Nature of symmetry breaking in the superconducting ground state

Kazue Matsuyama and Jeff Greensite
Physics and Astronomy Department, San Francisco State University, San Francisco, California 94132, USA

(Received 10 June 2019; revised manuscript received 13 September 2019; published 26 November 2019)

The order parameters which are thought to detect U(1) gauge symmetry breaking in a superconductor are both nonlocal and gauge dependent. For that reason, they are also ambiguous as a guide to phase structure. We point out that a global subgroup of the local U(1) gauge symmetry may be regarded, in analogy to non-Abelian theories, as a "custodial" symmetry affecting the matter field alone, and construct, along the lines of our previous work, a gauge-invariant criterion for breaking symmetries of this kind. It is shown that spontaneous breaking of custodial symmetry is a necessary condition for the existence of spontaneous symmetry breaking of a global subgroup of the (Abelian or non-Abelian) gauge group in any given gauge, and a sufficient condition for the existence of spontaneous breaking of a global subgroup of the gauge group in some gauge. As an illustration, we compute numerically, in a lattice version of the Ginzburg-Landau model, the phase boundaries of the theory and the order parameters associated with various symmetries in each phase.

DOI: 10.1103/PhysRevB.100.184513

I. INTRODUCTION

Superconductivity is the simplest example of a so-called dynamically broken gauge symmetry. In view of the Elitzur theorem, which states that gauge symmetry is unbreakable either dynamically or spontaneously, this characterization deserves closer scrutiny. What symmetry, exactly, is broken? And in which operators is that breaking manifest? The issue is largely a conceptual one since the BCS theory seems perfectly adequate for conventional (noncuprate) superconductors, but these questions seem relevant not just to a deeper understanding of superconductivity, but also to a better understanding of any theory which is claimed to break a gauge symmetry, whether spontaneously or dynamically.

Certainly the ground state of a superconductor, and in fact any physical state, must be invariant under infinitesimal and, more generally, local gauge transformations of the dynamical fields; this is required by the Gauss law condition, and the vanishing of locally noninvariant operators in the ground state is guaranteed by the Elitzur theorem [1]. But neither Gauss's law nor the Elitzur theorem forbids the breaking of a global symmetry, and in fact there is a global U(1) subgroup of the gauge symmetry which appears to be broken by the superconducting ground state. But, the order parameter which has been proposed to detect the breaking of this gauge symmetry is itself gauge dependent, and the magnitude of the order parameter, including whether it is zero or nonzero, depends on the gauge choice, as shown below in an effective model. In view of this fact, is it possible to construct a gauge-invariant criterion which distinguishes the symmetric phase from the symmetry-broken phase in U(1) gauge theories, and in gauge-Higgs theories in general? That is the question we would like to address here.

To fix notation, let $c_\sigma(x),\ c^\dagger_\sigma(x)$ denote the electron operators with spin index σ, transforming as

$$c_\sigma(x) \to e^{i\theta(x)}c_\sigma(x), \quad c^\dagger_\sigma(x) \to e^{-i\theta(x)}c^\dagger_\sigma(x),$$
$$A_\mu(x) \to A_\mu(x) + \frac{1}{e}\partial_\mu\theta(x) \tag{1}$$

under a local gauge transformation. A global U(1) subgroup of the gauge group is defined by the set of transformations with $\theta(x) = \theta$ independent of space. But, this can be regarded as a global symmetry pertaining to the matter sector of the theory alone since the A_μ gauge field is unaffected by such transformations. For this reason, adopting a term from the electroweak sector of the standard model, it may be regarded as a type of "custodial symmetry." Of course, the name we choose to assign to a symmetry may be just a matter of words (although we will support our preference in Sec. III), but the choice of order parameter to detect symmetry breaking is not just semantic. In the context of superconductivity, it is usually the expectation value of the Cooper pair creation operator which is said to detect the breaking of gauge invariance in the BCS ground state. But, since this operator transforms under local as well as global transformations, it can only serve as an order parameter for global symmetry breaking in a fixed gauge. The spontaneous breaking of a "remnant" gauge symmetry, i.e., a global symmetry which remains after gauge fixing, is known to be ambiguous, in that the symmetry-breaking transition depends on the gauge choice [2]. This ambiguity is consistent with the theorem proved by Osterwalder and Seiler [3], whose consequences were elaborated by Fradkin and Shenker [4].

This raises the question of whether we can find a gauge-invariant criterion for the breaking of a global symmetry characterized by $\theta(x) = \theta$. We will address this question, along the lines of our recent work [5], in non-Abelian gauge-Higgs theories, in the context of a lattice version of Ginzburg-Landau theory, i.e., the Abelian Higgs model. In Sec. II we show explicitly the ambiguity of spontaneous gauge symmetry breaking in this model, in the sense that the location (and even the existence) of a symmetry-breaking transition of this kind actually depends on the gauge choice required, either explicitly or implicitly, to define the order parameter.

The main point of this paper is presented in Sec. III. In that section we introduce the concept of custodial

 ©2019 American Physical Society

symmetry, together with a gauge-invariant criterion for custodial symmetry-breaking Abelian and non-Abelian gauge-Higgs theories, pointing out that in the non-Abelian case custodial and global gauge transformations belong to different groups. We then propose that the Higgs phase should be *defined*, in a gauge-invariant manner, as the phase of broken custodial symmetry. In support of this proposal, we show in Sec. IV that custodial symmetry breaking is a necessary condition for the spontaneous breaking of a global subgroup of the gauge symmetry in any given gauge, and a sufficient condition the existence of global gauge symmetry breaking in some gauge. In Sec. V we display the custodial symmetry-breaking transition line in the Abelian Higgs model with a double-charged scalar field, and discuss the possible correspondence of this transition, in the confined region, to a transition between different types of confinement (color confinement and "charge separation" confinement) in the confining phase. In this section we also address the question of Goldstone excitations. In non-Abelian theories the custodial transformations and global gauge transformations are distinct, which raises the question of how the Goldstone theorem is evaded when custodial symmetry is broken. Further questions relating to the Goldstone theorem in the BCS theory are discussed in an Appendix. Our conclusions are in Sec. VI.

In this paper we concentrate on an effective Abelian Higgs model, but for a discussion of related issues in the microscopic theory, when a quantized electromagnetic field is included in the Hamiltonian (see Ref. [6]).

II. AMBIGUITY OF SPONTANEOUS GAUGE SYMMETRY BREAKING

In the absence of gauge fixing, there are no local operators, transforming nontrivially under the gauge group, which can acquire a vacuum expectation value; this is essentially the content of Elitzur's theorem. On the other hand, certain gauges leave unfixed a global subgroup of the gauge group, which we call a "remnant" symmetry, and operators which transform under the remnant symmetry can, depending on the couplings, acquire a vacuum expectation value in the thermodynamic limit. In other words, a remnant symmetry can break spontaneously. But, is this what is meant by a "spontaneously broken" gauge symmetry? We believe this phrase is ambiguous, on the grounds of gauge dependence. Operators which are sensitive to a remnant symmetry in different gauges may not agree on exactly where in the phase diagram the symmetry is actually broken. They may not even agree on whether the symmetry is broken at all.

In this section we will elaborate on this point. For this purpose we will focus on the Abelian Higgs model, where the scalar field has charge qe, where q is an integer. The Abelian Higgs model at $q = 2$ is a relativistic version of the Ginzburg-Landau effective action with a lattice regularization and compact U(1) gauge group. The quantum mechanical model is described by

$$Z = \int DU_\mu D\phi\, e^{-S}, \tag{2}$$

with action

$$S = -\beta \sum_x \sum_{\mu<\nu} \mathrm{Re}[U_\mu(x)U_\nu(x+\hat{\nu})U^*_\mu(x+\hat{\nu})U^*_\nu(x)] - \gamma \sum_x \sum_{\mu=0}^{3} \mathrm{Re}[\phi^*(x)(U_\mu(x))^q \phi(x+\hat{\mu})]. \tag{3}$$

The gauge field is an element of the U(1) group, i.e., $U_\mu(x) = e^{i\chi_\mu(x)}$, and for simplicity we also take the Higgs field to have unit modulus, i.e., $\phi(x) = e^{i\delta(x)}$. Finite temperature is imposed by a finite extension N_t of the lattice in the time direction, i.e., $T = 1/(N_t a)$, where a is the lattice spacing. The action is invariant under U(1) gauge transformations

$$\begin{aligned} U_\mu(x) &\to U'_\mu(x) = e^{i\theta(x)}U_\mu(x)e^{-i\theta(x+\hat{\mu})}, \\ \phi(x) &\to \phi'(x) = e^{iq\theta(x)}\phi(x). \end{aligned} \tag{4}$$

There are two reasons to consider a compact U(1) gauge group. First, the gauge group of electromagnetism may actually be compact. This will be the case if electromagnetism is embedded in some larger compact gauge group, as in many beyond-the-standard model theories. Second, the compact Abelian Higgs model is also the appropriate formulation if we wish to understand symmetry breaking in the Abelian theory in the context of analogous phenomena in non-Abelian gauge Higgs theories. The unit modulus constraint is a convenience, which can be regarded as the $\lambda \to \infty$ limit of a "Mexican hat" potential $V(\phi) = \lambda(\phi\phi^* - \gamma)^2$. We can soften the unit modulus constraint by replacing it with the Mexican hat potential, but this means computing a phase diagram in a three- (rather than two-) dimensional β, γ, λ parameter space, and we believe this extra dimension will not affect the issues we are concerned with in any essential way.

A. "Gauge-invariant" order parameters

It is impossible to identify gauge symmetry breaking using local order parameters such as $\phi(x)$ in the effective theory or, in the (gauge-invariant) microscopic theory, the Cooper pair creation operator $c^\dagger_\uparrow(x)c^\dagger_\downarrow(x)$. These local operators transform under local gauge transformations, and their expectation values vanish in accordance with the Elitzur theorem. In order to detect symmetry breaking of a global gauge symmetry, there are two options, which are essentially equivalent. The first is to fix a gauge which leaves unfixed a global remnant symmetry. Elitzur's theorem will then not rule out an expectation value for $\phi(x)$ or the Cooper pair operator because Elitzur's theorem does not apply to global transformations of any kind. The second option is to introduce a nonlocal operator Q_x which transforms under the global symmetry $\theta(x) = \theta$, but is invariant under all other local gauge transformations. Again, Elitzur's theorem does not rule out a finite expectation value for various operators of this type. Consider an operator Q which is noninvariant under a gauge transformation carried out in a fixed finite volume V_Q. Elitzur's theorem states that $\langle Q \rangle = 0$ even when we carry out the usual procedure of adding a term to break the symmetry, take the infinite volume limit, and then remove the breaking. What is crucial, however, is that Q will vary under a local gauge transformation carried out only in a fixed volume, which remains fixed in the thermodynamic limit. What Elitzur showed is that in this situation the effect of the breaking term can be bounded, even in the infinite volume limit, by some small parameter which is taken to zero at the end. Details can be found in [1,7]. If, on the other hand, Q only varies under transformations carried out at every site on the lattice (or, in the present case, throughout the volume of the solid), the bound fails in the thermodynamic limit, and the theorem does not apply. And of course it must fail in this situation, for otherwise Elitzur's argument would also rule out the spontaneous breaking of ordinary global symmetries. But, it is not assured, for different Q_x operators or for remnant global symmetries in different gauges, that the global symmetry breaking occurs at the same place in the phase diagram.

As an example, let us introduce, in the microscopic theory, the phase factor

$$e^{i\gamma_x} = \exp\left[i\frac{e}{4\pi}\int d^3z\, A_i(z)\frac{\partial}{\partial z_i}\frac{1}{|\boldsymbol{x}-\boldsymbol{z}|}\right] \tag{5}$$

and define

$$\theta(x) = \theta_0 + \varphi(x) \text{ where } \int d^3x\,\varphi(x) = 0. \tag{6}$$

Under a local gauge transformation (1) we have

$$e^{i\gamma_x} \to e^{i\varphi(x)}e^{i\gamma_x}. \tag{7}$$

Next, introduce operators which are invariant under local transformations

$$\widetilde{c}_\sigma(x) = c_\sigma(x)e^{-i\gamma_x}, \quad \widetilde{c}^\dagger_\sigma(x) = c^\dagger_\sigma(x)e^{i\gamma_x}, \tag{8}$$

and define

$$Q_x = \widetilde{c}^\dagger_\uparrow(x)\widetilde{c}^\dagger_\downarrow(x). \tag{9}$$

Then, under an arbitrary gauge transformation $\theta(x)$, Q_x transforms only under the zero mode

$$Q_x \to e^{2i\theta_0}Q_x. \tag{10}$$

We then observe that

$$g_C(x;A) = e^{i\gamma_x} \tag{11}$$

is precisely the gauge transformation which takes the matter and gauge fields into Coulomb gauge. Exactly (and only) in Coulomb gauge, $e^{i\gamma_x} = 1$, and Q_x has the local form $Q_x = c^\dagger_\uparrow(x)c^\dagger_\downarrow(x)$. But, of course the locality is deceptive since Coulomb gauge fixing is a nonlocal operation. In any case, evaluating $c^\dagger_\uparrow(x)c^\dagger_\downarrow(x)$ in Coulomb gauge, and evaluating the "gauge-invariant" operator Q_x in the absence of gauge fixing, are completely equivalent.

Obviously, this construction generalizes. Let $g_G(x;A)$ [or, on the lattice $g_G(x;U)$ where U denotes lattice link variables] be a gauge transformation which takes the matter and gauge fields into some gauge G, and which leaves unfixed the remnant symmetry $\theta(x) = \theta$. We then define, e.g.,

$$Q_{G,x} = g_G(x;A) \circ \{c^\dagger_\uparrow(x)c^\dagger_\downarrow(x)\} \tag{12}$$

in the microscopic theory, or

$$Q_{G,x} = g_G(x;U) \circ \phi(x) \tag{13}$$

in the Abelian Higgs theory to serve as nonlocal order parameters for breaking of the remnant symmetry. Constructions of that type are found in the literature, e.g., in [8,9], where G is an axial gauge, or in [10], where G is lattice Landau gauge. Although these order parameters are described as (and in fact are) locally gauge invariant, it should be understood that a certain gauge choice, and therefore a certain arbitrariness, underlies these constructions. Evaluation of such Q observables, in the absence of gauge fixing, is completely equivalent to evaluating the expectation value $\langle\phi\rangle$ in a particular gauge.

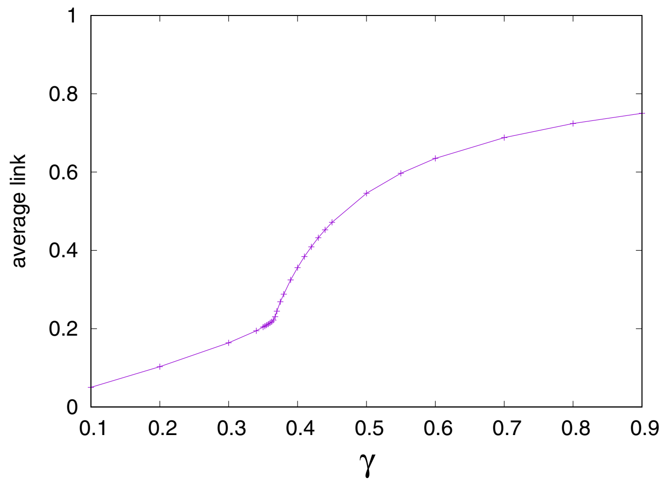


FIG. 1. Average link L [Eq. (14)] vs γ at $\beta = 2.0$ on a 16^4 lattice volume. The transition from the massless to the Higgs phase is located at $\gamma = 0.365$, where the slope changes abruptly.

B. Phase diagrams

The phase diagram of the Abelian Higgs model in the space of couplings β, γ and charges $q = 1, 2, 6$ was determined long ago by Ranft *et al.* [11], albeit on lattices which were tiny (4^4) by today's standards, with transition points located by a method (hysteresis curves) which has since been superseded by other methods. For this paper we have determined the transition points in the $q = 2$ theory, from the confinement to the Higgs or massless phases, from the location of peaks in the plot of plaquette susceptibility vs β, at fixed γ, on a 12^4 lattice volume. Transition points from the massless to Higgs phase are located from the position of a "kink," i.e., an abrupt change in slope, in a plot of the link action

$$L = \frac{1}{V}\sum_x\sum_\mu \left\langle \mathrm{Re}\left[\phi^*(x)U^2_\mu(x)\phi(x+\hat{\mu})\right]\right\rangle \tag{14}$$

vs γ. An example of data for L vs γ at $\beta = 2$, on a 16^4 lattice, is plotted in Fig. 1, and the kink is apparent at $\gamma \approx 0.365$. Since this behavior should reflect a nonanalyticity of the free energy in the thermodynamic limit, we would expect the change in slope at the transition to become increasingly abrupt as the volume increases. In Fig. 2 we show our data for L vs γ in the immediate neighborhood of the transition point, at lattice volumes 8^4, 12^4, 16^4, which agrees with this expectation.

In the end, our results for the thermodynamic phase structure of the $q = 2$ theory, displayed in Fig. 3, are not far off the old results of [11]. We should point out that in the confinement phase denoted "conf" in Fig. 3, what is really confined are test

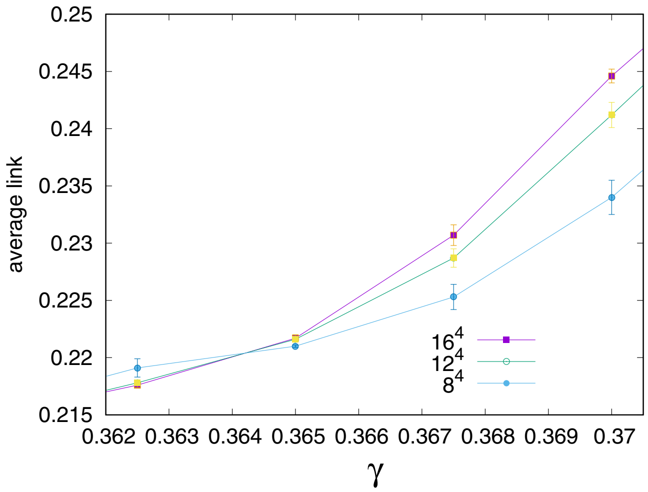


FIG. 2. Closeup of the L vs γ data at $\beta = 2.0$ in the immediate neighborhood of the transition, on 8^4, 12^4, 16^4 lattice volumes. Note that the change in slope at the transition near $\gamma = 0.365$ becomes more abrupt with increasing volume.

charges with ± 1 units ($q = 1$) of electric charge. The meaning of confinement in this region for $q = 2$ charges, and how the confinement phase for $q = 2$ charges is distinguished from the Higgs phase, is not at all trivial, and will be discussed in Sec. V B. The massless phase is continuously connected to the massless phase of the pure gauge theory at $\gamma = 0$, which is known to have a transition between the confined and massless phases at $\beta = 1$.

Let us define, in the $q = 2$ Abelian Higgs theory, two different order parameters Q_L and Q_T, each of which transforms under a global subgroup, defined by $\theta(x) = \theta$, of the local U(1) gauge symmetry, via

$$Q_L \to e^{2i\theta} Q_L, \quad Q_T \to e^{2i\theta} Q_T, \tag{15}$$

but which are invariant under any local gauge transformation. These operators are defined by the gauge transformations

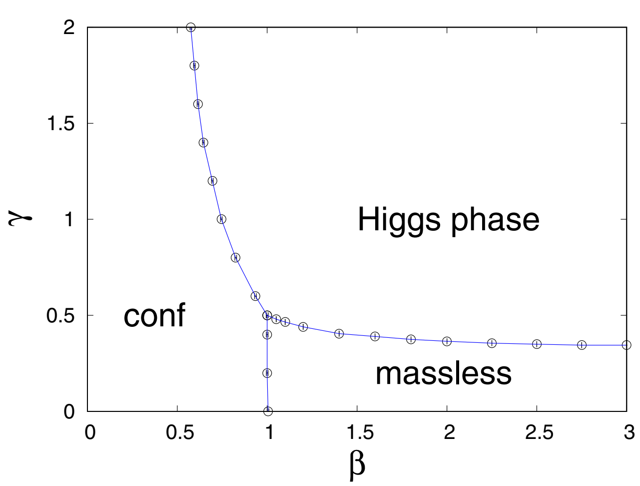


FIG. 3. The $q = 2$ phase diagram. The confining, massless, and Higgs phases are completed, separated by thermodynamic transitions.

$g_L(x;U)$ and $g_T(x;U)$ which take the system into Landau and temporal gauge, respectively, then operating on ϕ and averaging over volume,

$$Q_{L,T} = \frac{1}{V} \sum_x g_{L,T}(x;U) \circ \phi(x), \tag{16}$$

where V is the lattice volume. Both Landau and temporal gauge leave unfixed a remnant symmetry. Lattice Landau gauge is defined as the gauge that maximizes

$$R_L = \sum_x \sum_{\mu=1}^{4} \mathrm{Re}[U_\mu(x)]. \tag{17}$$

The subscript T denotes "maximal" temporal gauge. On a periodic lattice one cannot fix all timelike links to the identity, the most that can be done in temporal (or any axial) gauge is to fix links in a maximal "tree" since it is impossible to gauge fix all links in a loop. Maximal temporal gauge is defined by setting

$$\begin{aligned} &U_4(x) = 1 \text{ for } x_4 \neq 1, \\ &U_3(x) = 1 \text{ for } x_4 = 1,\ x_3 \neq 1, \\ &U_2(x) = 1 \text{ for } x_4 = 1,\ x_3 = 1,\ x_2 \neq 1, \\ &U_1(x) = 1 \text{ for } x_4 = 1,\ x_3 = 1,\ x_2 = 1,\ x_1 \neq 1. \end{aligned} \tag{18}$$

Landau and maximal temporal gauge fix all but a remnant global symmetry $\theta(x) = \theta$.

In lattice Landau gauge, however, we have to contend with the Gribov ambiguity, i.e., the fact that there are many local maxima of R, and therefore the full specification of $g_L(x;U)$ depends on the Gribov copy selected. Obviously, no fully gauge-invariant observable can depend on such a choice, but we are dealing here with order parameters which, as we shall see, most definitely depend on the gauge. The most natural choice in Landau gauge would be the transformation g_L which brings R to an absolute maximum. Numerically, this is impossible to achieve in practice, in fact the determination of the absolute maximum is believed to be NP hard. However, any deterministic algorithm will select a unique gauge copy corresponding to a local maximum of R, given a particular lattice configuration $U_\mu(x)$, so the specific gauge-fixing algorithm used by the computer may be regarded as part of the specification of the gauge choice.

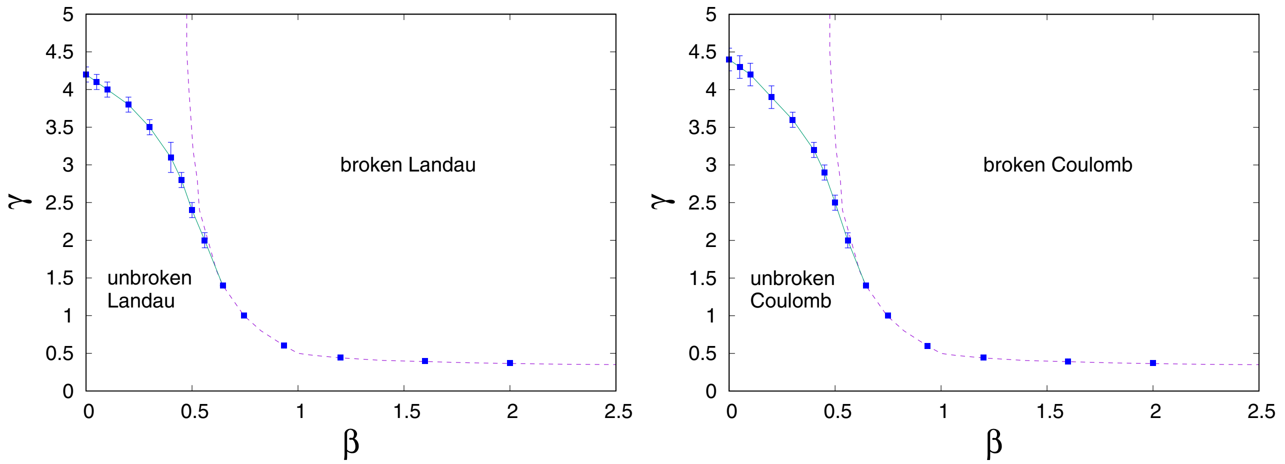
We may also define lattice Coulomb gauge as the gauge which maximizes

$$R_C = \sum_x \sum_{i=1}^{3} \mathrm{Re}[U_i(x)], \tag{19}$$

and $g_C(x;U)$ as the gauge transformation to Coulomb gauge. In Coulomb gauge there remains a symmetry under gauge transformations which depend only on time, i.e., $\theta(\mathbf{x},t) = \theta(t)$. On any given time slice, this is a remnant global symmetry, which may be spontaneously broken on that time slice. We therefore define the Q observable on each time slice as

$$Q_C(t) = \frac{1}{V_3} \sum_{\mathbf{x}} g_C(\mathbf{x},t;U) \circ \phi(\mathbf{x},t), \tag{20}$$

where V_3 is the ($D = 3$)-dimensional spatial volume of the time slice. Of course, there is no true phase transition on



(a) Landau gauge (b) Coulomb gauge

FIG. 4. Transition points for the breaking of a global remnant gauge symmetry in (a) Landau and (b) Coulomb gauges. The dashed line is the line of thermodynamic transition shown in Fig. 3.

a finite volume, and so in practice we compute, in a fixed volume V,

$$Q_{L,T}(V) = \frac{1}{V}\left|\sum_x \phi(x)\right|,$$

$$Q_C(V_3, t) = \frac{1}{V_3}\left|\sum_{\boldsymbol{x}} \phi(\boldsymbol{x}, t)\right|, \tag{21}$$

$$Q_C(V_3) = \frac{1}{N_t}\sum_{t=1}^{N_t} Q_C(V_3, t), \tag{22}$$

with $\phi(x)$ fixed to Landau, maximal temporal, or Coulomb gauge, respectively, and extrapolate the results to $V = \infty$. Transitions are located by peaks in the susceptibilities

$$\chi_L = V[\langle Q_L(V)^2\rangle - \langle Q_L(V)\rangle^2],$$
$$\chi_C = \frac{1}{N_t}\sum_{t=1}^{N_t} V_3[\langle Q_C(V_3, t)^2\rangle - \langle Q_C(V_3, t)\rangle^2]. \tag{23}$$

We have seen in Fig. 3 that in the $q = 2$ case there are three phases, which we denote as "massless," "Higgs," and "confinement," completely separated from one another by lines of thermodynamic transition. In the massless phase, all three of the order parameters Q_L, Q_C, Q_T extrapolate to zero at infinite volume, as one might expect. Within the Higgs phase, the remnant global gauge symmetry is spontaneously broken in the full volume, for Landau gauge, and in any time slice, in Coulomb gauge. However, the remnant symmetries in Landau and Coulomb gauges are *also* broken inside the confinement phase, at higher γ values, and moreover the Landau and Coulomb transition lines do not coincide within the confinement phase. The phase diagrams for remnant symmetry breaking, for Landau and Coulomb gauges, are shown in Fig. 4. In this figure the remnant symmetries break at the points shown, while the thermodynamic transition is indicated by the dashed line. We see that at small β there is a line of remnant symmetry breaking in the confined region which does not correspond to any thermodynamic transition, and which lies entirely in the confined phase. Moreover, the transition line in the confined phase is slightly different in Landau and Coulomb gauges, as seen in Fig. 5. Already, we can conclude that spontaneous breaking of remnant gauge symmetry is gauge dependent.

Even within the Higgs phase, spontaneous gauge symmetry breaking is not seen in all gauges which leave unfixed a global subgroup of the gauge symmetry. In Fig. 6 we display Q_L and Q_T vs $1/\sqrt{V}$, at a point $\beta = 1.2$, $\gamma = 0.7$ which is inside the Higgs phase (as determined by thermodynamic transitions, see Fig. 3). It is seen Q_T extrapolates to zero at infinite lattice volume inside the Higgs phase, while Q_L does not. Here again we have evidence of the gauge dependence of spontaneous symmetry breaking of remnant gauge symmetry.

The fact that Q_T extrapolates to zero in the Higgs phase, and in fact throughout the phase diagram, is a particularly clear indication of the ambiguity of spontaneous gauge

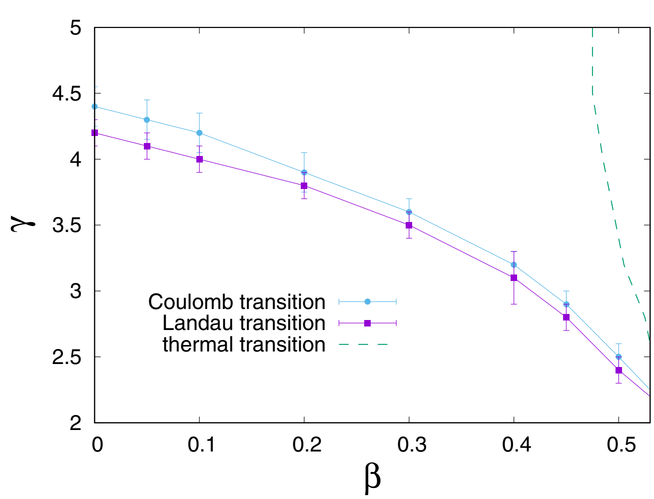


FIG. 5. Closeup of the remnant symmetry-breaking points in Landau and Coulomb gauges, away from the line of thermal transitions.

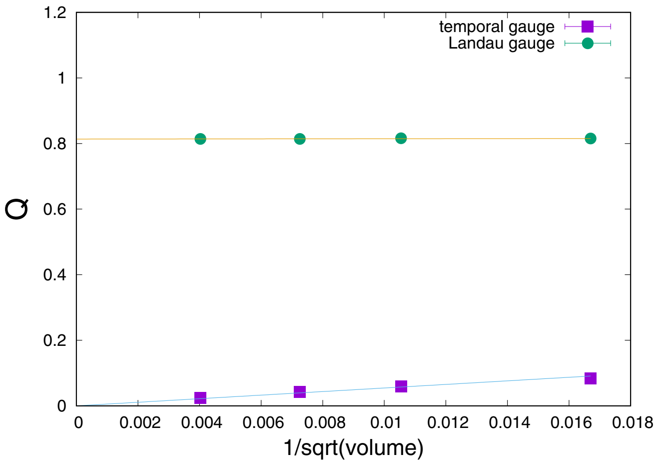


FIG. 6. The Landau and temporal gauge order parameters Q_L and Q_T vs inverse square root of the lattice volume $1/\sqrt{V}$, inside the Higgs phase at $\beta = 1.2$, $\gamma = 0.7$. The remnant global gauge symmetry is broken in this phase in Landau gauge, according to Q_L, but is not broken in temporal gauge, according to Q_T, which extrapolates to zero at infinite volume.

symmetry breaking. But, the vanishing of Q_T at infinite volume should not be a surprise. Suppose, instead of fixing to a maximal temporal gauge as in (18), we only impose $U_4(\boldsymbol{x}, t) = 1$ at all $t = x_4 \neq 1$ [the $U_4(\boldsymbol{x}, t = 1)$ links cannot be fixed also since this would require gauge fixing links on a closed contour, which is impossible]. Then, there is a residual *local* symmetry under time-independent gauge transformations $g(\boldsymbol{x}, t) = g(\boldsymbol{x})$, which can differ at every $\boldsymbol{x}$. The residual symmetry is reduced, in maximal temporal gauge, to a global space-time-independent symmetry $g(\boldsymbol{x}, t) = g = e^{i\theta}$ by the additional gauge fixing on the $t = x_4 = 1$ time slice shown in (18).[1] We interpret the vanishing of Q_T in the infinite volume limit to mean that away from $t = 1$ the local symmetry is effectively recovered, probably because the $t = 1$ boundary conditions are irrelevant, in the bulk of a large volume, in a system with a finite correlation length. The same situation is expected to hold, for the same reason, in other (maximal) axial gauges, where the links are fixed in a configuration such that only a global remnant symmetry remains. In fact, because of the Euclidean symmetry, maximal $A_z = 0$ gauge is no different in this respect from maximal temporal gauge. But, the fact that we can understand *why* Q_T should vanish, i.e., because of the irrelevance of a boundary condition which fixes the remnant local symmetry, does not affect the fact that the spontaneous breaking of a remnant global symmetry in some gauges is not manifest in other gauges. Or, to put it another way, the fact that an operator transforms nontrivially only under a global subgroup of the gauge symmetry does not guarantee that this operator shows a transition from the "symmetric" massless phase to the "broken" Higgs phase.

The ambiguity outlined here is certainly not limited to the Abelian Higgs model, in fact, it was first noted in Ref. [2]

[1] In calculating Q_T we have excluded $\phi(x)$ on the time slice $t = 1$, which should make no difference in the infinite volume limit.

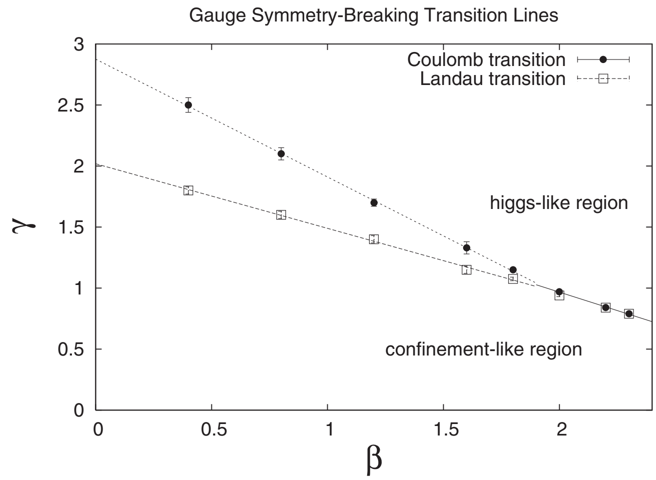


FIG. 7. The remnant symmetry-breaking lines for Coulomb and Landau gauges in SU(2) gauge-Higgs theory with action (24).

for the SU(2) gauge-Higgs model, with the Higgs field in the fundamental representation of the gauge group. The action in this case is

$$\begin{aligned} S &= -\beta \sum_x \sum_{\mu<\nu} \frac{1}{2} \mathrm{Tr}[U_\mu(x) U_\nu(x+\hat{\nu}) U_\mu^\dagger(x+\hat{\nu}) U_\nu^\dagger(x)] \\ &\quad -\gamma \sum_{x,\mu} \frac{1}{2} \mathrm{Tr}[\phi^\dagger(x) U_\mu(x) \phi(x+\hat{\mu})] \\ &= S_W + S_H, \end{aligned} \tag{24}$$

with $\phi(x)$ an SU(2)-valued field. It was found that the breaking of the residual gauge invariance in Coulomb and Landau gauges occurs along different transition lines, shown in Fig. 7. There is no thermodynamic transition in the region of the phase diagram where the Landau and Coulomb lines differ.

C. Z_2 symmetry breaking

Apart from gauge symmetry, the action of the $q = 2$ gauge-Higgs model is invariant under the following symmetry:

$$U_4(\mathbf{x}, 0) \rightarrow z U_4(\mathbf{x}, 0) \quad \text{all } \mathbf{x} \text{ at } t = 0, \tag{25}$$

where $z = \pm 1$ is an element of the Z_2 group. For pure gauge theory ($\gamma = 0$), z is an element of U(1), and the symmetry is known as "center symmetry." In the $q = 2$ model the U(1) center symmetry is broken down to Z_2, while in the $q = 1$ model the symmetry is absent entirely. A gauge-invariant observable which transforms nontrivially under the Z_2 symmetry is the Polyakov line

$$P(\mathbf{x}) = U_4(\mathbf{x}, 1) U_4(\mathbf{x}, 2) \dots U_4(\mathbf{x}, N_t), \tag{26}$$

where $P(\mathbf{x}) \rightarrow z P(\mathbf{x})$ under (25). Therefore, the expectation value of the Polyakov line is an order parameter for spontaneous breaking of global Z_2 symmetry. Moreover, since

$$\langle P \rangle \sim e^{-F/kT}, \tag{27}$$

where F is the free energy of a static source with a single unit of charge, $\langle P \rangle = 0$ implies confinement, and $\langle P \rangle \neq 0$ means nonconfinement, of particles with a single unit ($q = 1$) of charge. Thus, we expect $\langle P \rangle = 0$ in the region labeled "conf"

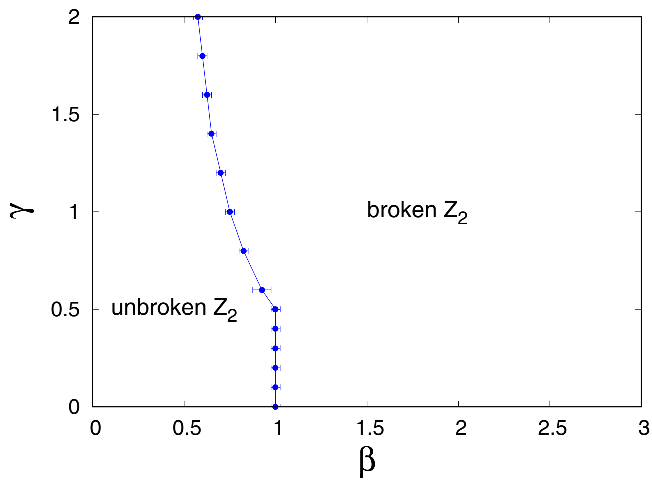


FIG. 8. The Z_2 transition line, as detected by Polyakov lines on a $12^3 \times 6$ lattice. This line coincides with thermodynamic transitions from the confinement phase for $q = 1$ (but not $q = 2$) test charges, to the massless and Higgs phases.

of the phase diagram shown in Fig. 3, and $\langle P \rangle \neq 0$ in the Higgs and massless phases. We have verified (on a $12^3 \times 6$ lattice volume) that the transition happens across the transition line shown in Fig. 8, separating the confinement from the Higgs and massless phases.

D. Numerical details

The numerical simulations are carried out by standard methods. The $|\phi| = 1$ constraint is implemented by expressing $\phi(x) = \exp[i\varphi(x)]$. At each update, one generates a random number $r(x)$ uniformly distributed in the interval [0,1], and an angle $\eta(x) = 2\pi\xi[0.5 - r(x)]$. Then, a trial $\phi_{\text{new}}(x)$ is generated, with

$$\phi_{\text{new}}(x) = e^{i\eta(x)}\phi(x) \tag{28}$$

and this trial value is accepted or rejected according to the Metropolis algorithm. The constant ξ is adjusted, after each update sweep of the lattice, in order to keep the acceptance rate in the range 40%–60%. The same procedure (with a different value of ξ) is used for updating link variables $U_\mu(x) = \exp[i\alpha_\mu(x)]$. We carried out 100 update sweeps through the lattice for each data-taking sweep.

Both Landau and Coulomb gauge are implemented via over-relaxation [12]. The procedure is to compute, at site x, the quantity

$$\begin{aligned} A(x) &= \sum_{\mu=1}^{4}[U_\mu(x) + U_\mu^\dagger(x-\hat{\mu})] \\ &= \rho(x)e^{i\theta(x)}, \end{aligned} \tag{29}$$

with $\rho(x)$ real and positive. Then, the gauge transformation $g(x) = \exp[-i\theta(x)]$, applied at site x only, maximizes the real part of $A(x)$. In over-relaxation we apply instead the transformation $g(x) = \exp[-i\kappa\theta(x)]$, with $\kappa > 1.0$. The procedure is to sweep through the lattice site by site, calculating and applying the transformation $g(x)$ at each site. After each sweep one computes the average value of $\text{Re}[U_\mu(x)]$ in the lattice volume. When a convergence criterion is satisfied (the fractional change from one sweep to the next is less than 10^{-7}), the gauge fixing is terminated, and the value of $\phi(x)$ in Landau gauge, averaged over the lattice volume, is computed. Over-relaxation speeds up convergence to the gauge, and we have found that the choice of $\kappa = 1.7$ works well. As explained above, in connection with the Gribov copy issue, the numerical gauge-fixing algorithm should really be regarded as part of the specification of the gauge condition.

The gauge-fixing procedure in Coulomb gauge is similar, except that one computes, on each time slice

$$\begin{aligned} A(\boldsymbol{x}, t) &= \sum_{\mu=1}^{3}[U_\mu(\boldsymbol{x}, t) + U_\mu^\dagger(\boldsymbol{x}, t)] \\ &= \rho(\boldsymbol{x}, t)e^{i\theta(\boldsymbol{x},t)}. \end{aligned} \tag{30}$$

The gauge-fixing procedure is applied to each time slice separately, with the gauge fixed at time t before moving on to fix the gauge on time slice $t + 1$. The advantage is that some time slices require more gauge-fixing sweeps than others, and since the gauge condition on one slice does not depend on other time slices, it is more efficient to fix each time slice separately.

III. CUSTODIAL SYMMETRY BREAKING

We have seen that operators such as Q_C, Q_L, which transform under some global subgroup of the gauge group, can have transition lines which do not coincide everywhere, and some operators of this kind, such Q_T may have no transition line at all. The expectation value of each of these operators amounts to the expectation value $\langle\phi\rangle$ of the scalar field, evaluated in some gauge which leaves unfixed a global remnant symmetry. Let us take the point of view that if there is even one gauge of this type in which $\langle\phi\rangle$ is nonzero, then some symmetry is, in some sense, broken. But which symmetry and in what sense, given that $\langle\phi\rangle$ is nonzero in some gauges but not in others? What we are looking for is a gauge-invariant criterion for the breaking of some (not necessarily gauge) symmetry such that, if there exists *any* gauge-fixing condition $F(U) = 0$ in which $\langle\phi\rangle \neq 0$, then this symmetry is spontaneously broken. In this section we will introduce the concept of custodial symmetry, and in the following section we will show that it has the property just mentioned.

The term "custodial symmetry" is derived from the electroweak sector of the standard model. We will define a custodial symmetry to be a global symmetry of one or more matter fields which (i) does not transform the gauge field; and for which (ii) any local operator which transforms nontrivially under the custodial symmetry also transforms nontrivially under the local gauge symmetry. The spontaneous or dynamical breaking of such a symmetry is therefore masked by the unbroken gauge symmetry, which makes it difficult to see how to construct an order parameter for the custodial symmetry breaking without first fixing the gauge symmetry in some way. We have already encountered one such symmetry, namely, the transformation (1) with $\theta(x) = \theta$ independent of space. Another symmetry of this kind is well known in the electroweak sector of the standard model. Returning to the

SU(2) lattice gauge-Higgs theory (24), we note that the action is invariant under

$$U_\mu(x) \to L(x)U_\mu(x)L^\dagger(x+\hat{\mu}), \quad \phi(x) \to L(x)\phi(x)R, \tag{31}$$

where $L(x) \in \mathrm{SU}(2)_{\mathrm{gauge}}$ is a local gauge transformation, while $R \in \mathrm{SU}(2)_{\mathrm{global}}$ is a global transformation. $\mathrm{SU}(2)_{\mathrm{global}}$ is sometimes referred to as the "custodial" symmetry of the theory (cf. [13]) and it is obviously distinct from the gauge group.

It should be noted that if we choose a gauge (e.g., unitary gauge) in which the Higgs field acquires a vacuum expectation value

$$\langle\phi\rangle = \begin{bmatrix} \upsilon & 0 \\ 0 & \upsilon \end{bmatrix}, \tag{32}$$

then the $\mathrm{SU}(2)_{\mathrm{gauge}} \times \mathrm{SU}(2)_{\mathrm{global}}$ symmetry is broken down to a diagonal global subgroup

$$\mathrm{SU}(2)_{\mathrm{gauge}} \times \mathrm{SU}(2)_{\mathrm{global}} \to \mathrm{SU}(2)_D, \tag{33}$$

corresponding to transformations

$$L(x) = R^\dagger = g, \quad \phi(x) \to g\phi(x)g^\dagger, \quad U_\mu(x) \to gU_\mu(x)g^\dagger. \tag{34}$$

Some authors refer to transformations in this diagonal subgroup, which preserve the vacuum expectation value of ϕ in a fixed gauge, as the custodial symmetry group. Whatever the terminology, custodial symmetry has a role to play in the phenomenology of the electroweak interactions, and is reviewed in many places, e.g., [13–15]. Here, however, we wish to focus first on the $\mathrm{SU}(2)_{\mathrm{global}}$ group of R transformations in the absence of gauge fixing, moving from there to the $\theta(x) = \theta$ global U(1) symmetry group in the Abelian theory.

Does it make any sense to describe the Higgs phase of the theory as a phase of spontaneously broken $\mathrm{SU}_{\mathrm{global}}$ symmetry, what we call here custodial symmetry? Local gauge symmetries cannot break according to the Elitzur theorem, and the breaking of a global subgroup of the gauge symmetry appears to depend on the gauge choice, as we have seen in the previous section. There is also no gauge-invariant local order parameter for custodial symmetry breaking, so it cannot break spontaneously in the usual sense (and if it did, one would have to contend with the Goldstone theorem). On the other hand, the full partition function of the SU(2) gauge-Higgs theory can be regarded as a sum of partition functions of a spin system in an external gauge field, i.e.,

$$Z = \int DU\, Z_{\mathrm{spin}}[U]e^{-S_W(U)}, \tag{35}$$

where

$$Z_{\mathrm{spin}}[U] = \int D\phi\, e^{-S_H(U,\phi)}, \tag{36}$$

and, depending on U, custodial symmetry *can* break in the system described by $Z_{\mathrm{spin}}(U)$.

Let us define the expectation value of an operator $\Omega[U,\phi]$ in the spin system

$$\overline{\Omega}(U) = \frac{1}{Z_{\mathrm{spin}}(U)} \int D\phi\, \Omega(\phi, U)e^{-S_H}, \tag{37}$$

with the full expectation value

$$\langle\Omega\rangle = \int DU\, P(U)\overline{\Omega}(U) = \frac{1}{Z}\int DU\, D\phi\, \Omega(\phi, U)e^{-S}. \tag{38}$$

This means that the expectation value in the spin system is to be evaluated from ensembles with U chosen from the probability distribution

$$P(U) = \frac{1}{Z} Z_{\mathrm{spin}}[U]e^{-S_W(U)}. \tag{39}$$

So, the question becomes the following: Is $Z_{\mathrm{spin}}[U]$ in the broken or the unbroken phase for gauge field configurations selected from this probability distribution? It is not hard to devise a gauge-invariant operator $\Phi(U)$ which is nonzero in the broken phase, and which vanishes in the unbroken phase in the thermodynamics limit. Then, $\langle\Phi\rangle \neq 0$, i.e., custodial symmetry breaking, is our proposed *definition* of the Higgs phase of a gauge-Higgs theory.

In a numerical simulation we may determine whether $Z_{\mathrm{spin}}(U)$ is in the broken phase in the probability distribution defined by (39) by a "Monte Carlo-within-a-Monte Carlo simulation." The procedure is to update the $U_\mu(x)$, $\phi(x)$ fields in the full gauge-Higgs theory in the usual way for, e.g., 100 update sweeps, which is followed by the data-taking procedure, which is itself a lattice Monte Carlo simulation of $Z_{\mathrm{spin}}(U)$, keeping the link variables fixed at whatever they were at the end of the last update sweep. The $Z_{\mathrm{spin}}(U)$ simulation proceeds for n_{spin} sweeps, updating only the $\phi(x)$ variables. Let $\phi(x,n)$ denote $\phi(x)$ at the nth update sweep of the spin system, and let

$$\overline{\phi}_{n_{\mathrm{spin}}}(x) = \frac{1}{n_{\mathrm{spin}}} \sum_{n=1}^{n_{\mathrm{spin}}} \phi(x,n). \tag{40}$$

We then define

$$\Phi_{n_{\mathrm{spin}},V}[U] = \frac{1}{V}\sum_x \left|\overline{\phi}_{n_{\mathrm{spin}}}(x)\right|, \tag{41}$$

where $|\phi| = \det^{\frac{1}{2}}(\phi)$, and

$$\Phi[U] = \lim_{n_{\mathrm{spin}}\to\infty}\lim_{V\to\infty} \Phi_{n_{\mathrm{spin}},V}[U]. \tag{42}$$

Averaging $\Phi_{n_{\mathrm{spin}},V}[U]$ over many data-taking sweeps at large n_{spin}, and extrapolating to infinite volume, provides a numerical estimate of $\langle\Phi[U]\rangle$. Then, the Higgs phase of the full gauge-Higgs theory is distinguished from the unbroken phase by

$$\langle\Phi[U]\rangle = \begin{cases} \text{zero}, & \text{unbroken phase} \\ \text{nonzero}, & \text{Higgs phase.} \end{cases} \tag{43}$$

This procedure was carried out for the SU(2) and SU(3) gauge-Higgs models in Ref. [5], where we have determined the transition line between the phases of broken and unbroken custodial symmetry, as defined above. The custodial symmetry-breaking transition in the SU(2) theory is shown

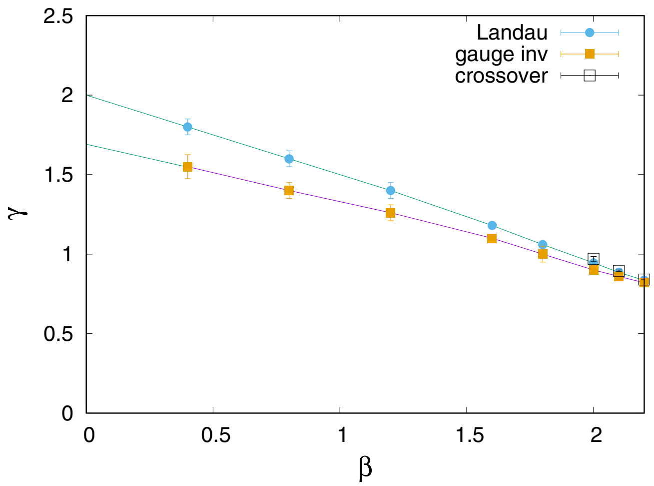


FIG. 9. The custodial symmetry (labeled "gauge inv") and Landau gauge remnant symmetry transition points in SU(2) gauge-Higgs theory. Points labeled "crossover" locate a sharp thermodynamic crossover, but not a phase transition. Note that the Landau transition lies above the line of custodial symmetry breaking.

in Fig. 9, together with the remnant symmetry-breaking line for Landau gauge. At the larger β values, the two transitions coincide, and also coincide with a sharp crossover in the action vs γ, which is also shown. The Coulomb transition line (not shown, but see Fig. 7), lies above the Landau transition.

We may define the gauge-invariant observable $\Phi[U]$ more formally, without any appeal to numerical simulations, by introducing a small perturbation which is removed after taking the thermodynamic limit. Let

$$\overline{\phi}_{JV}[x;U,\eta] = \frac{1}{Z_{\text{spin}}[U]}\int D\phi\,\phi(x) \times \exp\left[-S_H + J\sum_x \text{Tr}[\eta^\dagger(x)\phi(x)]\right], \tag{44}$$

where $\eta(x)$ is a unimodular field $|\eta| = 1$, which is chosen to be any one of an equivalent set of configurations, related by the $\text{SU}_{\text{global}}$ symmetry, which maximizes the averaged sum of moduli

$$\Phi_{JV}[U] = \max_\eta \frac{1}{V}\sum_x |\overline{\phi}_{JV}[x;U,\eta]|. \tag{45}$$

We then define the order parameter for symmetry breaking

$$\langle\Phi\rangle = \lim_{J\to 0}\lim_{V\to\infty}\langle\Phi_{JV}[U]\rangle, \tag{46}$$

with the order of limits as shown. This parameter is nonzero if the $\text{SU}_{\text{global}}$ symmetry of the spin system is spontaneously broken, and zero otherwise.

The maximization condition in (45) guarantees the gauge invariance of $\Phi_{JV}[U]$, even at finite J. This is not hard to see, and goes as follows. Let $\eta = \eta_{\max}$ be any one of the set η's that maximizes the right-hand side of (45). The members of this set are easily seen to be related to one another by custodial symmetry transformations, i.e., if $\eta_{\max}(x)$ is a maximizing configuration, then so is $\eta_{\max}(x)R$, where R is an element of the custodial group. Let $g(x)$ be any local gauge transformation, and let $\phi(x) = g(x)\phi'(x)$, $\eta'(x) = g^\dagger(x)\eta(x)$. Then,

$$\begin{aligned}
\Phi_{JV}[g\circ U] &= \max_\eta \frac{1}{V}\sum_x |\overline{\phi}_{JV}[x; g\circ U,\eta]| \\
&= \max_\eta \frac{1}{V}\sum_x \frac{1}{Z_{\text{spin}}(g\circ U)}\left|\int D\phi\,\phi(x)\exp\left[-S_H(g\circ U,\phi) + J\sum_x \text{Tr}[\eta^\dagger(x)\phi(x)]\right]\right| \\
&= \max_\eta \frac{1}{V}\sum_x \frac{1}{Z_{\text{spin}}(U)}\left|\int D\phi' g(x)\phi'(x)\exp\left[-S_H(g\circ U, g\circ\phi') + J\sum_x \text{Tr}[\eta^\dagger(x)g(x)\phi'(x)]\right]\right| \\
&= \max_{\eta'} \frac{1}{V}\sum_x \frac{1}{Z_{\text{spin}}(U)}\left|g(x)\int D\phi'\phi'(x)\exp\left[-S_H(U,\phi') + J\sum_x \text{Tr}[\eta'^\dagger(x)\phi'(x)]\right]\right| \\
&= \Phi_{JV}[U],
\end{aligned} \tag{47}$$

where we have used the gauge invariance of $Z_{\text{spin}}(U)$ and $S_H(U,\phi)$, the fact that the modulus of the functional integration is independent of the unitary matrix $g(x)$ in front of the integral, and the fact that it does not matter whether the expression is maximized with respect to η or η'. Obviously, the last line is maximized at $\eta' = \eta_{\max}$ or $\eta(x) = g(x)\eta_{\max}(x)$. In other words, the η fields which satisfy the maximization condition transform covariantly under a gauge transformation, and the order parameter itself is gauge invariant.

A field $\eta(x)$ which maximizes the right-hand side of (45) for a given $U_\mu(x)$ configuration is very difficult to determine in practice. Since no gauge is fixed, $U_\mu(x)$ varies wildly in space, and the same will be true of $\eta(x)$. Were we to define the spatial average of $\overline{\phi}_J[x;U,\eta]$ before taking the modulus, it would average to zero in general. In practice we use the lattice Monte Carlo procedure, described above, to determine $\langle\Phi\rangle$.

Unitary gauge

One might wonder what happens to the $\text{SU}_{\text{global}}$ symmetry in unitary gauge, where there is no longer any freedom to transform ϕ. In fact, the relevant degrees of freedom are still there in unitary gauge, but they are now found in the gauge sector. Let us write $U_\mu(x) = g(x)U^F_\mu(x)g^\dagger(x+\hat\mu)$, where U^F_μ is the gauge field in the gauge $F(U) = 0$. Custodial symmetry is now a global transformation on the $g(x)$ field. To see this, begin by fixing to $\phi = \mathbb{1}$. Then,

$$Z = \int DU\ \exp\left[-S_W + \gamma \sum_{x,\mu} \frac{1}{2}\mathrm{Tr}[U_\mu(x)]\right]. \tag{48}$$

Now let $F[U] = 0$ be any gauge-fixing condition, and we insert unity in the usual way:

$$\begin{aligned} Z &= \int DU \left\{ \Delta_{FP}[U] \int Dg\, \delta(F[g\circ U])\right\} \\ &\quad \times \exp\left[-S_W + \gamma \sum_{x,\mu}\frac{1}{2}\mathrm{Tr}U_\mu(x)\right] \\ &= \int DU\ \Delta_{FP}[U]\delta(F[U])e^{-S_W} \\ &\quad \times \int Dg\ \exp\left[\gamma \sum_{x,\mu}\frac{1}{2}\mathrm{Tr}[g^\dagger(x)U_\mu(x)g(x+\hat\mu)\right] \\ &= \int DU\ \Delta_{FP}[U]\delta(F[U])Z_{\mathrm{spin}}(\gamma, U)e^{-S_W}, . \end{aligned} \tag{49}$$

where

$$\begin{aligned} Z_{\mathrm{spin}}(\gamma, U) &= \int Dg\, e^{-S_H[U,g]}, \\ S_H[U,g] &= -\gamma \sum_{x,\mu}\frac{1}{2}\mathrm{Tr}[g^\dagger(x)U_\mu(x)g(x+\hat\mu)]. \end{aligned} \tag{50}$$

Obviously, the Higgs action S_H is again invariant under custodial transformations of the g field. The only difference, as compared to (35) and (36), is that the gauge link variables U_μ are in the $F(U) = 0$ gauge. Since the order parameter Φ for symmetry breaking in $Z_{\mathrm{spin}}(\gamma, U)$ is gauge invariant, the choice of gauge does not matter, and we recover the original formulation with $\phi(x)$ replaced by $g(x)$.

IV. SIGNIFICANCE

As we have emphasized repeatedly, the spontaneous breaking of a global subgroup of the gauge group is a gauge-dependent (or, equivalently, operator-dependent) phenomenon, and the question is whether the spontaneous breaking of a global gauge symmetry in some gauge, or seen in some Q_G operator, has any gauge-independent significance. In this section we will show that if global gauge symmetry appears to be spontaneously broken in any gauge at all, then it means that custodial symmetry is broken, and the breaking of custodial symmetry has a gauge-independent meaning. In other words, spontaneous breaking of remnant gauge symmetry in some gauge is a sufficient but not necessary condition for the spontaneous breaking of custodial symmetry. Moreover, if custodial symmetry is broken, then global gauge symmetry is spontaneously broken in at least one gauge. Taken together, custodial symmetry breaking is both a necessary condition for gauge symmetry breaking in any gauge, and a sufficient condition for the existence of global gauge symmetry breaking in some gauge. We will prove these statements below; the proof is general and applies to both Abelian and non-Abelian gauge theories. In this section we will continue to use the notation of the SU(2) gauge theory, but the arguments apply to U(1) and SU($N > 2$) gauge-Higgs theories as well.

Start with the necessary condition. Stated a little more precisely, consider any gauge condition $F(U) = 0$ which leaves unfixed a global subgroup of the gauge symmetry, and let

$$\begin{aligned} |\langle\phi\rangle_{JV}| = \frac{1}{Z}\bigg| \int DU\, D\phi\ \Delta[U]\delta[F(U)]\left(\frac{1}{V}\sum_x \phi(x)\right)e^{-S} \\ \times \exp\left[J\sum_x \mathrm{Tr}[\phi(x)]\right]\bigg| \end{aligned} \tag{51}$$

in volume V where $\Delta[U]$ is the Faddeev-Popov term. The global subgroup of the gauge symmetry is said to be spontaneously broken in this gauge if

$$|\langle\phi\rangle| = \lim_{J\to 0}\lim_{V\to\infty} |\langle\phi\rangle_{JV}| \neq 0. \tag{52}$$

The statement is that symmetry breaking of that kind is only possible if $\langle\Phi\rangle > 0$, i.e., if constituent symmetry is also spontaneously broken. This can be seen from the definition of constituent symmetry breaking. Since $\Phi_{JV}(U)$ is gauge invariant, it can of course be evaluated with or without gauge fixing, and in particular in the gauge $F(U) = 0$. Then,

$$\begin{aligned} \langle\Phi_{JV}\rangle &= \frac{1}{Z}\int DU\ \Delta[U]\delta[F(U)]e^{-S_W}Z_{\mathrm{spin}}[U]\max_\eta \frac{1}{V}\sum_x |\overline{\phi}_J[x; U, \eta]| \\ &= \frac{1}{Z}\int DU\ \Delta[U]\delta[F(U)]e^{-S_W}Z_{\mathrm{spin}}[U]\left\{\frac{1}{Z_{\mathrm{spin}}[U]}\max_\eta \frac{1}{V}\sum_x \left|\int D\phi\ \phi(x)e^{-S_H}\exp\left[J\sum_x \mathrm{Tr}[\eta^\dagger(x)\phi(x)]\right]\right|\right\} \\ &= \frac{1}{Z}\int DU\ \Delta[U]\delta[F(U)]e^{-S_W}\max_\eta \frac{1}{V}\sum_x \left|\int D\phi\ \phi(x)e^{-S_H}\exp\left[J\sum_x \mathrm{Tr}[\eta^\dagger(x)\phi(x)]\right]\right|. \end{aligned} \tag{53}$$

However, from (51)

$$\begin{aligned}|\langle\phi\rangle_{JV}| &\leqslant \frac{1}{Z}\int DU\,\Delta[U]\delta[F(U)]e^{-S_W}\left|\frac{1}{V}\sum_x\int D\phi\,\phi(x)e^{-S_H}\exp\left[J\sum_x \mathrm{Tr}[\phi(x)]\right]\right| \\ &\leqslant \frac{1}{Z}\int DU\,\Delta[U]\delta[F(U)]e^{-S_W}\frac{1}{V}\sum_x\left|\int D\phi\,\phi(x)e^{-S_H}\exp\left[J\sum_x \mathrm{Tr}[\phi(x)]\right]\right| \\ &\leqslant \frac{1}{Z}\int DU\,\Delta[U]\delta[F(U)]e^{-S_W}\max_\eta\frac{1}{V}\sum_x\left|\int D\phi\,\phi(x)e^{-S_H}\exp\left[J\sum_x \mathrm{Tr}[\eta^\dagger(x)\phi(x)]\right]\right| \\ &\leqslant \langle\Phi_{JV}\rangle.\end{aligned} \quad (54)$$

Taking first the infinite volume and then the $J \to 0$ limits, it follows that

$$\langle\Phi\rangle \geqslant |\langle\phi\rangle|. \quad (55)$$

So, although remnant gauge symmetry may or may not be broken at some point in the space of couplings, depending on the choice of gauge, we can conclude that the existence of spontaneous gauge symmetry breaking for those couplings in *some* gauge is only possible if custodial symmetry is also spontaneously broken. This means, in particular, that the custodial symmetry-breaking line must lie below the remnant gauge symmetry-breaking lines in Coulomb and Landau gauges, which is indeed what we see in Fig. 9, taken together with Fig. 7.

Moving on to the sufficient condition, let us define $\overline{\phi}_{JV}(x;U)$ as $\overline{\phi}_{JV}(x;U,\eta)$ with η chosen to maximize the right-hand side of (45). Let

$$\widehat{\phi}_{JV}(x;U) = \frac{\overline{\phi}_{JV}(x;U)}{|\overline{\phi}_{JV}(x;U)|} \quad (56)$$

and we consider the gauge

$$\widehat{\phi}_{JV}(x;U) = \mathbb{1}. \quad (57)$$

Since this condition is imposed only on the gauge field, there is obviously a remnant symmetry under those transformations which leave U invariant, and these are the elements z of the center of the gauge group, i.e., $g(x) = z$. For an SU(N) gauge group, the center subgroup is Z_N, while for the U(1) gauge group, the center is the group U(1) itself. Introducing an explicit breaking term

$$\begin{aligned}|\langle\phi\rangle| &= \lim_{J\to 0}\lim_{V\to\infty}\frac{1}{Z}\left|\int DU\,\Delta[U]\delta[\widehat{\phi}_{JV}(x;U)-\mathbb{1}]e^{-S_W}\max_\eta\int D\phi\frac{1}{V}\sum_x\phi(x)e^{-S_H}\exp\left[J\sum_x\mathrm{Tr}[\eta^\dagger(x)\phi(x)]\right]\right| \\ &= \lim_{J\to 0}\lim_{V\to\infty}\frac{1}{Z}\int DU\,\Delta[U]\delta[\widehat{\phi}_{JV}(x;U)-\mathbb{1}]e^{-S_W}\max_\eta\frac{1}{V}\sum_x\left|\int D\phi\,\phi(x)e^{-S_H}\exp\left[J\sum_x\mathrm{Tr}(\eta^\dagger(x)\phi(x))\right]\right| \\ &= \langle\Phi\rangle,\end{aligned} \quad (58)$$

where, in going from the first to the second equality, we have used the fact that in the gauge (57) the integral over ϕ is equal to a positive number times the unit element. If custodial symmetry is spontaneously broken, then $|\langle\phi\rangle| > 0$, and the remnant global gauge symmetry is also spontaneously broken.

A custodial symmetry in a non-Abelian theory is not necessarily a continuous symmetry. Let us consider a lattice version of an SU(N) gauge-Higgs theory, this time with the unimodular Higgs field in the adjoint representation. A lattice action with the correct continuum limit is [16]

$$S = -S_W - \gamma\sum_{x,\mu}\mathrm{Tr}[\Gamma(x)U_\mu(x)\Gamma^\dagger(x+\hat{\mu})U_\mu^\dagger(x)], \quad (59)$$

where $\Gamma(x)$ is an SU(N)-valued Higgs field, and S_W is the usual Wilson action. The custodial symmetry in this case is the discrete global symmetry

$$\Gamma(x) \to \Gamma'(x) = z_n\Gamma(x), \quad (60)$$

where

$$z_n = e^{2\pi i n/N}, \quad n = 0, 1, \dots, N-1 \in Z_N \quad (61)$$

and the set of elements $\{z_n\mathbb{1}\}$ constitute the center subgroup of SU(N).

V. CUSTODIAL SYMMETRY BREAKING IN THE ABELIAN HIGGS MODEL

After this excursion into non-Abelian gauge theory we return to the example relevant to superconductivity, i.e., the lattice Abelian Higgs model (3) with a double-charged Higgs field, corresponding to $q = 2$. We observe that the action is invariant under a global U(1) transformation $\phi(x) \to e^{i\alpha}\phi(x)$ of the Higgs field alone. By our definition, this is a custodial symmetry, which in the Abelian case is indistinguishable from a global gauge transformation with $\theta(x) = \alpha/2$. The spontaneous breaking of this custodial symmetry can be detected by the methods outlined above. In numerical simulations we use

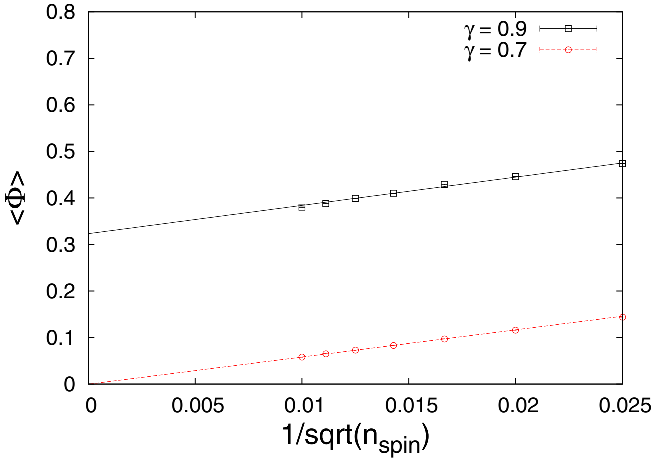


FIG. 10. The parameter $\langle\Phi_{n_{\rm spin}}\rangle$ vs $1/\sqrt{n_{\rm spin}}$ at $\beta = 0.5$. The custodial symmetry-breaking transition is at $\gamma_c \approx 0.84$. The plot displays our values of $\langle\Phi\rangle$ below ($\gamma = 0.7$) and above ($\gamma = 0.9$) the critical value. For all $\gamma < \gamma_c$, $\langle\Phi\rangle$ extrapolates to zero as $n_{\rm spin} \to \infty$, while at all $\gamma > \gamma_c$ $\langle\Phi\rangle$ extrapolates to a nonzero value.

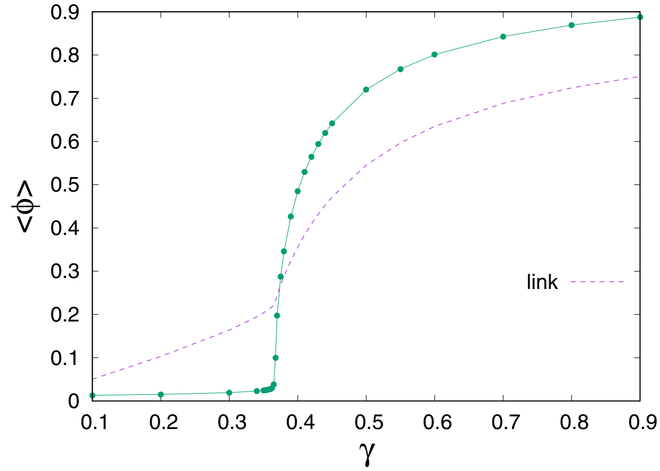


FIG. 11. Order parameter $\langle\Phi\rangle$ for custodial symmetry breaking vs γ at $\beta = 2.0$ on a 16^4 lattice volume, with $n_{\rm spin} = 6400$. Also shown (dashed line) is the corresponding data for L vs γ, already shown in Fig. 1. The thermodynamic and custodial symmetry-breaking transitions coincide at $\gamma = 0.365$.

the Monte-Carlo-within-a-Monte-Carlo approach, calculating $\Phi_{n_{\rm spin},V}[U]$ during the data-taking process using $n_{\rm spin}$ update sweeps of the ϕ field at fixed U, and averaging over the values obtained at every set of data-taking sweeps at fixed U to arrive at $\langle\Phi_{n_{\rm spin},V}[U]\rangle$. This quantity is computed at a range of $n_{\rm spin}$ on a $V = 12^4$ lattice volume, and extrapolated to $n_{\rm spin} = \infty$ by fitting the data to

$$\langle\Phi_{n_{\rm spin},V}[U]\rangle = \langle\Phi_V[U]\rangle + \frac{\rm const}{\sqrt{n_{\rm spin}}}. \tag{62}$$

Below the transition line, $\langle\Phi_V[U]\rangle = 0$, while above the line $\langle\Phi_V[U]\rangle > 0$. An example of this procedure is shown in Fig. 10, where we present data for $\langle\Phi_{n_{\rm spin},V}[U]\rangle$ vs $n_{\rm spin}$ at $\beta = 0.5$, at γ values above ($\gamma = 0.9$) and below ($\gamma = 0.7$) the transition.

At points where the custodial symmetry transition coincides with the thermodynamic transition, there is an abrupt rise in $\langle\Phi_{n_{\rm spin}}\rangle$ even at moderate values of $n_{\rm spin}$ as illustrated in Fig. 11, where we plot $\langle\Phi\rangle$ vs γ at $\beta = 2$ on a 16^4 lattice. Also shown in this figure, as a dashed line, is the corresponding data for the average link variable L, already displayed in Fig. 1. It is clear that the thermodynamic transition (the "kink") and custodial breaking transition, signaled by a sudden rise in $\langle\Phi\rangle$, occur at the same point, namely, $\gamma = 0.365$ at $\beta = 2$.

The custodial symmetry transition line in the β-γ plane is shown in Fig. 12. Note that this transition line lies well below the Landau and Coulomb transition lines. However, there are an infinite number of gauges which preserve some global remnant gauge symmetry; no doubt some of them have transition lines closer to the custodial transition. In fact, as we have shown in Sec. IV, there exists one gauge of this type whose transitions coincide with the custodial symmetry transition line.

A. Absence of Goldstone excitations

The reason that spontaneous breaking of custodial symmetry does not result in physical gapless excitations is essentially the same reason given long ago [17], when a similar question was raised regarding the spontaneous breaking of (remnant) gauge symmetries. In the case of an Abelian theory, it is obvious that the same reasoning must apply because, in that case, the custodial symmetry is identical to the remnant gauge symmetry $\theta(x) = \theta$.

In a little more detail, spontaneous breaking of custodial symmetry in a given $Z_{\rm spin}(U)$ for some U may very well be associated with gapless excitations. However, there is no reason to believe that such excitations appear in correlation functions associated with physical states. For instance, if custodial symmetry is broken in $Z_{\rm spin}(U)$, with order parameter $\phi(x)$, then for fixed U there might be a long-range part to a connected correlator such as

$$\overline{\phi(x)\phi(y)} - \overline{\phi(x)} \times \overline{\phi(y)}. \tag{63}$$

Such a correlator, however, being locally gauge noninvariant, would necessarily vanish in the full theory, i.e.,

$$\langle\phi(x)\phi(y)\rangle - \langle\phi(x)\rangle\langle\phi(y)\rangle = 0. \tag{64}$$

In order to apply the Goldstone theorem to custodial symmetry, and restrict to physical excitations, it is necessary to fix to a gauge which eliminates extraneous degrees of freedom, leaving only physical degrees of freedom. Examples are Coulomb gauge and axial gauge. In such gauges it is necessary to impose Gauss's law as an operator identity, and solve for E_L^2 in the Hamiltonian. Gauges of this type are Lorentz noninvariant, and the E_L^2 term gives rise to long-range interactions in the Hamiltonian. Nonlocal terms in general violate one of the assumptions of the Goldstone theorem. This observation was made originally in reference to the breaking of remnant gauge symmetries [17], but it applies equally well to the current associated with any continuous custodial symmetry. The conclusion is that spontaneous breaking of a global custodial symmetry does not necessarily imply gapless physical excitations, which might have been expected from the Goldstone theorem.

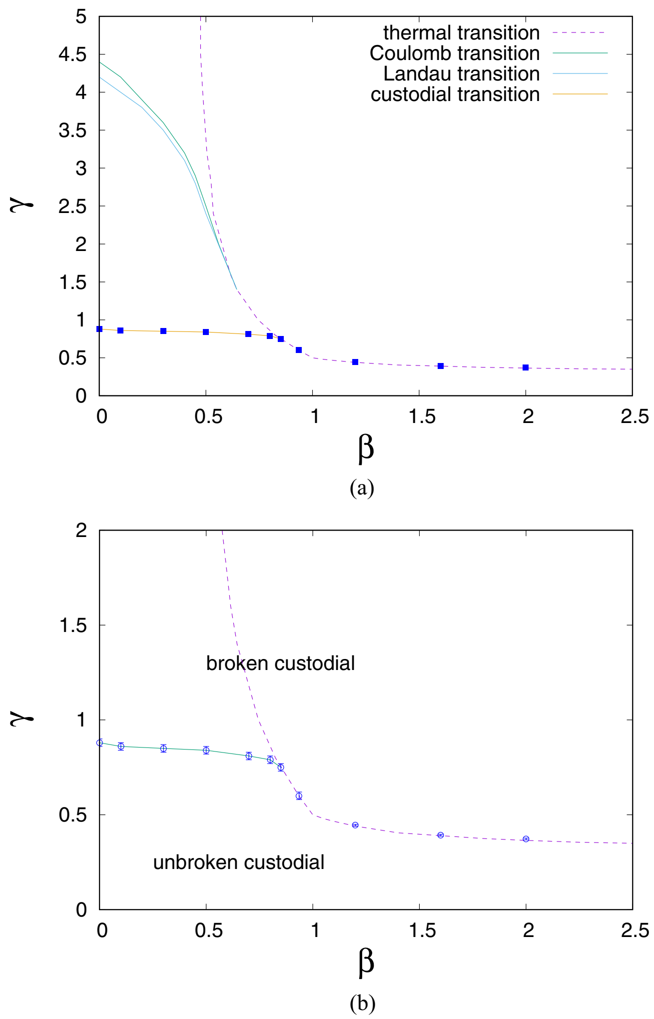


FIG. 12. Custodial symmetry-breaking transition points. The dashed line is a line of thermodynamic transition shown in Fig. 3, which coincides with custodial symmetry breaking at $\beta > 0.85$. For lower values of β, the custodial symmetry-breaking transition occurs in the region labeled "conf" in Fig. 3. (a) Also shown (solid lines) are the transition lines for Coulomb and Landau remnant symmetry breaking in the confined region. (b) Closeup of the custodial transition region.

B. C vs S_c confinement

Custodial symmetry, and also remnant gauge symmetry in Coulomb and Landau gauges, have transition lines in the confinement region of the $q = 2$ phase diagram. Usually, $\langle\phi\rangle \neq 0$ is associated with a Higgs phase, so how can this happen in a confined phase? In this case, it is helpful to consider unitary gauge at large γ, and write the link variables in the form $U_\mu(x) = \widetilde{U}_\mu(x)Z_\mu(x)$, where $\text{Re}[\widetilde{U}_\mu(x)] > 0$ and $Z_\mu(x) = \pm 1$. As $\gamma \to \infty$, then $\widetilde{U}_\mu(x) \to 1$, and the Abelian Higgs model goes over to Z_2 lattice gauge theory, which has a confined and unconfined phase. But what is confined, in the confined phase, are $q = 1$ test charges, i.e., sources with ± 1 units of electric charge. Test charges with $q = 2$ are insensitive to the $Z_\mu(x)$ degrees of freedom, and couple only to $\widetilde{U}_\mu(x)$. Away from unitary gauge, the remnant gauge symmetry which is broken spontaneously by $\langle\phi\rangle \neq 0$ is global $U(1)/Z_2$, and from the point of view of $q = 2$ sources the theory is actually in a Higgs phase. This raises the question of the nature of the transition, as seen by $q = 2$ sources, from the confined phase into the Higgs phase, since by criteria such as Wilson loops and Polyakov lines the $q = 2$ sources are not really confined anywhere in the phase diagram.

We have addressed the same question in Ref. [5], in the context of SU(2) gauge-Higgs theory with the Higgs field in the fundamental representation of the gauge group. In this theory, as in any gauge theory with matter in the fundamental representation (such as QCD), Wilson loops fall off asymptotically with a perimeter law, and Polyakov lines have a nonzero vacuum expectation value. Then, what is meant by the word "confinement" in such theories? A common answer is that confinement means that only color-singlet particle states appear in the asymptotic spectrum, a property which we will refer to as "C confinement." It is well known that this property holds not only in confinementlike region of an SU(2) gauge-Higgs theory, but also deep in the Higgs regime [3,4,18,19]. Nevertheless, there seems to be a qualitative difference between these regions since in the confinementlike region there are color electric flux tube formation, linear Regge trajectories, and a linear potential up to string breaking, as in QCD, while in the Higgs region there is no electric flux tube formation in any distance regime, no linear Regge trajectories, and only Yukawa forces among particles.

In a pure SU(2) gauge theory, the word "confinement" includes but goes beyond the property of C confinement. Certainly, the asymptotic spectrum consists only of color singlets, i.e., glueballs. But, it also has the property that the energy $E(R)$ above the vacuum energy, of any physical state containing a static quark-antiquark pair, is bounded from below by a linear potential. In other words, let $V_{ab}(x, y; A)$ be any functional of the gauge field A which transforms covariantly under the gauge group, and we consider physical states of the form

$$\Psi_V = \overline{q}^a(\mathbf{x})V_{ab}(\mathbf{x}, \mathbf{y}; A)q^b(\mathbf{y})\Psi_0, \tag{65}$$

where Ψ_0 is the ground state. Let $E_V(R) = \langle\Psi_V|H - E_0|\Psi_V\rangle$ be the expectation value of energy, above the vacuum energy E_0, in state Ψ_V, where $R = |\boldsymbol{x} - \boldsymbol{y}|$. We define "separation-of-charge" confinement, or "S_c" confinement for short, to mean that $E_V(R)$ is bounded from below, asymptotically, by a linear potential

$$\lim_{R\to\infty} \frac{dE_V}{dR} > \sigma \tag{66}$$

for *any* choice of V. Pure SU(N) gauge theories in $D \leqslant 4$ dimensions certainly have this property. We have suggested in [5] that this same definition extends to gauge theories with matter fields, with the essential requirement that $V(x, y; A)$ depends only on the gauge field, and not on the matter fields. This restriction essentially tests whether the dynamics would form a flux tube between sources if we exclude string breaking by matter fields. In the cited reference we have shown that there must exist a transition between the C and S_c confinement regions, and we have also computed, in SU(2) and SU(3) gauge-Higgs theories, the line of custodial symmetry breaking. Our conjecture, for which we have presented some evidence, is that the S_c-to-C confinement transition, and the custodial symmetry-breaking transition, coincide.

That is also our conjecture regarding the custodial symmetry-breaking transition inside the confinement phase of the $q = 2$ Abelian Higgs model, with this modification: For the $q = 2$ theory we have confinement of single charged ($q = 1$) sources by a linear potential whenever the Z_2 global symmetry defined in Sec. II C is unbroken, which is the entire region labeled "conf" in Fig. 3. The C-vs-S_c transition in the $q = 2$ theory concerns the nature of the confined phase for *double*-charged ($q = 2$) objects, which are insensitive to the Z_2 degrees of freedom. Double-charged Wilson loops have a perimeter-law falloff and double-charged Polyakov lines are nonzero inside the confined phase, as in SU(N) gauge theories (such as QCD) with matter in the fundamental representation. We can define S_c confinement for double-charged sources in the same way: Consider operators $V(\mathbf{x}, \mathbf{y}; A)$, and $q = 2$ matter fields $\psi(x)$ which transform under a gauge transformation $g(x) = \exp[i\theta(x)]$ as

$$\begin{aligned} V(\mathbf{x}, \mathbf{y}; A) &\to e^{2i\theta(\mathbf{x})} V(\mathbf{x}, \mathbf{y}; A) e^{-2i\theta(\mathbf{y})}, \\ \psi(x) &\to e^{2i\theta(\mathbf{x})}\psi(x), \\ \Psi_V &= \overline{\psi}(\mathbf{x}) V(\mathbf{x}, \mathbf{y}; A)\psi(\mathbf{y})\Psi_0. \end{aligned} \tag{67}$$

Then, the theory is S_c confining when the condition (66) is satisfied. As in the non-Abelian theory, our conjecture is that custodial symmetry breaking at small β coincides with the transition from S_c to C confinement for $q = 2$ charges.

Our point is this: From the standpoint of $q = 2$ charged matter in a $q = 2$ Abelian Higgs theory, the transition from a confined phase (which we define as S_c confinement) to a Higgs phase need not coincide everywhere with the transition from a confined to a Higgs phase for $q = 1$ test charges. What we are proposing is that the spontaneous breaking of custodial symmetry is a gauge-invariant criterion which sets the boundary of the Higgs region, as seen by $q = 2$ matter in the $q = 2$ Abelian Higgs theory.

We should finally note that the custodial and remnant gauge symmetry-breaking lines in the confinement region of the $q = 2$ gauge-Higgs model, and also in the SU(2) gauge-Higgs theory, are not lines of thermodynamic transition. As we have just argued, this does not imply irrelevance. Recall that there are other physically meaningful transitions in statistical systems which, like custodial and remnant symmetry breaking, are not necessarily associated with thermodynamic transitions. We here have in mind the geometric transition lines, also known as Kertesz lines, in Ising and Potts models, which are associated with percolation transitions [20,21].

VI. CONCLUSIONS

In this paper we have pointed out that "spontaneous breaking of gauge symmetry" is an ambiguous concept, and we have proposed that it is spontaneous breaking of custodial symmetry which characterizes the Higgs phase. The ambiguity of spontaneous gauge symmetry breaking is due to the fact that local gauge symmetries cannot break spontaneously, as we know from the Elitzur theorem, which means that only a global subgroup of the gauge symmetry can break spontaneously, and this is visible only in a gauge which leaves this global subgroup unfixed. This means that the order parameter for spontaneous gauge symmetry breaking is gauge dependent. As shown previously for SU(2) gauge-Higgs theory, and as shown here in the lattice Abelian Higgs model, spontaneous gauge symmetry breaking can occur at different places in the phase diagram in different gauges, and in some gauges it may even disappear entirely. This is true also for locally gauge-invariant operators whose construction is based on an implicit gauge choice.

Adopting a term from the electroweak theory, we have defined "custodial symmetry" to be (i) a group of transformations of the matter fields which does not transform the gauge field, and (ii) a symmetry for which there is no gauge-invariant order parameter, in the sense that any operator which transforms under the custodial symmetry also transforms under the gauge group. The custodial symmetry group and global gauge transformations share symmetry transformations which belong to the center of the gauge group, which for U(1) gauge theory is the group itself. Despite the absence of a gauge-invariant order parameter, we have shown here how spontaneous breaking of the custodial symmetry can be defined and observed in a gauge-invariant manner, without recourse to either an explicit or implicit choice of gauge.

The relation of custodial symmetry breaking to gauge symmetry breaking is as follows: First, custodial symmetry breaking is a necessary condition for gauge symmetry breaking in any particular gauge. Second, custodial symmetry is a sufficient condition for the existence of some gauge in which the gauge symmetry breaks spontaneously. If we identify the Anderson-Brout-Englert-Higgs mechanism with the existence of spontaneous gauge symmetry breaking in some gauge, then this mechanism occurs if and only if custodial symmetry is spontaneously broken. In the $q = 2$ Abelian Higgs model, we have seen numerically that custodial symmetry breaks along the line separating the massless and Higgs phases.

In some regions of the phase diagram, gauge symmetries and custodial symmetry can break without a corresponding thermodynamic transition, as is the case for the geometric (Kertesz) transition in the Ising and Potts models. We believe that custodial symmetry breaking in the absence of a thermodynamic transition is related to what we have elsewhere described as the transition between separation-of-charge confinement and color confinement [5]. This correspondence is so far a conjecture, and calls for further investigation.

ACKNOWLEDGMENTS

We thank A. Beekman and E. Fradkin for helpful correspondence. This work is supported by the US Department of Energy under Grant No. DE-SC0013682.

APPENDIX: GOLDSTONE MODES AND THE SUPERCONDUCTOR PHASE

One reason that the Goldstone theorem can fail is the presence of nonlocal interactions, which show up in the Hamiltonian in any physical gauge. The nonlocality invalidates one of the assumptions underlying the standard derivation of the Goldstone theorem [17]. There is, however, a simple derivation of the Goldstone theorem, which can be found in standard textbooks [22], and which seems oblivious to the presence or absence of nonlocal interactions. In this Appendix we seek to understand where that simple argument goes wrong, in the case of the BCS Hamiltonian.

The argument goes as follows: Suppose we have an operator Q (a "charge" operator) which commutes with the Hamiltonian, and that the set of transformations $\exp[i\theta Q]$ is a U(1) symmetry group (possibly a subgroup of a larger symmetry), and $Q\Psi_0 \neq 0$. We also suppose that Q is associated with a conserved current and can be expressed as the spatial integral of a charge density

$$Q = \int d^3x\, J_0(x). \tag{A1}$$

Because $[H, Q] = 0$, the state $Q\Psi_0$ has the same energy as Ψ_0, namely, the ground-state energy E_0. Now, consider a state with momentum $\boldsymbol{k}$:

$$|\boldsymbol{k}\rangle = \int d^3x\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} J_0(x)|\Psi_0\rangle. \tag{A2}$$

As $\boldsymbol{k} \to 0$, the energy of this state converges to the energy of $Q\Psi_0$, which is E_0. The conclusion is that the excitation energy $E_{ex}(k)$ of state $|\boldsymbol{k}\rangle$ above the ground-state energy vanishes as $\boldsymbol{k} \to 0$, i.e., there exist gapless (or, in particle physics language, massless) excitations. This is the Goldstone theorem.

Perhaps surprisingly, this argument correctly predicts the existence of gapless excitations in the normal state, which is generally not considered to be a state of broken symmetry. It may be of interest to see explicitly how this works out in the normal state, and how that conclusion is evaded in superconducting state. In the present case, it is the number operator

$$N = \int d^3x\, c^\dagger_\sigma(x) c_\sigma(x) \tag{A3}$$

which commutes with the Hamiltonian, and in fact eN is the electric charge operator. It is easy to see that

$$[e^{i\theta N}, c^\dagger_\sigma(x)] = e^{i\theta} c^\dagger_\sigma(x) \tag{A4}$$

and, as a consequence, operating on the BCS ground state

$$\begin{aligned} e^{i\theta N}\Psi^0_{\rm BCS} &= e^{i\theta N}\prod_k (u_k + v_k c^\dagger_\uparrow(\boldsymbol{k}) c^\dagger_\downarrow(-\boldsymbol{k}))|0\rangle \\ &= \prod_k (u_k + v_k e^{2i\theta} c^\dagger_\uparrow(\boldsymbol{k}) c^\dagger_\downarrow(-\boldsymbol{k}))|0\rangle, \end{aligned} \tag{A5}$$

which we recognize as a global U(1) transformation $\theta(x) = \theta$ acting on the ground state. As usual,

$$\begin{aligned} u(k) &= \sqrt{\frac{1}{2}\left(1 + \frac{\epsilon_k}{E_k}\right)}, \quad v(k) = \sqrt{\frac{1}{2}\left(1 - \frac{\epsilon_k}{E_k}\right)}, \\ E_k &= \sqrt{\epsilon_k^2 + |\Delta_k|^2}, \quad \epsilon_k = \frac{k^2}{2m} - \epsilon_F, \\ 1 &= \frac{g}{V}\sum_k{}' \frac{1}{2\sqrt{\epsilon_k^2 + |\Delta|^2}}, \\ \sum_k{}'[\ldots] &\equiv \sum_k \theta(\omega_D - |\epsilon_k|)[\ldots], \end{aligned} \tag{A6}$$

and

$$\Delta_k = \begin{cases} \Delta & |\epsilon_k| < \omega_D, \\ 0 & |\epsilon_k| \geqslant \omega_D, \end{cases} \tag{A7}$$

where ω_D, ϵ_F, Δ are the Debye frequency, Fermi energy, and gap, respectively.

Let us ignore photon and ion degrees of freedom, and consider only the usual BCS Hamiltonian

$$\begin{aligned} H &= \int d^3x\, c^\dagger_\sigma(x)\left[\frac{1}{2m}(-\nabla^2) - \epsilon_F\right] c_\sigma(x) \\ &\quad - \frac{g}{V}\sum_k{}' \sum_{k'}{}' c^\dagger_\uparrow(k) c^\dagger_\downarrow(-k) c_\downarrow(k') c_\uparrow(-k') - \mathscr{E}_{grd}, \end{aligned} \tag{A8}$$

where $\mathscr{E}_{grd}$ is the ground-state energy, so that $H|\Psi_0\rangle = 0$. H is still invariant under the global transformations $c_\sigma(x) \to e^{i\theta} c_\sigma(x)$, $c^\dagger_\sigma(x) \to e^{-i\theta} c^\dagger_\sigma(x)$, and commutes with the generator of those transformations, i.e., the number operator N:

$$N = \int d^3x\, J_0(x), \quad J_0(x) = c^\dagger_\sigma(x) c_\sigma(x). \tag{A9}$$

Let us define

$$\begin{aligned} N_q &= \int d^3x\, c^\dagger_\sigma(x) c_\sigma(x) e^{-i\boldsymbol{q}\cdot\boldsymbol{x}} \\ &= \sum_k c^\dagger_\sigma(\boldsymbol{k}) c_\sigma(\boldsymbol{k} + \boldsymbol{q}), \end{aligned} \tag{A10}$$

so in this case

$$|q\rangle = N_q|\Psi_0\rangle. \tag{A11}$$

Note that since N_q cannot change electron number, any excitations above the ground state must correspond to the creation of a particle-hole pair. Introducing the usual Bogoliubov quasiparticle operators

$$\begin{aligned} c_\uparrow(\boldsymbol{k}) &= u_k a_\uparrow(\boldsymbol{k}) - v_k a^\dagger_\downarrow(-\boldsymbol{k}), \\ c_\downarrow(-\boldsymbol{k}) &= v_k a^\dagger_\uparrow(\boldsymbol{k}) + u_k a_\downarrow(-\boldsymbol{k}), \\ c^\dagger_\uparrow(\boldsymbol{k}) &= u_k a^\dagger_\uparrow(\boldsymbol{k}) - v_k a_\downarrow(-\boldsymbol{k}), \\ c^\dagger_\downarrow(-\boldsymbol{k}) &= v_k a_\uparrow(\boldsymbol{k}) + u_k a^\dagger_\downarrow(-\boldsymbol{k}), \end{aligned} \tag{A12}$$

with the property that $a_\sigma(k)\Psi_0 = 0$, we have for $\boldsymbol{q} \neq 0$

$$\begin{aligned} |q\rangle &= \sum_k u_k v_{k+q}\{a^\dagger_\downarrow(\boldsymbol{k}) a^\dagger_\uparrow(-\boldsymbol{k} - \boldsymbol{q}) \\ &\quad - a^\dagger_\uparrow(\boldsymbol{k}) a^\dagger_\downarrow(-\boldsymbol{k} - \boldsymbol{q})\}|\Psi_0\rangle, \end{aligned} \tag{A13}$$

and from here on our notational convention for momentum subscripts is that $u_{k\pm q}$, $v_{k\pm q}$, $E_{k\pm q}$ means $u_{|\boldsymbol{k}\pm\boldsymbol{q}|}$, $v_{|\boldsymbol{k}\pm\boldsymbol{q}|}$, $E_{|\boldsymbol{k}\pm\boldsymbol{q}|}$, respectively. We find the norm

$$\langle q|q\rangle = 2\sum_k \left(u_k^2 v_{k+q}^2 + u_k u_{k+q} v_k v_{k+q}\right), \tag{A14}$$

and then evaluate $\langle q|H|q\rangle$ in the mean field approximation, replacing H in (A8) by

$$H_{mf} = \sum_k E_k a^\dagger_\sigma(k) a_\sigma(k), \tag{A15}$$

which leads, in this approximation, to

$$\mathscr{E}_q = \frac{\langle q|H_{mf}|q\rangle}{\langle q|q\rangle} = \frac{\sum_k (E_k + E_{k+q})\big(u_k^2 v_k^2 + u_k u_{k+q} v_k v_{k+q}\big)}{\sum_k \big(u_k^2 v_{k+q}^2 + u_k u_{k+q} v_k v_{k+q}\big)}. \quad \text{(A16)}$$

Now, in the normal phase $\Delta_k = 0$, we have $u_k v_k = 0$ for all $\boldsymbol{k}$, and $u_k v_{k+q} \neq 0$ only for $\boldsymbol{k}$ and $\boldsymbol{k}+\boldsymbol{q}$ on opposite sides of the Fermi surface. Then, as $q \to 0$, the sum over $\boldsymbol{k}$ is nonzero only in the immediate region of the Fermi surface, where $E_k = 0$. This means that $\mathscr{E}_q \to 0$ as $\boldsymbol{q} \to 0$, i.e., there are gapless excitations in this phase which follow from application of the Goldstone argument, whether or not one cares to describe this case as a phase of spontaneously broken symmetry.

In the superconducting phase $\Delta_k \neq 0$, the situation is different. In this case, $u_k v_k$ and $u_k v_{k+q}$ are nonzero, as $\boldsymbol{q} \to 0$, for k roughly in the range $|\epsilon_k| < \omega_D$, and in this range $E_k, E_{k+q} > \Delta$. Hence, $\mathscr{E}_q \approx 2\Delta$ as $q \to 0$, and there are no gapless excitations. Excited states (quasiparticle pairs) have a minimum energy of 2Δ.

But, this raises a question, since at $q = 0$ *exactly* we must have $\mathscr{E}_q = 0$. This is because $|q = 0\rangle = N|\Psi_0\rangle$, and since $[H, N] = 0$ it must be that $|q = 0\rangle$ has the same energy as the ground state, i.e., $\mathscr{E}_0 = 0$, not $\mathscr{E}_0 \approx 2\Delta$. The apparent paradox is resolved by the realization that exactly at $q = 0$ there is an additional contribution to N_q which does not annihilate the ground state, namely,

$$\sum_k v_k^2 a_\sigma(k) a_\sigma^\dagger(k) = \sum_k v_k^2 [2 + a_\sigma^\dagger(k) a_\sigma(k)]. \quad \text{(A17)}$$

Redoing the calculation including these contributions, we have for the norm

$$\langle q = 0|q = 0\rangle = 4\left(\sum_k v_k^2\right)^2 + 4\sum_k u_k^2 v_k^2. \quad \text{(A18)}$$

The first term on the right-hand side is proportional to the square of the number of electrons in the system, i.e., to the square of the volume, while the second term grows only linearly with volume, and in addition only momenta in the neighborhood of the Fermi surface contribute to the second sum. Therefore, up to $O(1/V)$ corrections,

$$\langle q = 0|q = 0\rangle = 4\left(\sum_k v_k^2\right)^2. \quad \text{(A19)}$$

Then, since $H_{mf}|\Psi_0\rangle = 0$, we have

$$\mathscr{E}_0 = \frac{\langle q = 0|H_{mf}|q = 0\rangle}{\langle q = 0|q = 0\rangle} = \frac{\sum_k 2E_k u_k^2 v_k^2}{\big(\sum_k v_k^2\big)^2} = 0 + O(1/V), \quad \text{(A20)}$$

where the last line follows since the numerator in the second line is $O(V)$, while the denominator is $O(V^2)$. The fact that $\mathscr{E}_0$ is not exactly zero, but differs from zero by a term of order $1/V$, can be attributed to the mean field approximation, which in the BCS case is also only accurate up to corrections of this order.

So, we have seen that the textbook argument [22] can be applied to both the normal and superconducting phases. The normal phase has gapless excitations in accordance with this argument. The superconducting phase, however, evades this conclusion in an interesting way, via a discontinuity in $\mathscr{E}_q$ (the energy of the low-momentum $|q\rangle$ state) precisely at $q = 0$.[2]

The superconducting and normal phases are of course distinguished by the expectation value of the Cooper pair operator $c_\uparrow^\dagger(k)c_\downarrow^\dagger(-k)$, which vanishes in the normal phase and is nonzero in the superconducting phase.[3] However, the simple model described by the Hamiltonian in (A8) has no coupling to gauge fields, and no local gauge invariance. In a gauge theory, order parameters such as the Cooper pair creation operator in the BCS theory, or the charged scalar field in the Ginzburg-Landau effective theory, transform under local gauge transformations. Hence, their expectation values vanish unless either (i) the gauge is fixed or (ii) we employ a construction which is equivalent to gauge fixing. As we will have seen in Sec. II, this introduces an ambiguity, in the sense that the vanishing or finiteness of the order parameter turns out to be gauge dependent.

[2]In an insulator the proof is evaded by the fact that there are no small-q particle-hole excitations near the Fermi surface, as N_q in (A10) annihilates the ground state for small q. Hence, there is no smooth limit to $q = 0$.

[3]Actually, this operator also vanishes in a system with a definite number of electrons. In that case, one may consider instead correlators such as $\langle c_\uparrow(x)c_\downarrow(x)c_\uparrow^\dagger(y)c_\downarrow^\dagger(y)\rangle$ in the limit of large separation $|x - y|$, which would still vanish in the normal phase and be nonzero in the superconducting phase.

[1] S. Elitzur, Phys. Rev. D **12**, 3978 (1975).
[2] W. Caudy and J. Greensite, Phys. Rev. D **78**, 025018 (2008).
[3] K. Osterwalder and E. Seiler, Ann. Phys. **110**, 440 (1978).
[4] E. H. Fradkin and S. H. Shenker, Phys. Rev. D **19**, 3682 (1979).
[5] J. Greensite and K. Matsuyama, Phys. Rev. D **98**, 074504 (2018).
[6] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.100.184513 for a discussion of the ground state and order parameters in the microscopic theory, which includes Refs. [23–25].
[7] C. Itzykson and J. M. Drouffe, *Statistical Field Theory, Vol. 1*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, 1989).
[8] A. M. Schakel, *Boulevard of Broken Symmetries: Effective Field Theories of Condensed Matter* (World Scientific, Singapore, 2008).

[9] J. van Wezel and J. van den Brink, Phys. Rev. B **77**, 064523 (2008).
[10] T. Kennedy and C. King, Commun. Math. Phys. **104**, 327 (1986).
[11] J. Ranft, J. Kripfganz, and G. Ranft, Phys. Rev. D **28**, 360 (1983).
[12] J. E. Mandula and M. Ogilvie, Phys. Lett. B **248**, 156 (1990).
[13] A. Maas, Prog. Part. Nucl. Phys. **106**, 132 (2019).
[14] S. Willenbrock, Symmetries of the standard model, in *Proceedings of TASI 2004, Boulder, Colorado, USA, June 6-July 2, 2004*, edited by J. Terning, C. Wagner, and D. Zeppenfeld (World Scientific, 2006), arXiv:hep-ph/0410370.
[15] S. Weinberg, *The Quantum Theory of Fields*, Vol. 2 (Cambridge University Press, Cambridge, 2013).
[16] J. M. Drouffe, J. Jurkiewicz, and A. Krzywicki, Phys. Rev. D **29**, 2982 (1984).
[17] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Adv. Part. Phys. **2**, 567 (1968); Phys. Rev. Lett. **13**, 585 (1964).
[18] J. Frohlich, G. Morchio, and F. Strocchi, Nucl. Phys. B **190**, 553 (1981).
[19] G. 't Hooft, NATO Sci. Ser. B **59**, 117 (1980).
[20] J. Kertesz, Phys. A (Amsterdam) **161**, 58 (1989).
[21] P. Blanchard, D. Gandolfo, L. Laanait, J. Ruiz, and H. Satz, J. Phys. A: Math. Gen. **41**, 085001 (2008).
[22] A. Zee, *Quantum Field Theory in a Nutshell*, 2nd ed. (Princeton University Press, Princeton, NJ, 2010); M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, Cambridge, 2014).
[23] Y. Nambu, Phys. Rev. **117**, 648 (1960).
[24] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
[25] C. Timm, Theory of Superconductivity, https://www.physik.tu-dresden.de/~timm/personal/teaching/thsup_w11/Theory_of_Superconductivity.pdf.