course divergent, but their divergences can be separated out by extracting their values at $\theta = 0$:

$$I_B(\theta) = I_B(0) + \tfrac{1}{12}\theta^2 ,$$
$$I_F(\theta) = I_F(0) - \tfrac{1}{24}\theta^2 . \qquad (3.24)$$

This is important because it shows that the same infinite counterterm which removes divergences at zero temperature continues to remove them at all temperatures; in fact, we may recognize the divergences here as just the same ones we encounter in field theory:

$$I_B(0) = I_F(0)$$
$$= (2\pi)^{-3} \int \frac{d^3k}{2|\vec{k}|}$$
$$= \frac{-i}{(2\pi)^4} \int \frac{d^4k}{k^2 - i\epsilon} \quad .$$

Using (3.24) in (3.23), we have finally

$$Q_{ij}(\theta) = Q_{ij}(0) + \tfrac{1}{24}\theta^2\{f_{ijkk} + 6(\theta_\alpha\theta_\alpha)_{ij} + \mathrm{Tr}[\Gamma_i\gamma_4\Gamma_j\gamma_4]\} \quad . \qquad (3.25)$$

[We note that f_{ijkl} is of order e^2, while θ_α and Γ_i are of order e, so $Q - Q(0)$ is of order $e^2\theta^2$, as anticipated.] The term $Q_{ij}(0)$ in (3.25) just serves to provide a temperature-independent quadratically divergent renormalization of the mass parameters in the polynomial, so we may write (3.3) as

$$P_{\text{eff}}(\phi) = P_{\text{ren}}(\phi) + \tfrac{1}{48}\theta^2\{f_{ijkk} + 6(\theta_\alpha\theta_\alpha)_{ij} + \mathrm{Tr}[\Gamma_i\gamma_4\Gamma_j\gamma_4]\}\phi_i\phi_j \ , \qquad (3.26)$$

where $P_{\text{ren}}(\phi)$ is just the original polynomial $P(\phi)$, but with masses replaced by renormalized values. This formula will be used in the next section to determine the critical temperature.

Even though this has so far been a one-loop calculation, it is actually valid to lowest order in e but to all orders in $e\theta$. We could insert another loop in the single-loop diagram used to calculate the tadpoles or scalar self-energies, and if this new loop were quadratically divergent it would contribute a non-negligible factor $e^2\theta^2$, but the *old* loop would then not be quadratically divergent, and therefore would be suppressed at least by a factor e (see Fig. 3). More generally, we expect multiloop as well as single-loop diagrams for T_i to involve only a *single* factor of θ^2, so it is only the one-loop diagram that survives when $e \ll 1$ and $\theta \approx \mathfrak{M}/e$.

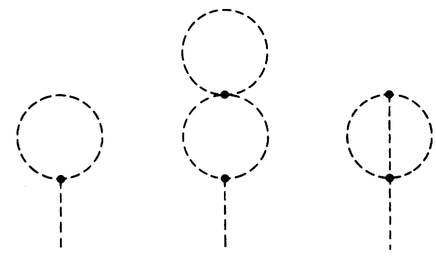


FIG. 3. One- and two-loop graphs for the tadpole T in a scalar field theory. The order of magnitude of the various contributions is indicated below each graph.

IV. THE CRITICAL TEMPERATURES

We have learned in the last section that the leading effect of multiloop graphs at temperatures θ of order $\mathfrak{M}/e$ is to change the polynomial $P(\phi)$ in the Lagrangian to the effective polynomial given by Eq. (3.26). The symmetry group of the Green's functions at a given temperature consists simply of that subgroup of the invariance group of the Lagrangian which leaves invariant the point λ_i at which $P_{\text{eff}}(\phi)$ has its minimum. We can therefore locate the various critical temperatures of the theory by asking at what temperature the symmetries of λ_i are gained or lost.

In particular, we note that if the temperature-dependent part of $P_{\text{eff}}(\phi)$ is a positive-definite function of ϕ, then at sufficiently high temperatures the minimum of $P_{\text{eff}}(\phi)$ must be at $\phi = 0$. This is because the quartic part of $P_{\text{eff}}(\phi)$ is in any case positive-definite (otherwise the energy would be unbounded below for large ϕ), while for sufficiently large θ the total quadratic term in $P_{\text{eff}}(\phi)$ will also be positive-definite (and large enough to overwhelm any cubic term that is not overwhelmed by the quartic terms). Thus we conclude that if the θ^2 term in P_{eff} is positive-definite then there is always a highest critical temperature, above which the Green's functions exhibit the full symmetry group of the theory.

The $\theta_\alpha\theta_\alpha$ and $\Gamma_i\Gamma_i$ in (3.26) are indeed positive matrices, because the θ and Γ matrices satisfy Hermiticity conditions,

$$\theta_\alpha^\dagger = \theta_\alpha, \quad \Gamma_i^\dagger = \gamma_4\Gamma_i\gamma_4 \ .$$

In fact, these terms are positive-*definite*, unless there are no gauge couplings or Yukawa couplings in the theory at all. However, the f term in (3.26), while usually positive-definite, is not always so.[19] We shall take a look at two examples where symmetry is restored at high temperature, and one example where it is not.

Example 1: Global O(n) with one n-vector

Let us consider a nongauge theory, invariant under a global group O(n), involving a single n-vector multiplet of scalar fields ϕ_i. The polynomial P will be of the form

$$P(\phi)=\tfrac{1}{2}\mathfrak{M}_0{}^2\phi_i\phi_i+\tfrac{1}{4}e^2(\phi_i\phi_i)^2 ,$$

where $\mathfrak{M}_0{}^2$ and e^2 are real quantities, with $e^2>0$ but $\mathfrak{M}_0{}^2$ of arbitrary sign. The quadrilinear coupling coefficient here is

$$f_{ijkl}=2e^2(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}) ,$$

so (3.25) gives

$$P_{\rm eff}(\phi)=\tfrac{1}{2}\mathfrak{M}^2(\theta)\phi_i\phi_i+\tfrac{1}{4}e^2(\phi_i\phi_i)^2 ,$$

where

$$\mathfrak{M}^2(\theta)=\mathfrak{M}^2(0)+\tfrac{1}{12}(n+2)e^2\theta^2 ,$$

and $\mathfrak{M}^2(0)$ is $\mathfrak{M}_0{}^2$ plus renormalization counterterms (see Fig. 4). If $\mathfrak{M}^2(0)$ is negative then for sufficiently low temperatures $\mathfrak{M}^2(\theta)$ will also be negative, and $P_{\rm eff}(\phi)$ will have an O(n)-noninvariant minimum at $\phi_i=\lambda_i$, with

$$e^2\lambda_i\lambda_i=-\mathfrak{M}^2(\theta)>0 .$$

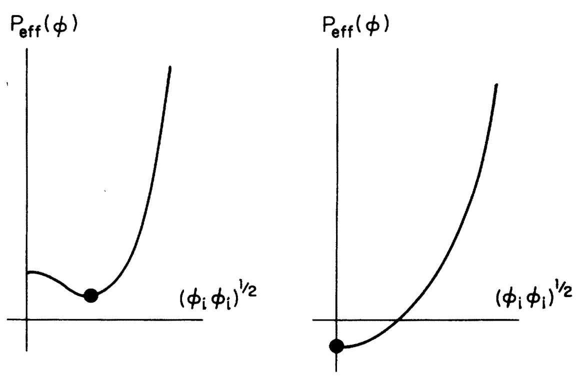
The full O(n) symmetry is therefore restored at a temperature θ_c such that

$$\mathfrak{M}^2(\theta_c)=0 ,$$

or[20]

$$\theta_c=\left(\frac{12}{n+2}\right)^{1/2}\frac{|\mathfrak{M}(0)|}{e} .$$

In order to express this in terms of observables,



(a) $\theta<\theta_c$ (b) $\theta>\theta_c$

FIG. 4. Schematic representation of the effective polynomial in Examples 1 and 2 of Sec. IV, below and above the critical temperature. The dark dot indicates the state of thermal equilibrium.

we may note that the physical zeroth-order mass matrix of the scalar fields is

$$\begin{aligned}M^2{}_{ij}(\theta)&=\left.\frac{\partial^2P_{\rm eff}(\phi)}{\partial\phi_i\,\partial\phi_j}\right|_{\phi=\lambda}\\&=\mathfrak{M}^2(\theta)\delta_{ij}+e^2(\delta_{ij}\lambda_k\lambda_k+2\lambda_i\lambda_j)\\&=2e^2\lambda_i\lambda_j ,\end{aligned}$$

so for $\theta<\theta_C$ the excitation spectrum consists of $n-1$ Goldstone bosons of zero mass and one boson of mass

$$M^2(\theta)=2e^2\lambda_i\lambda_i=-2\mathfrak{M}^2(\theta) .$$

We can therefore rewrite the critical temperature in terms of the single nonzero boson mass at zero temperature:

$$\theta_c=\left(\frac{6}{n+2}\right)^{1/2}\left(\frac{M(0)}{e}\right) .$$

Above the critical temperature the excitation spectrum consists of n degenerate bosons with mass $\mathfrak{M}(\theta)$.

Example 2: Local O(n) with one n-vector

Next, let us consider a *gauge* theory based on the group O(n), again with a single n-vector multiplet of scalar fields. The generators of the gauge group may be represented by matrices,

$$(\theta_{kl})_{ij}=ie'(\delta_{ki}\delta_{lj}-\delta_{kj}\delta_{li}) ,\quad 1\leqslant k<l\leqslant n$$

with a prime on the gauge coupling constant to distinguish it from the boson self-coupling constant e. The Casimir operator here is

$$\sum_{k<l}(\theta_{kl}\theta_{kl})_{ij}=(n-1)e'^2\delta_{ij} .$$

We take the polynomial in the Lagrangian again of the form

$$P(\phi)=\tfrac{1}{2}\mathfrak{M}_0{}^2\phi_i\phi_i+\tfrac{1}{4}e^2(\phi_i\phi_i)^2 .$$

Equation (3.26) now gives, as in case (a),

$$P_{\rm eff}(\phi)=\tfrac{1}{2}\mathfrak{M}^2(\theta)\phi_i\phi_i+\tfrac{1}{4}e^2(\phi_i\phi_i)^2 ,$$

where

$$\mathfrak{M}^2(\theta)=\mathfrak{M}^2(0)+\tfrac{1}{12}(n+2)e^2\theta^2+\tfrac{1}{4}(n-1)e'^2\theta^2 .$$

For $\mathfrak{M}^2(0)<0$ there is again a critical temperature θ_c, determined by the condition that $\mathfrak{M}^2(\theta_c)$ should vanish[20]:

$$\theta_c=[\tfrac{1}{12}(n+2)e^2+\tfrac{1}{4}(n-1)e'^2]^{-1/2}|\mathfrak{M}(0)| .$$

We see here an example of the general phenomenon, that adding gauge fields lowers the critical temperature. For $\theta<\theta_c$, this theory has an excitation spectrum consisting of one scalar boson of mass

$$M^2(\theta) = -2\mathfrak{M}^2(\theta)$$

(the remaining $n-1$ massless scalars are now unphysical) plus $n-1$ vector bosons with mass given by Eq. (3.12):

$$\mu^2(\theta) = e'^2\lambda_i(\theta)\lambda_i(\theta)$$
$$= \left(\frac{e'}{e}\right)|\mathfrak{M}^2(\theta)|$$

plus $(n-1)(n-2)/2$ vector bosons of zero mass, corresponding to the unbroken $O(n-1)$ subgroup. For $\theta > \theta_c$ the theory has n scalar bosons of mass $\mathfrak{M}(\theta)$ plus $n(n-1)/2$ vector bosons of zero mass. At θ_c there is evidently a transmutation of the zero-helicity states of $n-1$ vector bosons into $n-1$ scalar bosons, with the mass of all these bosons vanishing at θ_c to make the transmutation possible (see Fig. 5).

Example 3: Global $O(n) \times O(n)$ with two n-vectors

As an example with a very different behavior, let us consider an $O(n)\times O(n)$-invariant theory with two independent scalar multiplets χ_A and η_a transforming according to the representations $(n, 1)$ and $(1, n)$. The polynomial in the Lagrangian must take the form

$$P(\chi,\eta) = \tfrac{1}{2}\mathfrak{M}_\chi{}^2\chi_A\chi_A + \tfrac{1}{2}\mathfrak{M}_\eta{}^2\eta_a\eta_a + \tfrac{1}{4}e_{\chi\chi}{}^2(\chi_A\chi_A)^2 - \tfrac{1}{2}e_{\chi\eta}{}^2(\chi_A\chi_A)(\eta_a\eta_a) + \tfrac{1}{4}e_{\eta\eta}{}^2(\eta_a\eta_a)^2\,,$$

with parameters subject to the positivity constraints

$$e_{\chi\chi}{}^2 > 0,\quad e_{\eta\eta}{}^2 > 0,\quad e_{\chi\eta}{}^2 < |e_{\chi\chi}e_{\eta\eta}|\,.$$

(We use capital indices $A, B, \ldots$ for the χ's and lower-case indices $a, b, \ldots$ for the η's but all these indices run from 1 to n.) The nonvanishing elements of the quadrilinear coupling coefficient f_{ijkl} are now

$$f_{ABCD} = 2e_{\chi\chi}{}^2(\delta_{AB}\delta_{CD} + \delta_{AC}\delta_{BD} + \delta_{AD}\delta_{BC})\,,$$
$$f_{ABab} = -2e_{\chi\eta}{}^2\delta_{AB}\delta_{ab}\,,$$
$$f_{abcd} = 2e_{\eta\eta}{}^2(\delta_{ab}\delta_{cd} + \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc})\,,$$

together with other elements obtained by permutation of indices. The effective polynomial (3.26) therefore has the form

$$P(\chi,\eta) = \tfrac{1}{2}\mathfrak{M}_\chi{}^2(\theta)\chi_A\chi_A + \tfrac{1}{2}\mathfrak{M}_\eta{}^2(\theta)\eta_a\eta_a + \tfrac{1}{4}e_{\chi\chi}{}^2(\chi_A\chi_A)^2 - \tfrac{1}{2}e_{\chi\eta}{}^2(\chi_A\chi_A)(\eta_a\eta_a) + \tfrac{1}{4}e_{\eta\eta}{}^2(\eta_a\eta_a)^2\,,$$

where

$$\mathfrak{M}_\chi{}^2(\theta) = \mathfrak{M}_\chi{}^2(0) + \tfrac{1}{12}\theta^2[(n+2)e_{\chi\chi}{}^2 - ne_{\chi\eta}{}^2]\,,$$
$$\mathfrak{M}_\eta{}^2(\theta) = \mathfrak{M}_\eta{}^2(0) + \tfrac{1}{12}\theta^2[(n+2)e_{\eta\eta}{}^2 - ne_{\chi\eta}{}^2]\,.$$

For $e_{\chi\eta}{}^2$ positive, which as we shall see is the interesting case, there are four possible phases (see Fig. 6):

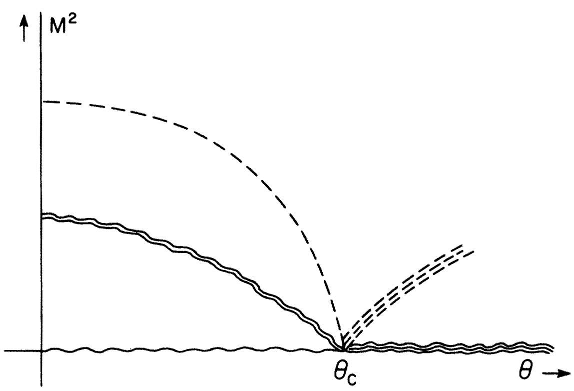


FIG. 5. Schematic representation of the excitation spectrum as a function of temperature for the gauge field theory discussed in Example 2 of Sec. IV. The gauge group here is taken as O(3). Wavy lines indicate particles of spin 1; dashed lines particles of spin 0. Note the continuity in the total numbers of helicity states at the critical temperature.

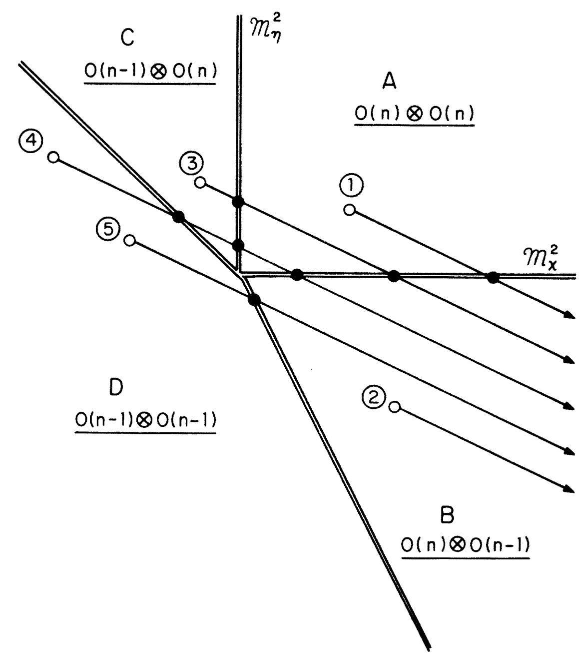


FIG. 6. Phase diagram for the theory described in Example 3 of Sec. IV. Phase boundaries are indicated by double lines. The values of $\mathfrak{M}_\chi{}^2$ and $\mathfrak{M}_\eta{}^2$ at zero temperature are indicated by open circles; the arrows indicate the behavior of $\mathfrak{M}_\chi{}^2$ and $\mathfrak{M}_\eta{}^2$ for large temperature. Critical temperatures are indicated by dark circles. The numbers in circles refer to the cases listed in the text.

(A) $\mathfrak{M}_\chi^2(\theta)>0, \quad \mathfrak{M}_\eta^2(\theta)>0$.

In this case the only minimum of $P_{\rm eff}(\chi,\eta)$ is at $\chi_A=\eta_a=0$, so the symmetry $O(n)\times O(n)$ is unbroken.

(B) $\mathfrak{M}_\chi^2(\theta)>0$,

$$0>\mathfrak{M}_\eta^2(\theta)>-(e_{\eta\eta}^2/e_{\chi\eta}^2)\mathfrak{M}_\chi^2(\theta)\ .$$

In this case the only minimum of $P_{\rm eff}(\chi,\eta)$ is at $\chi_A=0$, $\eta_a\neq 0$, so the symmetry is broken down to $O(n)\times O(n-1)$.

(C) $\mathfrak{M}_\eta^2(\theta)>0$,

$$0>\mathfrak{M}_\chi^2(\theta)>-(e_{\chi\chi}^2/e_{\chi\eta}^2)\mathfrak{M}_\eta^2(\theta)\ .$$

In this case the only minimum of $P_{\rm eff}(\chi,\eta)$ is at $\eta_a=0$, $\chi_A\neq 0$, so the symmetry is broken down to $O(n-1)\times O(n)$.

(D) For all other values of $\mathfrak{M}_\chi^2(\theta)$ and $\mathfrak{M}_\eta^2(\theta)$ the *deepest* minimum is at $\chi_A\neq 0$, $\eta_a\neq 0$, so the symmetry is broken down to $O(n-1)\times O(n-1)$.

The phase for $\theta\to\infty$ is always (A), (B), or (C), depending on the relative values of the coupling constants. If for example we choose

$$(n+2)e_{\chi\chi}^2>ne_{\chi\eta}^2>(n+2)e_{\eta\eta}^2$$

[which is consistent with the positivity requirements on $P(\chi,\eta)$ for large fields] then for $\theta\to\infty$ the system is necessarily in phase (B). We see that with this choice of coupling constants, *the symmetry is necessarily broken down to* $O(n)\times O(n-1)$ *at high temperature*. (The same is true if we introduce gauge fields, providing that the gauge coupling constant is sufficiently small compared with $|e_{\chi\eta}|$.) The critical points encountered at lower temperature depend on the signs and relative magnitude of $\mathfrak{M}_\chi^2(\theta)$ and $\mathfrak{M}_\eta^2(\theta)$ at zero temperature. We may distinguish the following cases (see Fig. 6):

(1) $\mathfrak{M}_\chi^2(0)>0, \quad \mathfrak{M}_\eta^2(0)>0$.

There is a single critical temperature, at which the phase changes from type (A) at low temperature to type (B) at high temperature.

(2) $\mathfrak{M}_\chi^2(0)>0$,

$$0>\mathfrak{M}_\eta^2(0)>-(e_{\eta\eta}^2/e_{\chi\eta}^2)\mathfrak{M}_\chi^2(0)\ .$$

There are no critical temperatures; the phase is of type (B) at all temperatures.

(3) $\mathfrak{M}_\eta^2(0)>0$,

$$0>\mathfrak{M}_\chi^2(0)>-(e_{\chi\chi}^2/e_{\chi\eta}^2)\mathfrak{M}_\eta^2(0)\ .$$

There are *two* critical temperatures, at which the phase changes from type (C) at low temperature to type (A) at medium temperature to type (B) at high temperature.

(4) $\mathfrak{M}_\eta^2(0)>0$,

$$-(e_{\chi\chi}^2/e_{\chi\eta}^2)\mathfrak{M}_\eta^2(0)>\mathfrak{M}_\chi^2(0)>-\left(\frac{(n+2)e_{\chi\chi}^2-ne_{\chi\eta}^2}{ne_{\chi\eta}^2-(n+2)e_{\eta\eta}^2}\right)\times\mathfrak{M}_\eta^2(0)\ .$$

There are *three* critical temperatures, at which the phase changes from type (D) at low temperature, to type (C) and then to type (A) at medium temperature, and finally to type (B) at high temperature.

(5) For other values of $\mathfrak{M}_\chi^2(0)$ and $\mathfrak{M}_\eta^2(0)$, there is a single critical temperature, at which the phase changes from type (D) at low temperature to type (B) at high temperature.

The existence of an (A)$\to$(B) critical point, at which the symmetry shifts with increasing temperature from the group $O(n)\times O(n)$ to a *smaller* group $O(n)\times O(n-1)$, runs counter to most of our experience with macroscopic systems. For instance, heating a superconductor restores gauge invariance; heating a ferromagnet restores rotational invariance; heating a crystal restores translational invariance. However, the example of Rochelle salt[6] reassures us that there is nothing impossible about a loss of symmetry with increasing temperature.

V. COMPARISON WITH UNITARITY GAUGE CALCULATIONS

It is instructive to compare the results we have obtained by calculating in the renormalizable ξ gauges with the corresponding results that would be obtained in the unitarity gauge, for which $\xi=0$. In this case, we would have had no reason to expect that the counterterm (3.7) would cancel all the $e^2\theta^2$ terms. In particular, the quadratically divergent part of the tadpole (3.8) would have contained an additional term,

$$T_{iU}^\infty=T_i^\infty+\tfrac{1}{2}(\mu^{-2})_{\alpha\beta}(M^2\theta_\alpha\theta_\beta\lambda)_i\int d^4k(k^2)^{-1}. \tag{5.1}$$

(A subscript U denotes the use of the unitarity gauge; quantities without this subscript are calculated in the renormalizable gauges with $\xi\neq 0$.) This implies a new temperature-dependent term,

$$T_{iU}=T_i+\tfrac{1}{2}i(2\pi)^4I_B(\theta)(\mu^{-2})_{\alpha\beta}(M^2\theta_\alpha\theta_\beta\lambda)_i\ . \tag{5.2}$$

This cannot in general be canceled by a counterterm of the form (3.7), because

$$\frac{\partial T_{iU}}{\partial\lambda_j}\neq\frac{\partial T_{jU}}{\partial\lambda_i}\ .$$

The counterterm (3.7) does in general cancel the quadratic divergences and $e^2\theta^2$ terms even for $\xi=0$ in a general *gauge invariant* Green's function, but this cancellation occurs because for $\xi=0$ there are quadratic divergences and $e^2\theta^2$ terms contributed by a wide variety of diagrams besides tadpoles and scalar self-energies. Thus it is not possible to determine the matrix Q_{ij} in (3.7) by inspection of tadpole graphs evaluated in the unitarity gauge.

The inadequacy of the unitarity gauge for our purposes may be obscured by the fact that in certain specially simple gauge theories[21] T_{iU} does take the form of a λ gradient. These theories are characterized by the condition that the scalar fields belong to a representation of the gauge group which is "transitive on the sphere," i.e., for which any direction in the representation space of the scalar fields may be rotated into any other direction by a gauge transformation. In this case $P_{\text{eff}}(\phi)$ must be of the same form as in an O(n)-invariant theory,

$$P_{\text{eff},U}(\phi)=\tfrac{1}{2}\mathfrak{M}_U{}^2(\theta)\phi_i\phi_i+\tfrac{1}{4}e^2(\phi_i\phi_i)^2, \qquad (5.3)$$

so we can take over the results of examples 1 and 2 of the last section. In particular, we again have

$$M^2{}_{ij}=2e^2\lambda_i\lambda_j\,,$$

so (5.2) gives

$$T_{iU}=T_i+ie^2(2\pi)^4I_B(\theta)\lambda_i\,. \qquad (5.4)$$

The new term here can be canceled by a new term in the matrix Q_{ij} in (3.7):

$$Q_{ijU}(\theta)=Q_{ij}(\theta)-e^2I_B(\theta)\delta_{ij}\,. \qquad (5.5)$$

This new term in $Q_{ij}(\theta)$ leads to a decrease in the temperature-dependent mass term in P_{eff}, and hence to an increase in the critical temperature. For instance, the mass term calculated in example 2 of the last section using the renormalizable gauges was

$$\mathfrak{M}^2(\theta)=\mathfrak{M}^2(0)+\tfrac{1}{12}(n+2)e^2\theta^2+\tfrac{1}{4}(n-1)e'^2\theta^2, \qquad (5.6)$$

so the new term in (5.5) changes this to

$$\mathfrak{M}_U{}^2(\theta)=\mathfrak{M}_U{}^2(0)+\tfrac{1}{12}(n+1)e^2\theta^2+\tfrac{1}{4}(n-1)e'^2\theta^2, \qquad (5.7)$$

and the critical temperature for $\mathfrak{M}_U{}^2(0)<0$ is[20]

$$\theta_c=|\mathfrak{M}_U(0)|\left[\tfrac{1}{12}(n+1)e^2+\tfrac{1}{4}(n-1)e'^2\right]^{-1/2}. \qquad (5.8)$$

Since unitarity-gauge calculations give different results from renormalizable-gauge calculations, which should we believe? The answer is provided by our discussion in Sec. III: It is only in the renormalizable gauge that a change in the scalar-field polynomial restores the validity of perturbation theory at high temperatures, and therefore only in the renormalizable gauges can we use the effective polynomial to study the pattern of symmetry breaking. When the scalar field representation is transitive on a sphere, the introduction of an effective polynomial does eliminate the $e^2\theta^2$ terms found in the unitarity gauge in tadpole graphs, but there are plenty of other quadratic divergences and $e^2\theta^2$ terms in this gauge which are not thereby eliminated, and there is no reason to regard the effective polynomial as being of any special importance.

VI. HIGHER-ORDER EFFECTS

As frequently emphasized, these calculations of the effective polynomial are valid to all orders in $e^2\theta^2$ but only to lowest order in e^2. Suppose we wish to locate the critical temperature more exactly, or try to determine the precise nature of the phase transition. How would we go about it?

The introduction of the counterterm (3.7) had the purpose of restoring the validity of perturbation theory for $e^2\ll 1$ at temperatures of order $\mathfrak{M}/e$. As long as perturbation theory is valid, we can infer the symmetry properties of the theory by a study of its lowest-order terms, i.e., of $P_{\text{eff}}(\phi)$. However, even with the counterterm (3.7) working to cancel the $e^2\theta^2$ terms, the perturbation theory still breaks down when θ is very near θ_c, because in this case one or more of the mass terms in $P_{\text{eff}}(\phi)$ becomes very small, and so powers of e can be canceled by factors which become infrared-divergent for $\theta=\theta_c$.

At first sight, this problem appears very similar to that studied by Coleman and E. Weinberg.[8] Instead of adjusting the temperature so that the mass term in an effective polynomial vanishes, they considered a relativistic quantum field theory at zero temperature, with the bare scalar masses adjusted so that the renormalized scalar masses vanish. In order to decide whether the symmetries of the theory were spontaneously broken, they searched for the minima of an effective potential, the first term of which is just $P(\phi)$, using renormalization-group methods to sum up the logarithms associated with infrared divergences.

The difference here is that the infrared divergences are profoundly affected by a finite temperature.[22] According to the Feynman rules discussed in Sec. II, the only denominators which can ever vanish are those in boson propagators carrying zero energy. Since the energy is a discrete variable, the degree of infrared divergence must be determined counting *three* rather than four powers of momentum for each loop.

Hence the infrared divergences are those expected in a superrenormalizable rather than an ordinary renormalizable theory, and the appropriate methods needed here are those of Wilson[23] rather than those of Coleman and Weinberg.[8] The difference between our problem and that studied by Wilson is that we do not have a cutoff; ultraviolet divergences are eliminated by the renormalization procedure.

Without pursuing this problem too far, we can at least estimate how close the actual critical temperature, at which a broken symmetry is restored, is to the approximate critical temperature θ_c calculated in Sec. IV. For temperatures well below or well above θ_c (say, $|\theta-\theta_c|>\frac{1}{2}\theta_c$) perturbation theory is presumably valid [with the counterterm (3.7) canceling the $e^2\theta^2$ terms], so the symmetry of the theory must be just that of the lowest-order terms. On the other hand, when $\theta-\theta_c$ is small, the mass terms $\mathfrak{M}(\theta)$ in $P_{\rm eff}(\phi)$ vanish, with

$$\mathfrak{M}(\theta)\approx[e^2(\theta^2-\theta_c^2)]^{1/2}\approx e\theta_c^{1/2}(\theta-\theta_c)^{1/2}.$$

Our order-of-magnitude analysis at the beginning of Sec. III is then valid only if we replace the characteristic mass $\mathfrak{M}$, which is roughly $\mathfrak{M}(0)$, with the much smaller quantity $\mathfrak{M}(\theta)$. In particular, if we consider a Green's function all of whose external lines are zero-energy bosons with momenta of order $\mathfrak{M}(\theta)$, and allow as internal lines only zero-energy bosons with masses of order $\mathfrak{M}(\theta)$, then each loop contributes a factor e^2, a factor θ_c from Eq. (2.17), and, since the dimensionality of all graphs must be the same, also a factor $\mathfrak{M}^{-1}(\theta)$. The condition for the validity of perturbation theory is therefore that

$$e^2\left(\frac{\theta_c}{\mathfrak{M}(\theta)}\right)\ll 1,$$

or in other words

$$|\theta-\theta_c|\gg e^2\theta_c.$$

The true critical temperature is therefore expected to lie somewhere in the range

$$|\theta-\theta_c|\lesssim e^2\theta_c.$$

However, to locate it more precisely in this range, or to determine whether the phase transition is of first or second order,[24] we would need to carry out a renormalization-group analysis beyond the scope of the present article.

VII. OBSERVABILITY OF THE PHASE TRANSITION IN GAUGE THEORIES

In gauge theories, the mean value of the field ϕ_i is not an observable, because it is not gauge-invariant. Indeed, we saw in Sec. V that even the quadratically divergent and $e^2\theta^2$ terms in $\langle\phi_i\rangle$ are different in the unitarity gauge and the renormalizable ξ gauges, though they are the same in all the renormalizable ξ gauges. A suspicion may therefore arise as to the reality of the phase transition, which seems to occur when the invariance group of the lowest-order term λ_i in $\langle\phi_i\rangle$ suddenly expands or contracts. Could physical measurements really reveal a different symmetry group just above the critical temperature than just below it?

To some extent, this question already arises in gauge field theories at zero temperature. There, also, one cannot rely on the properties of Green's functions of gauge-noninvariant operators to tell us which symmetries are broken and which are not. However, at zero temperature these Green's functions have poles whose residues, the S-matrix elements, are gauge-invariant, and can be used to diagnose the symmetries of the theory. At finite temperature there is no such thing as a single collision (particles interact with the thermal background on their way into and out of any encounter), so there are no S-matrix elements as such, and this approach fails us. In particular, we cannot *directly* measure the particle masses discussed in Sec. IV.

We can, however, presumably measure the partition function in the presence of various gauge-invariant perturbing operators of the form $\phi_i\phi_i$, $F_{\alpha\mu\nu}F_\alpha{}^{\mu\nu}$, $\bar\psi\psi$, etc., and from these measurements we can infer values for the temperature Green's functions (2.11) for these operators. It is not immediately obvious, though, that we can use such Green's functions to learn about λ values and masses, because the gauge-invariant operators are necessarily at least bilinear, and so the lowest-order graphs for their Green's functions are not trees but loops, and we cannot adjust the four-momenta carried by internal lines of such loops to arbitrary values. In particular, if the typical energy or momentum carried by the internal lines of the loop is of order θ, then the insertion of a quadratically divergent subgraph of order $e^2\theta^2$ in an internal line of the loop produces two more powers of θ in the denominator, so the over-all effect is to introduce a factor e^2. From this standpoint, it does not seem that the "failure" of perturbation theory, which forced us to introduce the effective polynomial in Sec. III, is real at all.

The answer is that we must consider not the Green's functions of arbitrary gauge-invariant operators, but the Green's functions of gauge-invariant operators carrying moderate momenta ($P\approx\mathfrak{M}$ rather than $P\approx\theta$) and *zero* energy. As long as there are enough operators so that the over-all dimensionality D of the Green's function is less than 1, the leading lowest-order terms for $\theta \gg \mathfrak{M}$ will be the graphs in which all internal lines have zero energy and moderate momenta. (As discussed in Sec. III, the temperature dependence of such terms in entirely contained in the single multiplicative factor θ; all other terms are then suppressed by factors of order $\theta/\mathfrak{M}$.) These leading graphs will be sensitive functions of the masses of their internal lines, and can be used to study the symmetries of the theory. In particular, if we did not introduce our effective polynomial, then the insertion of quadratically divergent subgraphs in the zero-energy internal lines of such leading graphs would introduce factors of $e^2\theta^2/\mathfrak{M}^2$ rather than e^2, and hence would lead to a breakdown of perturbation theory for $\theta \approx \mathfrak{M}/e$. The masses and λ values which can be inferred from a study of zero-energy gauge-invariant Green's functions are therefore the ones derived from our temperature-dependent effective polynomial, and these are the ones which show the phase transition.

Of course, all we can ever measure in this way are invariant quantities, such as $\lambda_i\lambda_i$, $\mathrm{Tr}[M^2]$, $\mathrm{Tr}[\mu^2]$, etc. However, these can easily be used to tell whether the symmetries of the theory are broken. For instance, if a symmetry requires that λ_i or μ vanish, or that M^2 be proportional to the unit matrix, and our measurements reveal that $\lambda_i\lambda_i$ or $\mathrm{Tr}[\mu^2]$ is not zero, or that $\mathrm{Tr}[M^4]$ is not equal to $(\mathrm{Tr}[M^2])^2$ then we know the symmetry is broken, although of course we never find out in which direction the symmetry breaking occurs.

It is also possible to infer the existence of a phase transition by studying the partition function itself. The lowest-order graph is the "no-loop" term

$$[\ln Q]_{\text{no loop}} = -\frac{\Omega}{\theta} P_{\text{eff}}(\lambda), \tag{7.1}$$

with Ω the volume of the system, and P_{eff} evaluated at its minimum. The one-loop corrections involve λ-dependent terms which are canceled by the last term in Eq. (3.7), plus λ-independent terms of order θ^4 and $\mathfrak{M}^2\theta^2$. The latter terms are quite large, and in a sense represent a breakdown of perturbation theory which has so far been ignored because it occurs only in $\ln Q$ rather than in the Green's functions. [These terms are calculated for scalar field theories in the Appendix, and included there in P_{eff}; see Eq. (A36).] However, Eq. (7.1) correctly represents the leading terms in the part of $\ln Q$ which is λ-dependent and hence nonanalytic at the critical temperature. We will therefore rewrite Eq. (7.1) as

$$\{\ln Q\}_{\text{NA}} = -\frac{\Omega}{\theta} P_{\text{eff}}(\lambda), \tag{7.2}$$

with the subscript NA denoting the nonanalytic part. Familiar thermodynamic arguments then give the nonanalytic part of the pressure

$$\{P\}_{\text{NA}} = -P_{\text{eff}}(\lambda), \tag{7.3}$$

the entropy density

$$\{s\}_{\text{NA}} = -\frac{\partial P_{\text{eff}}(\lambda)}{\partial\theta}, \tag{7.4}$$

and the energy density

$$\{u\}_{\text{NA}} = P_{\text{eff}}(\lambda) - \theta\frac{\partial P_{\text{eff}}(\lambda)}{\partial\theta}. \tag{7.5}$$

To see how this works in practice, let us return to example 2 of Sec. IV. The effective polynomial was

$$P_{\text{eff}}(\phi) = \tfrac{1}{2}\mathfrak{M}^2(\theta)\phi_i\phi_i + \tfrac{1}{4}e^2(\phi_i\phi_i)^2,$$

where

$$\mathfrak{M}^2(\theta) = \mathfrak{M}^2(0) + \tfrac{1}{12}(n+2)e^2\theta^2 + \tfrac{1}{4}(n-1)e'^2\theta^2.$$

For $\theta < \theta_c$, the value of the polynomial at its minimum is

$$P_{\text{eff}}(\lambda) = -\frac{1}{4e^2}\mathfrak{M}^4(\theta)$$
$$= -\frac{1}{4e^2}\left[\tfrac{1}{12}(n+2)e^2 + \tfrac{1}{4}(n-1)e'^2\right]^2(\theta^2-\theta_c^2)^2.$$

On the other hand, for $\theta > \theta_c$, the minimum of the polynomial is at $\phi_i = 0$, where

$$P_{\text{eff}}(0) = 0.$$

Thus the pressure and the energy and entropy densities are continuous at $\theta = \theta_c$, but their derivatives are not. In particular, the specific heat shows a discontinuity:

$$\Delta C_V \equiv \left(\frac{\partial u}{\partial\theta}\right)_{\theta=\theta_c+\epsilon} - \left(\frac{\partial u}{\partial\theta}\right)_{\theta=\theta_c-\epsilon}$$
$$= \theta_c\left[\frac{\partial^2 P_{\text{eff}}(\lambda)}{\partial\theta^2}\right]_{\theta=\theta_c-\epsilon}$$
$$= -\frac{\theta_c^3}{e^2}\left[\tfrac{1}{12}(n+2)e^2 + \tfrac{1}{4}(n-1)e'^2\right]^2.$$

This certainly reveals the presence of a phase transition, though we need to go beyond thermodynamics to determine what symmetries are restored in this transition.

APPENDIX: OPERATOR APPROACH TO THE POTENTIAL

In order to take account of the higher-order effects discussed in Sec. VI, we need to generalize the effective polynomial by defining a potential,[8] the symmetries of whose minima will truly determine the symmetries of the theory. In this appendix, I will describe an operator approach to this problem, which incidentally leads to an interesting interpretation of the radiative corrections in field theories at zero temperature in terms of the zero-point energies of the various degrees of freedom. Unfortunately, it will be seen that this operator approach breaks down badly in gauge theories.

We will consider a system with a large but finite volume Ω, and define the spatial average of the Schrödinger representation operator $\phi(\vec{x})$ as

$$\overline{\phi}_i \equiv \frac{1}{\Omega}\int_\Omega d^3x\, \phi_i(\vec{x}). \tag{A1}$$

In order to allow us to vary the thermodynamic mean value of $\overline{\phi}_i$, we include in the Hamiltonian a perturbation $\Omega J_i \overline{\phi}_i$, with J_i a variable c-number "current." The mean value of $\overline{\phi}_i$ is then

$$\eta_i(J) \equiv \langle \overline{\phi}_i \rangle_J = \frac{\mathrm{Tr}[\overline{\phi}_i \exp\{-(1/\theta)(H+\Omega J_i\overline{\phi}_i)\}]}{\mathrm{Tr}[\exp\{-(1/\theta)(H+\Omega J_i\overline{\phi}_i)\}]}\,. \tag{A2}$$

This may be expressed in terms of the Helmholtz free energy per unit volume

$$A(J) \equiv -\frac{\theta}{\Omega}\ln \mathrm{Tr}[\exp\{-(1/\theta)(H+\Omega J_i\overline{\phi}_i)\}] \tag{A3}$$

as

$$\eta_i(J) = \frac{\partial A(J)}{\partial J_i}\,. \tag{A4}$$

(Of course, in addition to its dependence on J_i, the free energy also depends on θ and Ω and on any other parameters appearing in H.) Our potential is defined as a function of η rather than J, using a Legendre transformation to introduce the analog of the Gibbs free energy:

$$V(\eta) \equiv A(J) - J_i\eta_i\,. \tag{A5}$$

Using (A4), we may now define J as a function of η by the condition

$$\frac{\partial V(\eta)}{\partial \eta_i} = -J_i(\eta). \tag{A6}$$

In particular, the possible mean values of $\overline{\phi}_i$ when the current vanishes are given by the points where $V(\eta)$ is stationary,

$$\frac{\partial V(\eta)}{\partial \eta_i} = 0 \text{ if } J_i = 0, \tag{A7}$$

so $V(\eta)$ is a suitable potential for our purposes. In fact, the actual mean value of $\overline{\phi}_i$ for $J=0$ must be at a point where $V(\eta)$ is not only stationary but also a local *minimum*, because we can easily show that

$$\frac{\partial^2 V(\eta)}{\partial\eta_i\partial\eta_j} = \frac{\theta}{\Omega}\Delta^{-1}{}_{ij}(\eta),$$

with Δ the positive matrix

$$\Delta_{ij}(\eta) \equiv \langle(\overline{\phi}_i - \eta_i)(\overline{\phi}_j - \eta_j)\rangle_J\,.$$

The function $V(\eta)$ must be defined by analytic continuation at η values where $\partial^2V/\partial\eta_i\partial\eta_j$ is not positive, because no current J can produce such η values in a state of thermal equilibrium.

So far, this has been quite general. Let us now consider the simple scalar field theory described by the Lagrangian (2.1). The Hamiltonian here is

$$H = \int d^3x\,[\tfrac{1}{2}\pi_i\pi_i + \tfrac{1}{2}\vec{\nabla}\phi_i\cdot\vec{\nabla}\phi_i + P(\phi)], \tag{A8}$$

where π_i is the canonical conjugate to ϕ_i. As in Sec. II, we assume the parameters in $P(\phi)$ to be characterized by a typical mass $\mathfrak{M}$ and a typical dimensionless coupling $e \ll 1$, in the sense of Eq. (2.6). We will construct a perturbative expansion for $V(\eta)$ in powers of e, for temperatures θ ranging from 0 to order $\mathfrak{M}/e$.

First, in order to cancel the divergences and the attendant temperature-dependent terms in $V(\eta)$, we must again introduce a polynomial counterterm, writing

$$H = \int d^3x\,[\tfrac{1}{2}\pi_i\pi_i + \tfrac{1}{2}\vec{\nabla}\phi_i\cdot\vec{\nabla}\phi_i + P_{\text{eff}}(\phi) - \Delta P(\phi)], \tag{A9}$$

where

$$P_{\text{eff}}(\phi) = P(\phi) + \Delta P(\phi), \tag{A10}$$

with ΔP a quartic temperature-dependent polynomial to be constructed as we go along. To the order we will be studying here, it will be sufficient to treat $\Delta P(\phi)$ as a quantity of zeroth order in e for $\phi \approx \eta$; in higher order we would need to include other renormalization counterterms, including higher terms in ΔP.

Next, since we are interested in values of the variable η_i of order $\mathfrak{M}/e$ [compare Eq. (2.9)] we must shift ϕ_i, defining a new field ϕ_i' by

$$\phi_i \equiv \phi_i' + \eta_i\,. \tag{A11}$$

The Hamiltonian may now be written as a sum of terms $H^{(n)}$ of order n in e:

$$H = H^{(-2)} + H^{(-1)} + H^{(0)} + H^{(1)} + H^{(2)}, \tag{A12}$$

where

$$H^{(-2)} = \Omega P_{\text{eff}}(\eta), \tag{A13}$$

$$H^{(-1)} = \frac{\partial P_{\text{eff}}(\eta)}{\partial\eta_i}\int_\Omega d^3x\,\phi_i', \tag{A14}$$

$$H^{(0)}=\int_\Omega d^3x\left[\tfrac{1}{2}\pi_i\,\pi_i+\tfrac{1}{2}\vec\nabla\phi_i'\cdot\vec\nabla\phi_i'+\frac{1}{2}\,\frac{\partial^2 P_{\rm eff}(\eta)}{\partial\eta_i\partial\eta_j}\,\phi_i'\,\phi_j'\right]-\Omega\Delta P(\eta), \tag{A15}$$

$$H^{(1)}=\frac{1}{6}\,\frac{\partial^3 P_{\rm eff}(\eta)}{\partial\eta_i\partial\eta_j\partial\eta_k}\int_\Omega d^3x\,\phi_i'\phi_j'\phi_k'-\frac{\partial\Delta P(\eta)}{\partial\eta_i}\int_\Omega d^3x\,\phi_j', \tag{A16}$$

$$H^{(2)}=\frac{1}{24}\,\frac{\partial^4 P_{\rm eff}(\eta)}{\partial\eta_i\partial\eta_j\partial\eta_k\partial\eta_l}\int_\Omega d^3x\,\phi_i'\phi_j'\phi_k'\phi_l'-\frac{\partial^2\Delta P(\eta)}{\partial\eta_i\partial\eta_j}\int_\Omega d^3x\,\phi_i'\phi_j'. \tag{A17}$$

We shall attempt to calculate $V(\eta)$ as a sum of terms $V^{(n)}(\eta)$ of order n in e:

$$V(\eta)=\sum_n V^{(n)}(\eta), \tag{A18}$$

using the formula

$$V(\eta)=-\frac{\theta}{\Omega}\ln\mathrm{Tr}\left[\exp\left\{-\frac{1}{\theta}\left(H+J_i\int_\Omega d^3x\,\phi_i'\right)\right\}\right] \tag{A19}$$

[see Eqs. (A5) and (A3)]. The current here is given by Eq. (A6), so it may also be expanded as a sum of terms $J_i^{(n)}$ of order n in e:

$$J_i(\eta)=\sum_n J_i^{(n)}(\eta), \tag{A20}$$

with

$$J_i^{(n)}(\eta)=\frac{\partial V^{(n-1)}(\eta)}{\partial\eta_i}. \tag{A21}$$

(Recall that η is of order $\mathfrak{M}/e$.) Our calculation will therefore be recursive: Given $V^{(n)}(\eta)$, we use (A21) to calculate $J_i^{(n+1)}(\eta)$, and then insert the result back in Eq. (A19) to determine $V^{(n+1)}(\eta)$.

To start this recursive calculation, we tentatively assume that $J^{(n)}$ vanishes for $n\leq -2$, so (A19) gives the leading term in $V(\eta)$ as

$$V^{(-2)}(\eta)=-\frac{\theta}{\Omega}\ln\mathrm{Tr}\left[\exp\left(\frac{-H^{(-2)}}{\theta}\right)\right]=P_{\rm eff}(\eta). \tag{A22}$$

This can be used as in Sec. IV to study the symmetries of the theory to lowest order in e, but we do not yet know $Q_{ij}(\theta)$.

To go to the next order, we use Eq. (A21) to determine

$$J_i^{(-1)}(\eta)=-\frac{\partial P_{\rm eff}(\eta)}{\partial\eta_i}. \tag{A23}$$

Thus in first order the current term (A23) cancels the Hamiltonian (A14) in (A19), and therefore

$$V^{(-1)}(\eta)=0. \tag{A24}$$

In the next order, Eq. (A21) gives

$$J_i^{(0)}(\eta)=0, \tag{A25}$$

so from Eq. (A19)

$$V^{(0)}(\eta)=-\frac{\theta}{\Omega}\ln\mathrm{Tr}\left[\exp\left(\frac{-H^{(0)}}{\theta}\right)\right]. \tag{A26}$$

Aside from the c-number term in (A15), this is nothing but the free energy per unit volume of a noninteracting mixture of ideal Bose gases. It may therefore be written as a sum,

$$V^{(0)}(\eta)=-\Delta P(\eta)+\mathrm{Tr}[G(M^2(\eta))], \tag{A27}$$

where $G(M^2)$ is the free energy per unit volume of an ideal Bose gas with mass M,

$$G(M^2)=-\frac{\theta}{\Omega}\ln\prod_{\vec k}\sum_{N=0}^{\infty}\exp[-(N+\tfrac{1}{2})(\vec k^2+M^2)^{1/2}/\theta], \tag{A28}$$

and $M^2{}_{ij}(\eta)$ is the mass matrix in (A15)

$$M^2{}_{ij}(\eta)\equiv\frac{\partial^2 P_{\rm eff}(\eta)}{\partial\eta_i\partial\eta_j}. \tag{A29}$$

[Note that we are keeping the zero-point energy in the exponential in Eq. (A28), a point that will be of some importance later on.] Passing to the limit of infinite volume and performing the sum over N in (A28), we find

$$G(M^2)=\frac{1}{(2\pi)^3}\int d^3k\,(\tfrac{1}{2}(\vec k^2+M^2)^{1/2}+\theta\ln\{1-\exp[-(1/\theta)(\vec k^2+M^2)^{1/2}]\}). \tag{A30}$$

Before carrying out our renormalization, it is useful to compare our results so far with those of relativistic field theory. At zero temperature, Eqs. (A22), (A24), (A27), and (A30) give

$$V(\eta)\simeq P(\eta)+\frac{1}{2(2\pi)^3}\,\mathrm{Tr}\left[\int d^3k\,(\vec k^2+M^2(\eta))^{1/2}\right].$$

If $P(\eta)$ has a minimum at $\eta=\lambda$, then the true mean value of $\overline{\phi}_i$ for $J=0$ is determined by the condition

$$0=\left.\frac{\partial V(\eta)}{\partial\eta_i}\right|_{\eta=\langle\overline\phi\rangle}\simeq M^2{}_{ij}[\langle\overline\phi_j\rangle-\lambda_j]+\frac{1}{4(2\pi)^3}f_{ijk}\int d^3k\,(\vec k^2+M^2)^{-1/2}{}_{jk},$$

where

$$M^2{}_{ij}=\frac{\partial^2 P(\lambda)}{\partial\lambda_i\partial\lambda_j}\,,\quad f_{ijk}=\frac{\partial^3 P(\lambda)}{\partial\lambda_i\partial\lambda_j\partial\lambda_k}\,.$$

Therefore,

$$\langle\overline{\phi}_i\rangle=\lambda_i-\tfrac{1}{4}(2\pi)^{-3}M^{-2}{}_{il}f_{ljk}\int d^3k(\vec{k}^2+M^2)^{-1/2}{}_{jk}\,.$$

On the other hand, Eq. (3.8) shows that the mean value of the scalar field calculated by Feynman-diagram methods is

$$\begin{aligned}\langle\overline{\phi}_i\rangle&=\lambda_i-i(2\pi)^{-4}M^{-2}{}_{il}\,T_l\\&=\lambda_i+\tfrac{1}{2}i(2\pi)^{-3}M^{-2}{}_{il}f_{ljk}\\&\quad\times\int d^4k(k^2+M^2-i\epsilon)^{-1}{}_{jk}\,.\end{aligned}$$

By performing the k^0 integration, we easily see that this agrees with the result obtained by operator methods above. Thus *the one-loop diagrams simply represent the contribution of zero-point energies to the total free energy.*

Returning now to the main line of our calculation, we note that the first term of the free-energy function (A30) may be written

$$\frac{1}{2(2\pi)^3}\int d^3k(\vec{k}^2+M^2)^{1/2}=G_\infty(M^2)+\frac{1}{64\pi^2}M^4\ln M^2, \tag{A31}$$

where $G_\infty(M^2)$ is a quadratic polynomial in M^2 with divergent coefficients. Then $\mathrm{Tr}G_\infty(M^2(\phi))$ is a quartic polynomial in ϕ which satisfies all the symmetry requirements imposed on $P(\phi)$, so we can adjust the parameters in $P(\phi)$ to cancel this term, leaving over a finite-temperature-independent renormalized polynomial

$$P_{\mathrm{ren}}(\phi)\equiv P(\phi)+\mathrm{Tr}[G_\infty(M^2(\phi))]\,. \tag{A32}$$

For "moderate" temperatures, say $\theta\lesssim\mathfrak{M}$, the only counterterm we need is just the term $\mathrm{Tr}[G_\infty]$ in (A32), and we can take P_{ren} as our effective polynomial. Equations (A22), (A24), (A27), and (A30)–(A32) then give the potential as

$$V(\eta)\simeq P_{\mathrm{ren}}(\eta)+\mathrm{Tr}[G_1(M^2(\eta))]\,, \tag{A33}$$

with the function G_1 defined by

$$\begin{aligned}G_1(M^2)&\equiv G(M^2)-G_\infty(M^2)\\&=\frac{1}{64\pi^2}M^4\ln M^2\\&\quad+\frac{\theta}{(2\pi)^3}\int d^3k\ln\{1-\exp[-(1/\theta)(\vec{k}^2+M^2)^{1/2}]\}.\end{aligned} \tag{A34}$$

For $\theta\approx\mathfrak{M}$ the correction term $\mathrm{Tr}[G_1]$ is of order $\mathfrak{M}^4$, while $P_{\mathrm{ren}}(\eta)$ is of order $\mathfrak{M}^4/e^2$, so it is the temperature-independent polynomial P_{ren} that governs the pattern of broken symmetries.

On the other hand, for "high" temperatures, say $\theta\approx\mathfrak{M}/e$, the free energy (A30) becomes

$$G(M^2)\simeq G_\infty(M^2)-\tfrac{1}{90}\pi^2\theta^4+\tfrac{1}{24}\theta^2M^2 \qquad (\text{for } \theta\gg M). \tag{A35}$$

All these terms are at least as large as $P_{\mathrm{ren}}(\eta)$, so they all must be included in the effective polynomial, which now becomes

$$P_{\mathrm{eff}}(\phi)=P_{\mathrm{ren}}(\phi)-\tfrac{1}{90}B\pi^2\theta^4+\tfrac{1}{24}\theta^2\,\mathrm{Tr}[M^2(\phi)]\,, \tag{A36}$$

where B is the number of boson fields. By expanding in powers of ϕ, we find

$$\mathrm{Tr}[M^2(\phi)]=\tfrac{1}{2}f_{ijkk}\phi_i\phi_j+\text{constant}, \tag{A37}$$

so Eq. (A36) is the same as our previous result (3.26), except for a constant (the Stefan-Boltzmann term), which of course has no effect on the location of the minima of P_{eff}. We might try to improve this calculation by including the terms of order M^4 in (A36), but this improvement would be illusory; the "second-order" term $V^{(2)}(\eta)$ in the potential includes η-dependent contributions of order $e^2\theta^2\mathfrak{M}^2\approx\mathfrak{M}^4$, which are just as large as the corrections to Eq. (A36). In any case, even if we took these effects into account and correctly calculated all terms in V of order $\mathfrak{M}^4$, we still would not be able to calculate the behavior of the minimum of the potential near the critical temperatures, where one of the eigenvalues of $M^2(\eta)$ vanishes.

There is no problem in including fermions in this sort of calculation. Rather than go into such inessential complications here, let us turn immediately to the more challenging problem of a gauge theory described by the Lagrangian (2.4), leaving fermions aside for simplicity.

We need first to construct a Hamiltonian for this theory. To the best of my knowledge, this has so far been possible only in the "unitarity" gauges, defined by the condition that ϕ should have no components along the directions $\theta_\alpha\lambda$

$$\phi_i(\theta_\alpha\lambda)_i=0, \tag{A38}$$

where λ is some fixed vector. It is usual to choose λ to be a vector at which $P(\phi)$ is stationary,

$$\left.\frac{\partial P(\phi)}{\partial\phi_i}\right|_{\phi=\lambda}=0,$$

because then the directions $(\theta_\alpha\lambda)$ define the eigenvectors of the mass matrix with eigenvalue zero,

$$\left.\frac{\partial^2 P(\phi)}{\partial\phi_i\partial\phi_j}\right|_{\phi=\lambda}(\theta_\alpha\lambda)_j=0,$$

and the condition (A38) just amounts to the exclusion of Goldstone bosons from the theory. However, we shall leave λ arbitrary here.

The dynamical variables in the gauge (A38) are the spatial components $\vec{A}_\alpha$ of the gauge fields and the scalar field components

$$\phi_a = n_{ai}\phi_i\,, \tag{A39}$$

where n_{ai} forms a complete orthonormal set of vectors orthogonal to all $\theta_\alpha\lambda$:

$$n_{ai}n_{bi} = \delta_{ab}\,, \quad n_{ai}(\theta_\alpha\lambda)_i = 0. \tag{A40}$$

The Hamiltonian in this gauge is given by the unpleasant-looking formula[7]

$$\begin{aligned} H = \int d^3x \{ & \tfrac{1}{2}\omega^{-1}{}_{\alpha\beta}(\phi)[\vec{\nabla}\cdot\vec{P}_\alpha - C_{\alpha\gamma\delta}\vec{P}_\delta\cdot\vec{A}_\gamma + i(\theta_\alpha)_{ab}\pi_a\phi_b][\vec{\nabla}\cdot\vec{P}_\beta - C_{\beta\epsilon\zeta}\vec{P}_\zeta\cdot\vec{A}_\epsilon + i(\theta_\beta)_{cd}\pi_c\phi_d] \\ & + \tfrac{1}{2}\vec{P}_\alpha\cdot\vec{P}_\alpha + \tfrac{1}{2}\pi_a\pi_a + \tfrac{1}{2}(\vec{\nabla}\times\vec{A}_\alpha - C_{\alpha\beta\gamma}\vec{A}_\beta\times\vec{A}_\gamma)\cdot(\vec{\nabla}\times\vec{A}_\alpha - C_{\alpha\delta\epsilon}\vec{A}_\delta\times\vec{A}_\epsilon) \\ & + \tfrac{1}{2}\vec{\nabla}\phi_a\cdot\vec{\nabla}\phi_a + i\vec{A}_\alpha\cdot\vec{\nabla}\phi_a(\theta_\alpha)_{ab}\phi_b + \tfrac{1}{2}\vec{A}_\alpha\cdot\vec{A}_\beta[\omega_{\alpha\beta}(\phi) - (\theta_\alpha)_{ac}(\theta_\beta)_{bc}\phi_a\phi_b] + P(\phi)\}\,, \end{aligned} \tag{A41}$$

where $\vec{P}_\alpha$ and π_a are the "momenta" canonically conjugate to A_α and ϕ_α, and

$$(\theta_\alpha)_{ab} \equiv n_{ai}n_{bj}(\theta_\alpha)_{ij}\,, \tag{A42}$$

$$\omega_{\beta\delta}(\phi) \equiv (\theta_\beta\phi)_i(\theta_\alpha\lambda)_i(\nu^{-2})_{\alpha\gamma}(\theta_\gamma\lambda)_j(\theta_\delta\phi)_j\,, \tag{A43}$$

$$\nu^{-2}{}_{\alpha\beta} \equiv -(\theta_\alpha\lambda)_i(\theta_\beta\lambda)_i\,. \tag{A44}$$

At this point we face a problem. If we identify λ with the argument η of the potential V, then changes in λ will change V not only directly, through changes in the current $J(\lambda)$, but *also through changes in the choice of gauge*. The λ dependence of the Hamiltonian would invalidate the general formalism used here; indeed we already know[25] that the vacuum expectation value of ϕ_i at zero temperature in the one-loop approximation is not given by the minimum of *any* potential depending on the single variable λ. To avoid this problem, we fix λ_i and our choice of gauge once and for all,[26] and use as the independent variables in the potential the thermodynamic mean values η_a of the fields ϕ_a in this gauge.

With this understanding, our general formalism becomes applicable again. We define a shifted scalar field

$$\phi'_a = \phi_a - \eta_a\,, \tag{A45}$$

with η_a of order $\mathfrak{M}/e$. The Hamiltonian can then be written as a sum of terms $H^{(n)}$ of order e^n:

$$H = H^{(-2)} + H^{(-1)} + H^{(0)} + H^{(1)} + H^{(2)}, \tag{A46}$$

where, up to zeroth order,

$$H^{(-2)} = \Omega P_{\text{eff}}(\eta), \tag{A47}$$

$$H^{(-1)} = \frac{\partial P_{\text{eff}}(\eta)}{\partial\eta_i}\int d^3x\,\phi'_i\,, \tag{A48}$$

$$\begin{aligned} H^{(0)} = \int d^3x \{ & \tfrac{1}{2}\omega^{-1}{}_{\alpha\beta}(\eta)[\vec{\nabla}\cdot\vec{P}_\alpha + i(\theta_\alpha)_{ab}\pi_a\eta_b][\vec{\nabla}\cdot\vec{P}_\beta + i(\theta_\beta)_{cd}\pi_c\eta_d] + \tfrac{1}{2}\vec{P}_\alpha\cdot\vec{P}_\alpha + \tfrac{1}{2}\pi_a\pi_a + \tfrac{1}{2}(\vec{\nabla}\times\vec{A}_\alpha)\cdot(\vec{\nabla}\times\vec{A}_\alpha) \\ & + \tfrac{1}{2}\vec{\nabla}\phi'_a\cdot\vec{\nabla}\phi'_a + i\vec{A}_\alpha\cdot\vec{\nabla}\phi'_a(\theta_\alpha)_{ab}\eta_b + \tfrac{1}{2}\mu^2{}_{\alpha\beta}(\eta)\vec{A}_\alpha\cdot\vec{A}_\beta + \tfrac{1}{2}M^2{}_{ab}(\eta)\phi'_a\phi'_b\} - \Omega\Delta P(\eta). \end{aligned} \tag{A49}$$

with

$$P_{\text{eff}}(\phi) = P(\phi) + \Delta P(\phi), \tag{A50}$$

$$\mu^2{}_{\alpha\beta}(\eta) \equiv \omega_{\alpha\beta}(\eta) - (\theta_\alpha)_{ac}(\theta_\beta)_{bc}\eta_a\eta_b\,, \tag{A51}$$

$$M^2{}_{ab}(\eta) = \frac{\partial^2 P_{\text{eff}}(\eta)}{\partial\eta_a\partial\eta_b}\,. \tag{A52}$$

[In evaluating $\omega(\eta)$ and $P(\eta)$ and its derivatives, we set η_i equal to $n_{ai}\eta_i$.] For future reference, we note that since the n_{ai} span the space orthogonal to the $\theta_\alpha\lambda$, they satisfy the sum rule

$$n_{ai}n_{aj} = \delta_{ij} + (\theta_\alpha\lambda)_i\,\nu^{-2}{}_{\alpha\beta}(\theta_\beta\lambda)_j$$

[see (A44)]. Contracting this with $(\theta_\gamma\eta)_i(\theta_\delta\eta)_j$, we find

$$(\theta_\alpha)_{ab}(\theta_\delta)_{ac}\eta_b\eta_c = (\theta_\gamma\eta)_i(\theta_\delta\eta)_i + \omega_{\gamma\delta}(\eta),$$

and therefore

$$\mu^2{}_{\alpha\beta}(\eta) = -(\theta_\alpha\eta)_i(\theta_\beta\eta)_i\,.$$

Following precisely the same reasoning as for the scalar field theory discussed above, we find here for the potential up to zeroth order in e:

$$V(\eta) \simeq P_{\text{eff}}(\eta) - (\theta/\Omega)\ln\mathrm{Tr}[\exp(-H^{(0)}/\theta)]\,. \tag{A53}$$

The calculation of the second term is now not entirely straightforward, because $H^{(0)}$ is not in the familiar form of a free-particle Hamiltonian. In order to bring it to this form, we must first perform a canonical transformation to a new set of canonical variables $\vec{a}_\alpha$, Φ_a and their conjugates $\vec{p}_\alpha$, Π_a:

$$\vec{a}_\alpha = \vec{A}_\alpha + i(\mu^{-2}(\eta))_{\alpha\beta}(\theta_\beta)_{ab}\eta_b\vec{\nabla}\phi'_a\,, \tag{A54}$$

$$\vec{p}_\alpha = \vec{P}_\alpha, \tag{A55}$$

$$\Phi_a = (S^{-1}(\eta))_{ab}\phi_b', \tag{A56}$$

$$\Pi_a = (S(\eta))_{ab}[\pi_b + i(\mu^{-2}(\eta))_{\alpha\beta}(\theta_\beta)_{bc}\eta_c\vec{\nabla}\cdot\vec{P}_\alpha], \tag{A57}$$

where $S(\eta)$ is any matrix such that

$$S_{ba}(\eta)S_{ca}(\eta) = \delta_{bc} - (\theta_\alpha)_{ba}\eta_a(\theta_\beta)_{cd}\eta_d\omega^{-1}{}_{\alpha\beta}(\eta). \tag{A58}$$

It is straightforward, though tedious, to check that this *is* a canonical transformation, and that it brings $H^{(0)}$ to the diagonalized form:

$$\begin{aligned} H^{(0)} = \int d^3x\,[&\tfrac{1}{2}\mu^{-2}{}_{\alpha\beta}(\eta)(\vec{\nabla}\cdot\vec{p}_\alpha)(\vec{\nabla}\cdot\vec{p}_\beta) + \tfrac{1}{2}\vec{p}_\alpha\cdot\vec{p}_\beta \\ &+ \tfrac{1}{2}\Pi_a\Pi_a + \tfrac{1}{2}(\vec{\nabla}\times\vec{a}_\alpha)\cdot(\vec{\nabla}\times\vec{a}_\alpha) \\ &+ \tfrac{1}{2}\vec{\nabla}\Phi_a\cdot\vec{\nabla}\Phi_a + \tfrac{1}{2}\mu^2{}_{\alpha\beta}(\eta)\,\vec{a}_\alpha\cdot\vec{a}_\beta \\ &+ \tfrac{1}{2}\tilde{M}^2{}_{ab}(\eta)\,\Phi_a\Phi_b] - \Omega\Delta P(\eta), \end{aligned} \tag{A59}$$

where

$$\tilde{M}^2{}_{cd}(\eta) \equiv \frac{\partial^2 P_{\text{eff}}(\eta)}{\partial\eta_a\partial\eta_b}(S(\eta))_{ac}(S(\eta))_{bd}. \tag{A60}$$

We can now immediately write down the free energy:

$$\begin{aligned} V(\eta) \simeq P_{\text{eff}}(\eta) + 3\,\mathrm{Tr}[G(\mu^2(\eta))] + \mathrm{Tr}[G(\tilde{M}^2(\eta))] \\ - \Delta P(\eta), \end{aligned} \tag{A61}$$

where G is the function (A30).

At this point our calculation breaks down. The divergent part of the function $G(M^2)$ is a quadratic polynomial in M^2, so if $\tilde{M}^2(\eta)$ and $\mu^2(\eta)$ are quadratic polynomials in η the divergent parts of $\mathrm{Tr}[G(\mu^2(\eta))]$ and $\mathrm{Tr}[G(\tilde{M}^2(\eta))]$ will be quartic polynomials in η, and hence can be removed by the counterterm ΔP. However, although $\mu^2(\eta)$ is a quadratic polynomial in η, $\tilde{M}^2(\eta)$ is not [because of the matrices $S(\eta)$ in (A60)], and *therefore the infinite part of* $\mathrm{Tr}[G(\tilde{M}^2(\eta))]$ *is not a quartic polynomial in* η *and cannot be removed by renormalization.*

The reason for this difficulty is not hard to find. In general, the only reason that we would have to believe that renormalization should work in a calculation based on the unitarity gauge is that the results must be equivalent to those obtained in one of the renormalizable gauges, such as the ξ gauges used here in Secs. II–IV. However, in defining the potential, we have perturbed the Hamiltonian by a term linear in the scalar fields in the unitarity gauge, and these scalar fields are nonpolynomial functions of the scalar fields of the renormalizable gauge. [For instance, in an O(2) gauge theory with one 2-vector scalar field multiplet, the single U-gauge scalar field ϕ is related to the two R-gauge scalar fields ϕ_1, ϕ_2 by $\phi = (\phi_1{}^2 + \phi_2{}^2)^{1/2}$.] From the point of view of the R gauge, the perturbed Hamiltonian corresponds to a manifestly nonrenormalizable interaction, so of course renormalization theory does not work, whether we use the R gauge or the U gauge. [Actually the divergent part of $\mathrm{Tr}G(\tilde{M}^2(\eta))$ is a quartic polynomial in η in the special case where ϕ_i furnishes a representation transitive on the sphere, because in this case the matrix S_{ba} is the unit matrix. However, as discussed in Sec. V, even though we can calculate an effective potential in unitarity gauge in such simple theories, it is not particularly useful to do so.]

This analysis suggests two possible directions for construction of a suitable potential in gauge theories:

(a) Instead of perturbing the Hamiltonian by a linear function of the U-gauge scalar fields, we could use a quadratic or higher-order function which can be written as a gauge-invariant polynomial function of the R-gauge scalar fields. A preliminary analysis indicates that this would cure the nonrenormalizability we have found here, but the formalism needs further development.

(b) We could give up the operator methods altogether, and return to a diagrammatic R-gauge analysis along the lines of Secs. II–V. The difficulty here is in defining a suitable potential; it is known[25] that even at zero temperature, the tadpoles and other boson Green's functions with zero external four-momenta are not given by derivatives of any potential in any of the ξ gauges except the Landau ($\xi=\infty$) gauge. The solution would be to use either the Landau gauge or one of the ξ gauges defined by a fixed vector λ_i different from the argument η_i of the potential,[26] as in this appendix. The potential defined in this way would be gauge-dependent, because the perturbation $J_i\bar{\phi}_i$ is a gauge-dependent operator. However, the values of the potential at its local minima or maxima are gauge-independent, because J_i vanishes at these stationary points. Hence if there is a minimum of $V(\eta)$ with a value lower than $V(0)$ in one gauge, then there will be such a minimum in any gauge (although its position will generally be different), and the symmetries will definitely be broken.[27]

At any rate, the definition of a suitable potential is only a first step toward a solution of the real problem, the summation to all orders of the infrared divergences at the critical temperature.

ACKNOWLEDGMENTS

I am grateful for frequent valuable discussions throughout the course of this work with C. Bernard, S. Coleman, R. Jackiw, and P. Martin. I also wish to thank E. Brezin, A. Duncan, B. Harrington, J. R. Schrieffer, B. Vul, and K. Wilson for interesting conversations on this and related matters.

---

*Work supported in part by the National Science Foundation under Grant No. GP40397X.

[1]Y. Nambu and G. Jona-Lasinio, Phys. Rev. 122, 345 (1961). Also see W. Heisenberg, Z. Naturforsch. 14, 441 (1959).

[2]D. A. Kirzhnits and A. D. Linde, Phys. Lett. 42B, 471 (1972). Also see D. A. Kirzhnits, Zh. Eksp. Teor. Fiz. Pis'ma Red 15, 745 (1972) [JETP Lett. 15, 529 (1972)].

[3]The critical temperatures in various field theories have also been calculated, using rather different methods, by L. Dolan and R. Jackiw, this issue, Phys. Rev. D 9, 3320 (1974).

[4]For a general introduction and references to the original literature, see A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971), Chap. 7.

[5]S. Weinberg, Phys. Rev. Lett. 19, 1264 (1967). Also see A. Salam, in *Elementary Particle Theory: Relativistic Groups and Analyticity (Nobel Symposium No. 8)*, edited by N. Svartholm (Almqvist and Wiksells, Stockholm, 1968), p. 367.

[6]See, e.g., F. Jona and G. Shirane, *Ferroelectric Crystals* (Pergamon, Oxford, 1962), p. 280. Rochelle salt has a lower Curie point at −18 °C, below which the crystal is orthorhombic, and above which it is monoclinic.

[7]S. Weinberg, Phys. Rev. D 7, 1068 (1973).

[8]Potentials have been used for the study of broken symmetries in quantum field theories at zero temperature by J. Goldstone, A. Salam, and S. Weinberg [Phys. Rev. 127, 965 (1962)], G. Jona-Lasinio [Nuovo Cimento 34, 1790 (1964)], S. Coleman and E. Weinberg [Phys. Rev. D 7, 1888 (1973)], S. Weinberg (Ref. 25); J. M. Cornwall and R. E. Norton [Phys. Rev. D 8, 3338 (1973)], and others. Functional methods for the construction of potentials in gauge theories at zero temperature have recently been developed by R. Jackiw [Phys. Rev. D 9, 1686 (1974)] and have been applied to the problem of gauge symmetry breaking by L. Dolan and R. Jackiw [Phys. Rev. D 9, 2904 (1974)]. For a review of some of these developments, see S. Coleman, Lectures given at the 1973 International Summer School of Physics "Ettore Majorana" (unpublished).

[9]J. M. Cornwall, D. N. Levin, and G. Tiktopoulos, Phys. Rev. Lett. 30, 1268 (1973); 31, 572(E) (1973); C. H. Llewellyn Smith, Phys. Lett. 46B, 233 (1973); S. Joglekan, SUNY-Stony Brook report (unpublished); J. Sucher and C. H. Woo, Phys. Rev. D 8, 2721 (1973). Also see S. Weinberg, Phys. Rev. Lett. 27, 1688 (1971); A. I. Vainshtein and I. B. Khriplovich. Yad. Fiz. 13, 198 (1971) [Sov. J. Nucl. Phys. 13, 111 (1971)]; J. Bell Nucl. Phys. B60, 427 (1973).

[10]R. Hagedorn, Astron. Astrophys. 5, 184 (1970); K. Huang and S. Weinberg, Phys. Rev. Lett. 25, 895 (1971); S. Frautschi, Phys. Rev. D 3, 2821 (1971); etc.

[11]See, e.g., V. A. Belinskii, E. M. Lifschitz, and I. M. Khalatnikov, Usp. Fiz. Nauk. 102, 463 (1970) [Sov. Phys.–Usp. 13, 745 (1971)]; Zh. Eksp. Teor. Fiz. 62, 1606 (1972) [Sov. Phys.—JETP 35, 838 (1972)]; V. A. Belinskii and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. 63, 1121 (1972) [Sov. Phys.—JETP 36, 591 (1973).]

[12]These remarks are based on a conversation with J. R. Schrieffer.

[13]For a general review, see S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972), Chap. XV, Secs. V and VI.

[14]A. A. Abrikosov, L. P. Gorkov, and I. E. Dzvaloskinskii, Zh. Eksp. Teor. Fiz. 36, 900 (1959) [Sov. Phys.—JETP 9, 636 (1959)]; E. S. Fradkin, Zh. Eksp. Teor. Fiz. 36, 1286 (1959) [Sov. Phys.—JETP 9, 912 (1959)]; P. C. Martin and J. Schwinger, Phys. Rev. 115, 1342 (1959).

[15]K. Fujikawa, B. W. Lee, and A. I. Sanda, Phys. Rev. D 6, 2923 (1972). Also see Y.-P. Yao, Phys. Rev. D 7, 1647 (1973); G. 't Hooft and M. Veltman, Nucl. Phys. B50, 318 (1972); S. Weinberg, Ref. 25, Appendix A.

[16]This has been checked in detail by C. Bernard (private communication).

[17]C. Bernard, this issue, Phys. Rev. D 9, 3312 (1974); L. Dolan and R. Jackiw, Ref. 3.

[18]G. H. Hardy, *A Course of Pure Mathematics* (Cambridge Univ. Press, Cambridge, 1949), p. 330, Ex. 3.

[19]An example of a positive-definite quartic form $f_{ijkl} \times \phi_i \phi_j \phi_k \phi_l$ for which $f_{ijkk}\phi_i\phi_j$ is not a positive quadratic form was originally suggested to me by S. Coleman (private communication). Example 3 of Sec. IV is based on the quartic form suggested by Coleman.

[20]This result was also found by L. Dolan and R. Jackiw, Ref. 3, for the case $n = 2$.

[21]These are the simple theories considered in Sec. VII of Ref. 7.

[22]The effect of a background of black-body radiation on the infrared divergences in the quantum theory of photons or gravitons was considered by S. Weinberg, in *Contemporary Physics* (International Atomic Energy Agency, Vienna, 1969), Vol. I, p. 560.

[23]For a review, see K. G. Wilson and J. Kogut (unpublished); S.-K. Ma, Rev. Mod. Phys. 45, 589 (1973).

[24]The examples in Sec. IV exhibit at least an approximate second-order transition, in the sense that any discontinuity in the "order parameter" λ can only be due to effects of higher order in e. However, it remains an open possibility that the phase transition may be *weakly* of first order, i.e., that λ exhibits a small discontinuity at the true critical temperature. It should be noted that the presence of weakly coupled gauge fields in superconductors and certain liquid

crystals leads to phase transitions that are weakly of first order, as recently shown by B. I. Halperin, T. C. Lubensky, and S.-K. Ma, Phys. Rev. Lett. 32, 292 (1974).

[25]S. Weinberg, Phys. Rev. D 7, 2887 (1973).

[26]This was suggested to me by R. Jackiw (private communication).

[27]This remark is due to C. Bernard (private communication).