Publications and Research   City College of New York

7-15-1983

Order of Phase Transitions in a Field Theory

V. Parameswaran Nair
City College of New York

Recommended Citation

'Order of Phase Transitions in a Field Theory,' V.P. Nair, Phys. Rev. D28, 409 (1983).

Order of phase transitions in a field theory

V. P. Nair

Physics Department, Syracuse University, Syracuse, New York 13210

(Received 16 December 1982)

The relation between the gauge or global nature of a symmetry being broken and the order of the associated phase transition is clarified. First-order phase transitions in gauge theories are related to the asymptotic freedom of the Φ^3 interaction in six dimensions. Spontaneous violation of lepton number in the standard electroweak theory provides a nontrivial illustration of our results.

In this paper we propose a rule about the order of phase transitions in a field theory. Consider a four-dimensional field theory with a continuous symmetry group G which is spontaneously broken down to a subgroup H. If we have a gauged symmetry so that we have gauge bosons in G/H which acquire mass purely from the symmetry breakdown, the transition will be first order. If we have a global symmetry so that G/H corresponds to massless Goldstone bosons, the transition will be of second order. The theory should satisfy certain conditions in order that this rule be valid. We shall discuss these conditions below.

The precise statement of the rule does not seem to have appeared in the literature before, although it has been noticed in several contexts that gauge bosons tend to make the phase transition first order.[1−3] Our result is, in a certain sense, a clarification and refinement of some of the observations in these papers. That this refinement has some nontrivial content can be illustrated by an example. We consider the standard electroweak theory based on $SU(2)_L \times U(1)$. Consider spontaneous violation of lepton number which gives mass to the neutrinos.[4] Without the introduction of right-handed neutrinos this can be accomplished in a minimal way by a Higgs field H which transforms as a triplet under $SU(2)_L$:

$$H = \begin{pmatrix} \dfrac{h^+}{\sqrt{2}} & h^{++} \\ h^0 & -\dfrac{h^+}{\sqrt{2}} \end{pmatrix} . \tag{1}$$

We have indicated the charge assignments also. H carries lepton number -2 and couples to the standard lepton doublet l_L by

$$\hat{L}_I = -\alpha(\overline{l_L^c} H l_L + \text{H.c.}) . \tag{2}$$

The neutral component of H, viz., h^0 acquires a vacuum expectation value spontaneously breaking lepton number and giving Majorana masses to the neutrinos. Is this a second-order transition? Although lepton number is a global symmetry, the gauge bosons of $SU(2)_L \times U(1)$ couple to H and it is not *a priori* obvious if the transition remains second order. The rule we have stated assures us that it is of second order. [For completeness, we note that lepton number in the standard model has anomalies. More precisely, our remarks apply to $(B-L)$ which is the good global symmetry of the standard model.]

We shall attempt a constructive proof of our rule. Some of the arguments are given in a brief form in the appendix of Ref. 5. To include finite-temperature effects we work in Euclidean space with periodicity of bosonic fields (antiperiodicity of fermionic fields) along the imaginary time of period $\beta = 1/T$, where T is the temperature.[6] The propagator of a particle is then of the form $\sum_n 1/(\omega_n^2 + k^2)$, where $\omega_n = 2\pi n T$ for bosons and $\omega_n = 2\pi(n + \frac{1}{2})T$ for fermions. The $n=0$ mode of bosons has no infrared cutoff in k space leading to possible divergences. The accumulation of such divergences can affect the critical-point behavior. T provides an infrared cutoff for all fermionic modes as well as the $n \neq 0$ modes of bosons. To analyze the theory near the critical point, one can make a natural separation of the $n=0$ modes of the bosons. The other modes, viz., fermions and $n \neq 0$ bosons can be "integrated out." This means that one evaluates all diagrams with $n=0$ modes of bosons as external lines and the other modes as internal lines. This gives us an effective Lagrangian involving only the $n=0$ bosonic modes. The coupling constants of this Lagrangian have temperature-dependent corrections from the other modes. Since the $n=0$ bosonic modes do not depend on the Euclidean time variable we have a three-dimensional field theory. This technique of reduction to three dimensions is borrowed from Ref. 3 where more details can be found.

For a global symmetry the behavior near the critical point is modeled by a Lagrangian of the form

$$\mathcal{L} = -\frac{1}{T}\{|\nabla\Phi|^2 + \lambda[\Phi^*\Phi - V^2(T)^2]\} , \tag{3}$$

where

$$V^2 = a(T_{c_1}^2 - T^2), \quad a > 0 . \tag{4}$$

[We have suppressed possible internal-symmetry indices on Φ. Since most of the calculations leading to Eqs. (3) and (4) are standard we do not give details; see Ref. 7.] The temperature-dependent corrections to V^2 come from modes other than the $n=0$ mode of Φ. For $T > T_{c_1}$, Φ has positive mass squared and we have $\langle\Phi\rangle = 0$ (symmetric phase). For $T < T_{c_1}$ we have symmetry breaking, $\langle\Phi\rangle \neq 0$. At T_{c_1}, $\langle\Phi\rangle = 0$, we have a second-order phase transition. At T_{c_1}, Φ is a massless field.

The Lagrangian gives the lowest-order scaling behavior near the critical point. Higher-order corrections due to the self-interactions of the $n=0$ mode Φ can be treated in an ϵ expansion as in standard three-dimensional Φ^4 theory.[7,8]

We shall now contrast this with a gauged version of the model. The $n=0$ modes of the gauge bosons persist into the three-dimensional description. Of the four A_μ, A_0 acquires a mass (the plasmon effect) and decouples from critical-point behavior. We thus have a three-dimensional vector field $\vec{A}$. The effective Lagrangian is thus

$$\mathcal{L} = -\frac{1}{T}[|D\Phi|^2 + \lambda(\Phi^*\Phi - V^2)^2] - \frac{1}{2T}\text{Tr}(F_{jk}F_{jk}) . \tag{5}$$

$D\Phi$ denotes the covariant derivative of Φ. We shall now estimate the contribution of A to the effective potential. To one-loop order, this is equivalent to calculating the zero-point-energy contribution of the gauge bosons in an arbitrary background field Φ. We can neglect space variation of Φ since it does not affect critical-point behavior except as perturbative corrections. The gauge bosons in G/H acquire a mass^2 $M^2(\Phi) = \text{const} \times e^2\Phi^*\Phi$, where H is the little group of the background field Φ. Their contribution to the effective Lagrangian is

$$\Delta\mathscr{L} = \sum \int \tfrac{1}{2}\ln(k^2 + M^2)\frac{d^3k}{(2\pi)^3} . \tag{6}$$

The summation is over all helicity states of all bosons on G/H:

$$\Delta L = A + BM^2(\Phi) - \frac{(M^2)^{3/2}T}{12\pi} \tag{7}$$

$$= A + e^2B'\Phi^*\Phi - g(\Phi^*\Phi)^{3/2} , \tag{8}$$

where A, B, B' are divergent constants. g is proportional to $|e|^3T$, where e is the gauge coupling constant. A is absorbed by vacuum energy renormalization, and B by zero-temperature mass renormalization of Φ, i.e., into the definition of a and T_{c_1} of Eq. (4). The effective Lagrangian is thus of the form

$$\mathscr{L} = -\frac{1}{T}[|\partial\Phi|^2 + U(\Phi)] , \tag{9}$$

$$U(\Phi) = \lambda(\Phi^*\Phi - V^2)^2 - g(\Phi^*\Phi)^{3/2} . \tag{10}$$

The effect of the $n = 0$ modes of bosons is to give a $|\Phi|^3$ term, $|\Phi| = (\sum\Phi^*\Phi)^{1/2}$, which can also drive the symmetry breaking. For $T > T_{c_1}$ when $V^2 < 0$, $U(\Phi)$ increases near $\Phi = 0$. For larger values of Φ, the $|\Phi|^3$ term decreases $U(\Phi)$, and for still larger values the $\lambda|\Phi|^4$ term dominates giving increasing $U(\Phi)$. Thus $U(\Phi)$ has two local minima $\Phi = 0$ and $\Phi = \Phi_1 \neq 0$ separated by a potential barrier. The transition is of the first order (see Fig. 1). The critical temperature T_c is the point where the two minima have equal energies, i.e., $V(\Phi = 0, T_c) = V(\Phi = \Phi_1, T_c)$. For $T > T_c$, $\Phi = 0$ is the global minimum; for $T < T_c$, $\Phi = \Phi_1$ is the global minimum. One has $T_c > T_{c_1}$. Above a temperature
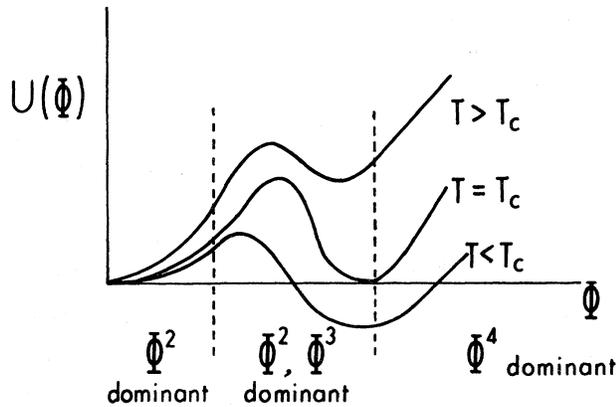

FIG. 1. The potential $U(\Phi)$ of Eq. (10) at different temperatures.

$T_{c_2} > T_c$, the $\Phi = \Phi_1$ minimum disappears altogether and full symmetry restoration occurs. Also at T_c, $\langle\Phi\rangle \neq 0$.

From Eqs. (7) and (8) it is clear that the gauge bosons which acquire a (mass)^2 proportional to $\Phi^*\Phi$ are making the transition first order. The significant point is that the gauge bosons associated with the symmetry being broken can acquire mass only from this breakdown. If there are gauge bosons which can acquire mass from other sources they do not give a $|\Phi|^3$ term. Consider the breakdown of lepton number again. The triplet Higgs field H gives a correction to the W-boson mass, $M_W^2 = e^2(\langle\Phi_D\rangle^2 + \langle H\rangle^2)$, where Φ_D is the doublet Higgs field of standard electroweak theory. Since $\langle\Phi_D\rangle \neq 0$ at T_{c_1}, $(M^2)^{3/2}$ can be expanded in powers of $H^\dagger H$, and we do not get an $|H|^3$ term which makes the transition first order. There may be other gauge bosons like those in H which remain massless above and below the transition. They also do not give cubic terms. On the other hand, for gauged lepton number the corresponding boson would acquire a (mass)$^2 \sim H^\dagger H$ giving a first-order transition.

One can construct theories where other bosons (scalars) are massless above the transition but acquire a (mass)$^2 \sim \Phi^*\Phi$ due to the transition. They can then make the transition first order. Since there is no requirement of such a structure based on symmetry-breaking patterns, such theories will require unnatural fine tuning of parameters. We shall exclude such theories by the following "naturalness" requirement: The only bosons which remain massless above the transition and acquire (mass)$^2 \sim \Phi^*\Phi$ due to the transition are those that are required to do so by symmetry-breaking patterns (viz., gauge bosons, if any, in G/H). Following is a comment about the stability of our result to higher-order A-loop corrections. The expression in Eq. (6) is exact for an Abelian vector potential since A occurs only quadratically in the Lagrangian. For the non-Abelian case there will be cubic and quartic interaction terms. These cannot contribute at one-loop order since we do not allow external A lines in the effective potential. The question then arises: Can higher-order corrections eliminate the cubic interaction which is crucial to our arguments? The question is made more acute by the fact that the gauge bosons in H which remain massless can lead to infrared divergences. The question is best understood from the point of view of the spectrum of the theory. The fact that the background field Φ breaks the symmetry from G to H is reexpressed as the statement that the vector bosons in G/H have a mass proportional to $\Phi^*\Phi$. The existence of massive states in the spectrum is a nonperturbative statement. The zero-point energy of these states will give the cubic interaction. The higher-loop effects will, of course, renormalize the cubic coupling constant and can be analyzed using the renormalization-group (RG) equations. To the order we shall compute (one-loop order for the RG), these effects can be neglected. In particular, they are negligible in the large-N limit we consider below.

We now turn to the question of Φ loops. Do higher-order corrections make g in Eq. (10) zero or negative so that the transition is restored to second order? This can be analyzed with use of the renormalization-group equations. A fixed point for the RG equations gives pure scaling behavior or a second-order phase transition. The critical dimension of the $|\Phi|^3$ interaction is six. The behavior of the three-dimensional theory can be explored within the ϵ expansion ($\epsilon = d_c - d = 3$). The $|\Phi|^3$ theory in six dimensions is asymptotically free. The RG equations in reverse

then imply the infrared slavery of g, i.e., we can have no fixed point and the transition is of the first order. The asymptotic freedom of g can be destroyed by the $\lambda(\Phi^*\Phi)^2$ interaction if there are too many Higgs multiplets. We shall therefore require as a second premise for our rule that the number of Higgs multiplets is small enough so as not to destroy the infrared slavery of the $g|\Phi|^3$ interaction. (The statement about infrared slavery, in contrast to asymptotic freedom, goes beyond perturbation theory.)

We can estimate the number of Higgs multiplets which destroys asymptotic freedom. With M copies of Φ, which is considered as an N-vector, we can write

$$\mathscr{L}_I = -g\,(\Phi_i^{\alpha}\Phi_i^{\alpha})^{3/2} + \lambda_1(\Phi_i^{\alpha}\Phi_i^{\alpha})^2 + \lambda_2(\Phi_i^{\alpha}\Phi_i^{\beta})^2 \qquad (11)$$

($i = 1,...,N$, $\alpha = 1,...,M$). We shall interpret the nonpolynomial interaction as $-\frac{3}{2}g\,\sigma(\Phi_i^{\alpha}\Phi_i^{\alpha})$, where σ is the composite field $(\Phi_i^{\alpha}\Phi_i^{\alpha})^{1/2}$. Notice that we can write the interaction as $g[\sigma^3/2 - \frac{3}{2}\sigma(\Phi_i^{\alpha}\Phi_i^{\alpha})]$, introducing σ as a Lagrangian multiplier field. The only ambiguity in doing perturbation theory will be with σ loops since σ does not have a propagator at this stage (although it will be generated in higher orders). Φ loops give a factor of N or M due to summation on the index i or α, but σ loops do not have such a factor. Thus we can ignore σ loops. The resulting error in perturbation theory would be negligible in the large-(N,M) limit. The one-loop vertex corrections are shown in Fig. 2. The loops are all Φ loops; σ's have been pushed to external lines wherever the cubic interaction vertex occurs. The RG equations are

$$\frac{d\lambda_1}{dt} = \frac{\lambda_1}{2} - \frac{\lambda_1^2 NM}{16\pi^3} - \frac{\lambda_1\lambda_2(N+M)}{8\pi^3}$$
$$+ \frac{9g^2}{32\pi^3}[NM\lambda_1 + (N+M)\lambda_2] \ ,$$
$$\frac{d\lambda_2}{dt} = \frac{\lambda_2}{2} - \frac{\lambda_2^2(N+M)}{16\pi^3} \ , \qquad (12)$$
$$\frac{dg}{dt} = \frac{3}{4}g + \frac{9g^3 NM}{128\pi^3} - \frac{3\lambda_1 gNM}{32\pi^3} - \frac{3\lambda_2 g(N+M)}{32\pi^3} \ .$$

$t \rightarrow \infty$ is the infrared limit. The infrared slavery (asymptotic freedom) of g is destroyed with λ_1, λ_2 developing stable fixed points for $NM \geq 42$. This number is comfortingly large.

The RG equations (12) are strictly legitimate only in $6-\epsilon$ dimensions. However, in the spirit of the ϵ expansion, we can use them to analyze our three-dimensional theory.

We now restate the rule with the caveats which have emerged from the discussion: Given (a) "naturalness" of the theory, i.e., the only bosons which acquire mass purely from
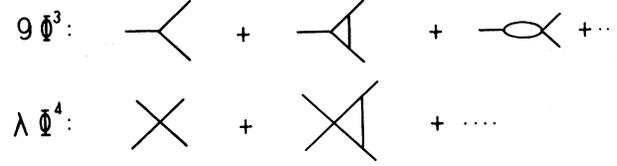
$9\,\Phi^3$:

FIG. 2. Radiative corrections to Φ^3 and Φ^4 interactions.

the symmetry breakdown are those required to do so by symmetry principles, and (b) superfluous replication of Higgs multiplets which produce the required symmetry breaking is small enough so that the $|\Phi|^3$ interaction retains infrared slavery, then a symmetry breaking $G \rightarrow H$ is (i) a first-order transition if it is a gauged symmetry (i.e., G/H corresponds to massive gauge bosons) and (ii) a second-order transition if it is a global symmetry (i.e., G/H corresponds to massless Goldstone bosons). [For global symmetries the restriction (ii) on the number of Φ_i's is irrelevant.]

The separation of the $n=0$ modes is crucial in making transparent the connection between symmetry and the order of the transition. The other modes give only perturbative corrections to this result. The number 42 for the destruction of infrared slavery is not very reliable since we are approximating a nonpolynomial interaction by a polynomial. Nevertheless, all methods of computation will give a value of this order of magnitude. (See Refs. 2 and 3 in this connection.)

There are some models like the SU(5) grand unified theory where the symmetry allows cubic Higgs interactions at the tree level giving a possible first-order transition. The entire discussion of this paper applies to theories where such terms are excluded at the tree level by discrete or continuous symmetires.

We do not have a rigorous proof of our rule. Since it is a statement that depends only on the symmetry structure of the theory, one suspects that it can be sharpened to a theorem and proved in a more general context. One can set up a rigorous axiomatic formulation of temperature-dependent field theory where, instead of the Poincaré group, one has the three-dimensional Euclidean group times a discrete translation group. It is conceivable that this rule can be proved within such a framework independent of the restrictions of perturbation theory.

[1]A. D. Linde, Phys. Lett. 99B, 391 (1981); Rep. Prog. Phys. 42, 389 (1979).

[2]B. I. Halperin, T. C. Lubensky, and S. K. Ma, Phys. Rev. Lett. 32, 292 (1974).

[3]P. Ginsparg, Nucl. Phys. B170 [FS1], 388 (1980).

[4]G. B. Gelmini and M. Roncadelli, Phys. Lett. 99B, 411 (1981).

[5]V. P. Nair, Syracuse Report No. COO-3533-232, 1982 (unpublished).

[6]D. A. Kirzhnits and A. D. Linde, Phys. Lett. 42B, 471 (1972); L. Dolan and R. Jackiw, Phys. Rev. D 9, 3320 (1974); S. Weinberg, ibid. 9, 3357 (1974); C. Bernard, ibid. 9, 3312 (1974).

[7]D. Amit, Field Theory: The Renormalization Group and Critical Phenomena (McGraw-Hill, New York, 1978).

[8]K. G. Wilson and J. Kogut, Phys. Rep. 12C, 75 (1974).