

What symmetry is actually broken in the Higgs phase of a gauge-Higgs theory?

Jeff Greensite and Kazue Matsuyama

*Physics and Astronomy Department, San Francisco State University,
San Francisco, California 94132, USA*

 (Received 8 May 2018; revised manuscript received 21 August 2018; published 5 October 2018)

In $SU(N)$ gauge-Higgs theories, with a single Higgs field in the fundamental representation, there exists in addition to the local gauge symmetry a global $SU(2)$ symmetry, at $N = 2$, and a global $U(1)$ symmetry, for $N \neq 2$. We construct a gauge-invariant order parameter for the breaking of these global symmetries in the Higgs sector and calculate numerically the transition lines, in coupling-constant space, for $SU(2)$ and $SU(3)$ gauge theories with unimodular Higgs fields. The order parameter is nonlocal, and, therefore, its nonanalyticity does not violate the theorem proved by Osterwalder and Seiler. We then show that there exists a transition, in gauge-Higgs theories, between two types of confinement: ordinary color neutrality in the Higgs region and a stronger condition, which we have called “separation-of-charge confinement”, in the confinement region. We conjecture that the symmetry-breaking transition coincides with the transition between these two physically different types of confinement.

DOI: [10.1103/PhysRevD.98.074504](https://doi.org/10.1103/PhysRevD.98.074504)

I. INTRODUCTION

Contrary to statements found in some textbooks, a local gauge symmetry cannot be broken spontaneously, as shown long ago by Elitzur [1]. In certain gauges, there are remnant global symmetries which *can* break spontaneously, but the locations of the corresponding transition lines are gauge dependent [2], which makes a physical interpretation of such transitions dubious. Of course, in a gauge-Higgs theory with the scalar field in the fundamental representation, there is a confinementlike region analogous to QCD, in which one finds color electric flux tube formation, Regge trajectories, and a linear static quark potential followed by string breaking. There is also a Higgs region with no flux tube formation, no Regge trajectories, and only Yukawa forces between static sources. It was shown many years ago by Osterwalder and Seiler [3], whose work was further elucidated by Fradkin and Shenker [4], that there is no thermodynamic transition which entirely isolates the Higgs regime from the confinementlike regime, meaning that the free energy is analytic along some path between any two points in coupling constant space. The implication is that, in the absence of a massless phase, there can be no transition from a color neutral to a color charged spectrum of asymptotic states. The gauge-invariant composite

operators which create color-neutral physical particles in the electroweak theory were first written down by Fröhlich, Morchio, and Strocchi [5] and by 't Hooft [6].

All these facts appear to imply that, in the absence of a gauge choice, there is no such thing as spontaneous symmetry breaking in the context of the Brout-Englert-Higgs (BEH) mechanism, and no gauge-invariant order parameter which could detect such a breaking. On the other hand, it is well known that in $SU(2)$ gauge-Higgs theory there exists a global $SU(2)$ symmetry, distinct from the local gauge symmetry, and it was likewise pointed out by Maas *et al.* [7] that for $SU(N > 2)$ gauge-Higgs theories the additional symmetry is global $U(1)$. But although global symmetries can break spontaneously, the absence of massless Goldstone excitations would seem to rule out that possibility. In this article, we point out that global symmetries in the Higgs sector can break in the Higgs sector, in the sense explained below, without introducing Goldstone particles in the full theory. We will construct a gauge-invariant order parameter which is sensitive to these symmetry breakings, and map out the transition line in coupling constant space for $SU(2)$ and $SU(3)$ gauge-Higgs theories with a single unimodular Higgs field.

This raises the question of the physical distinction between the symmetric and broken phases of a gauge-Higgs theory. In a recent article [8], we have suggested that gauge theories with matter fields in the fundamental representation may satisfy a confinement criterion which is stronger than the usual condition of a color-neutral spectrum. This stronger condition, which is a generalization of the Wilson area law criterion to gauge + matter theories, is

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

called “separation of charge” or S_c confinement, although its existence beyond pure gauge theory was only conjectured in Ref. [8]. In the present article, we show that S_c confinement actually does exist in at least some region of the gauge-Higgs phase diagram, which implies the existence of a transition between the stronger and weaker confinement phases. It is, therefore, natural to suppose, although we do not prove, that the symmetry-breaking transition in gauge-Higgs theories, which we have located here for the $SU(2)$ and $SU(3)$ gauge groups, corresponds to a transition between these two, physically distinct, types of confinement.

We should note that other criteria for confinement with matter fields can be found in the literature, namely the Kugo-Ojima criterion [9], nonpositivity or unphysical pole structure in quark/gluon propagators (an early reference is [10]), and the Fredenhagen-Marcu proposal [11]. The first two of these proposals rely on BRST symmetry, which is dubious at the nonperturbative level, while the Fredenhagen-Marcu criterion only distinguishes between massless and massive phases, rather than between Higgs and confinement. For a more detailed critique, see Sec. V of [8].

II. $SU(2)$ GAUGE-HIGGS THEORY

The symmetry of an $SU(2)$ gauge-Higgs theory with a single Higgs doublet is $SU(2)_{\text{gauge}} \times SU(2)_{\text{global}}$. The extra global symmetry is easiest to see by mapping the Higgs doublet (which we take, for simplicity, to be unimodular $|\phi| = 1$), onto an $SU(2)$ group element

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \Rightarrow \phi = \begin{bmatrix} \phi_2^* & \phi_1 \\ -\phi_1^* & \phi_2 \end{bmatrix}, \quad (1)$$

and the action can then be written in the form

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

which has the following invariance

$$\begin{aligned} U_\mu(x) &\rightarrow L(x) U_\mu(x) L^\dagger(x + \hat{\mu}) \\ \phi(x) &\rightarrow L(x) \phi(x) R, \end{aligned} \quad (3)$$

where $L(x) \in SU(2)_{\text{gauge}}$ is a local gauge transformation, while $R \in SU(2)_{\text{global}}$ is a global transformation.

If we choose a gauge (e.g., unitary gauge) in which the Higgs field acquires a vacuum expectation value (VEV)

$$\langle \phi \rangle = \begin{bmatrix} v & 0 \\ 0 & v \end{bmatrix}, \quad (4)$$

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

$$SU(2)_{\text{gauge}} \times SU(2)_{\text{global}} \rightarrow SU(2)_D, \quad (5)$$

corresponding to transformations

$$\begin{aligned} L(x) &= R^\dagger = G \\ \phi(x) &\rightarrow G \phi(x) G^\dagger, \quad U_\mu(x) \rightarrow G U_\mu(x) G^\dagger. \end{aligned} \quad (6)$$

Transformations in this diagonal subgroup, which is known as the group of “custodial symmetry,” preserve the VEV of ϕ . Custodial symmetry has a role to play in the phenomenology of the electroweak interactions, and is reviewed in many places, e.g., [12–14]. Here, however, we would like to focus on the R -transformations belonging to $SU(2)_{\text{global}}$.¹

We have already noted that characterizing the Brout-Englert-Higgs mechanism as a spontaneous breaking of gauge symmetry due to the nonzero VEV of ϕ is rather misleading, given that

- (i) $\langle \phi \rangle = 0$ at all β, γ in the absence of a gauge choice;
- (ii) $\langle \phi \rangle \neq 0$ at all β, γ in unitary gauge²;
- (iii) In other gauges, $\langle \phi \rangle$ may be zero or nonzero at a given β, γ , depending on the gauge choice.

But if the VEV of ϕ is misleading, at least outside the context of perturbation theory, we may still ask whether the Higgs phase of an $SU(2)$ gauge-Higgs theory can be distinguished from a non-Higgs phase by the spontaneous breaking of the $SU(2)_{\text{global}}$ symmetry. This question is motivated by the fact that a nonzero (but gauge-dependent) $\langle \phi \rangle$ always implies a broken $SU(2)_{\text{global}}$. The idea is to turn this around; i.e., the signature for the Higgs phase is spontaneously broken $SU(2)_{\text{global}}$, regardless of whether $\langle \phi \rangle$ is zero or nonzero in some gauge. If that idea makes sense, then we must be able to find a gauge-invariant order parameter which is sensitive to the symmetry breaking but insensitive to any gauge choice. Such an order parameter must be inherently nonlocal, since we know from the work of Osterwalder and Seiler [3] that the VEV of local gauge-invariant observables in a gauge-Higgs theory is analytic in the coupling constants, along a path joining the confinementlike to the Higgs regime. We must also confront the Goldstone theorem: if a global continuous symmetry is spontaneously broken, how can massless excitations be

¹The term “custodial symmetry” is sometimes used to refer to the group $SU(2)_{\text{global}}$ of R -transformations, rather than the diagonal subgroup $SU(2)_D$. See, e.g., Maas [14].

²To this we might add that the lattice Abelian-Higgs model in four dimensions has a massless phase in some region of the $\beta - \gamma$ plane [4,15], despite the fact that in unitary gauge $\langle \phi \rangle \neq 0$ also in that region.

avoided? The answer is that the global symmetry can be broken in the Higgs sector, without actually breaking (and giving rise to Goldstone modes) in the full theory.

To explain this point, let us begin by noting that the partition function $Z(\beta, \gamma)$ of the gauge-Higgs theory can be regarded as the weighted sum of partition functions $Z_{\text{spin}}(\gamma, U)$ of a spin system in a background gauge field, i.e.,

$$Z(\beta, \gamma) = \int DU Z_{\text{spin}}(\gamma, U) e^{-S_w[U]}, \quad (7)$$

where

$$Z_{\text{spin}}(\gamma, U) = \int D\phi e^{-S_H[\phi, U]} = e^{-\mathcal{F}_H[\gamma, U]}. \quad (8)$$

The only symmetry of the spin system, since $U_\mu(x)$ is fixed, is the $SU(2)_{\text{global}}$ symmetry $\phi(x) \rightarrow \phi(x)R$, and this symmetry may or may not be spontaneously broken, depending on the gauge field configuration $U_\mu(x)$. Our observation is that the symmetry may be spontaneously broken in every $Z_{\text{spin}}(\gamma, U)$ for which U is a thermalized configuration, without breaking the symmetry, or introducing a Goldstone mode, in the full theory. By ‘‘thermalized’’ we mean a member of the set of configurations which dominate the functional integral (7), samples of which are generated numerically in lattice Monte Carlo simulations.

How can we tell whether the global symmetry is spontaneously broken in these spin systems? If we denote the VEV of the $\phi(x)$ field in the background gauge field as $\bar{\phi}(x; U)$, where

$$\bar{\phi}(x; U) \equiv \frac{1}{Z_{\text{spin}}(\gamma, U)} \int D\phi \phi(x) e^{-S_H[\phi, U]}, \quad (9)$$

then in general, in a lattice volume V ,

$$\frac{1}{V} \sum_x \bar{\phi}(x; U) = 0, \quad (10)$$

and this is for two reasons. First, if no gauge is fixed so that $U_\mu(x)$ varies wildly in space, then $\phi(x)$ also varies wildly with position, and the spatial average vanishes. Still, at any given point x it could be that $\bar{\phi}(x; U) \neq 0$. But this is impossible for the second reason: In a finite volume and in the absence of any explicit $SU(2)_{\text{global}}$ breaking term, there can be no spontaneous symmetry breaking, and, since $\bar{\phi}$ transforms under $SU(2)_{\text{global}}$ symmetry, it follows that $\bar{\phi}(x; U) = 0$ at every point.

But of course real (and, therefore, finite volume) magnets can be magnetized at low temperatures, and in that case a global symmetry has been spontaneously broken, despite formal theorems to the contrary. The signature of a broken

symmetry in a real magnet, in the absence of an explicit source of symmetry breaking such as an external magnetic field, is the existence of long-lived metastable states of different but nonzero magnetization, with lifetimes that increase to infinity as $V \rightarrow \infty$. We can adopt this same principle to study broken symmetry in the spin system defined by (8). The idea is that the Boltzmann probability factor $\propto \exp[-S_H[\phi, U]]$ can be generated by long time evolution in a fictitious ‘‘fifth-time’’ t_5 , where the field $\phi(x, t_5)$ evolves according to, e.g., the Langevin equation, or the molecular dynamics approach, or via lattice Monte Carlo simulations. In the case of Monte Carlo simulations, t_5 is discrete and corresponds to the number of update sweeps through the lattice. But in any of these methods, the expectation value of an operator O is defined by

$$\bar{O} = \lim_{T_5 \rightarrow \infty} \frac{1}{T_5} \int_0^{T_5} dt_5 O[\phi(x, t_5)]. \quad (11)$$

We then use the fifth-time formalism, instead of (9), to define

$$\bar{\phi}(x; U) = \lim_{T_5 \rightarrow \infty} \lim_{V \rightarrow \infty} \frac{1}{T_5} \int_0^{T_5} dt_5 \phi(x, t_5), \quad (12)$$

with the order of limits as shown. If $\bar{\phi}(x; U) = 0$ at every point, then the symmetry is unbroken, otherwise the $SU(2)_{\text{global}}$ symmetry is broken spontaneously. Even if the symmetry is broken, it is still true that the spatial average of $\bar{\phi}(x; U)$ will vanish in general, as in (10). Moreover, $\bar{\phi}(x; U)$ is gauge-covariant rather than gauge-invariant, transforming as

$$\bar{\phi}(x; g \circ U) = g(x) \bar{\phi}(x; U). \quad (13)$$

However, this quantity has a gauge-invariant modulus

$$|\bar{\phi}(x; U)| = \sqrt{\frac{1}{2} \text{Tr}[\bar{\phi}^\dagger(x; U) \bar{\phi}(x; U)]}, \quad (14)$$

and the spatial average of the modulus is positive if $\bar{\phi}(x; U)$ is nonzero in general. We, therefore, define, as our gauge-invariant order parameter, the spatial average

$$\Phi[U] = \lim_{T_5 \rightarrow \infty} \lim_{V \rightarrow \infty} \frac{1}{V} \sum_x \left| \frac{1}{T_5} \int_0^{T_5} dt_5 \phi(x, t_5) \right|, \quad (15)$$

with $\Phi[U] = 0$ or $\neq 0$ in the unbroken and spontaneously broken cases respectively. If at given couplings β, γ we find that $\Phi[U] \neq 0$ for gauge field configurations contributing to $Z(\beta, \gamma)$ in the thermodynamic limit, i.e., if

$$\langle \Phi \rangle \equiv \frac{1}{Z(\beta, \gamma)} \int DU \Phi[U] e^{-(S_w[U] + \mathcal{F}_H[U])} > 0, \quad (16)$$

then by this definition $SU(2)_{\text{global}}$ is spontaneously broken in each of the Z_{spin} subsystems, at that point in the $\beta - \gamma$ phase diagram.

We can now understand the absence of Goldstone modes. The order parameter for symmetry breaking in a $Z_{\text{spin}}(\gamma, U)$ system is the gauge covariant quantity $\bar{\phi}(x; U)$, which vanishes when averaged over gauge-field configurations, i.e.,

$$\langle \bar{\phi}(x; U) \rangle = 0. \quad (17)$$

The same can be said of long-range correlations in various n -point functions. Such long-range correlations only exist, in a theory at fixed U and $\Phi[U] > 0$, in the n -point functions of gauge-noninvariant operators. These correlators vanish in the full theory. To pick a trivial example, the correlator

$$\frac{1}{2} \overline{\text{Tr}[\phi^\dagger(x)\phi(y)]} = \frac{1}{Z(\gamma, U)} \int D\phi \frac{1}{2} \text{Tr}[\phi^\dagger(x)\phi(y)] e^{-S_H[\phi, U]} \quad (18)$$

may have long range correlations for a particular gauge field U with $\Phi[U] > 0$, but this quantity vanishes when integrating over all gauge fields,

$$\langle \overline{\text{Tr}[\phi^\dagger(x)\phi(y)]} \rangle = 0, \quad (19)$$

as does $\langle \text{Tr}[\phi^\dagger(x)\phi(y)] \rangle$. One could, of course, construct a gauge-invariant quantity such as

$$G(x, y) = \overline{\langle \text{Tr}[\phi^\dagger(x)U(x, y)\phi(y)] \rangle}, \quad (20)$$

where $U(x, y)$ is a Wilson line with endpoints x, y , but there is no particular reason why this quantity should have a power-law falloff. The point here is that long-range correlations in the individual $Z_{\text{spin}}(\gamma, U)$, which are due to the Goldstone theorem, must cancel out in the full theory.

But the absence of Goldstone modes does not mean that gauge-Higgs theory in the “broken” phase (meaning that all the non-negligible spin systems are in the broken phase), is qualitatively similar to gauge-Higgs theory in the unbroken phase. We will elaborate on how these phases can differ in Sec. V.

A. $SU(2)_{\text{global}}$ in unitary gauge

One might wonder what happens to the $SU(2)_{\text{global}}$ symmetry in unitary gauge, where there is no longer any freedom to transform ϕ . In fact, nothing happens; the symmetry is still there. Let us fix to $\phi = 1$. Then

$$Z = \int DU \exp \left[-S_W + \gamma \sum_{x, \mu} \frac{1}{2} \text{Tr} U_\mu(x) \right]. \quad (21)$$

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

$$\begin{aligned} Z &= \int DU \left\{ \Delta_{FP}[U] \int Dg \delta(F[g \circ U]) \right\} \\ &\quad \times \exp \left[-S_W + \gamma \sum_{x, \mu} \frac{1}{2} \text{Tr} U_\mu(x) \right] \\ &= \int DU \Delta_{FP}[U] \delta(F[U]) e^{-S_W} \\ &\quad \times \int Dg \exp \left[\gamma \sum_{x, \mu} \frac{1}{2} \text{Tr} [g^\dagger(x) U_\mu(x) g(x + \hat{\mu})] \right] \\ &= \int DU \Delta_{FP}[U] \delta(F[U]) Z_{\text{spin}}(\gamma, U) e^{-S_W}. \end{aligned} \quad (22)$$

The last line is Eq. (7) in the gauge $F[U] = 0$. Since the order parameter Φ for symmetry breaking in $Z_{\text{spin}}(\gamma, U)$ is gauge invariant, we recover the original formulation, with $\phi(x)$ replaced by $g(x)$.

B. Numerical procedure

We calculate $\langle \Phi \rangle$ by a Monte Carlo-within-a-Monte Carlo procedure. That is to say, the usual update sweeps involve sweeping site by site through the lattice, and updating the four link variables and the Higgs field at each site. Since both the link and scalar field variables are elements of the $SU(2)$ group, the updates of both types of variables can be carried out using the Creutz heat bath method. In this method, one seeks to stochastically generate $SU(2)$ elements G according to a probability distribution

$$dP(G) \propto e^{\frac{1}{2} \text{Tr}[GA]} dG, \quad (23)$$

where A is a fixed matrix proportional to an $SU(2)$ group element. For updating a link variable $G = U_\mu(x)$, we have

$$\begin{aligned} A &= \beta \sum_{\nu \neq \mu} \{ U_\nu(x + \hat{\mu}) U_\mu^\dagger(x + \hat{\nu}) U_\nu^\dagger(x) \\ &\quad + U_\nu^\dagger(x + \hat{\mu} - \hat{\nu}) U_\mu^\dagger(x - \hat{\nu}) U_\nu(x - \hat{\nu}) \} \\ &\quad + \gamma \phi(x + \hat{\mu}) \phi^\dagger(x), \end{aligned} \quad (24)$$

while for updating a scalar field variable $G = \phi(x)$ we use

$$A = \gamma \sum_{\mu} (\phi^\dagger(x - \hat{\mu}) U_\mu(x - \hat{\mu}) + \phi^\dagger(x + \hat{\mu}) U_\mu^\dagger(x)) \quad (25)$$

The heat bath procedure for generating group elements G in a probability distribution (23) is described in standard texts such as [16], and in the seminal paper by Creutz [17].

The data-taking sweep, however, is a simulation of the spin system (8), and entails n_{sw} sweeps through the lattice, updating only the Higgs field by the heat bath method, while keeping the gauge field fixed. In the course

of this data-taking sweep, on a finite lattice volume V , we measure

$$\Phi_{n_{sw},V}[U] = \frac{1}{V} \sum_x \left| \frac{1}{n_{sw}} \sum_{t_5=1}^{n_{sw}} \phi(x, t_5) \right|, \quad (26)$$

where $\phi(x, t_5)$ is the Higgs field at point x after t_5 update sweeps, holding the U field fixed. The quantity we would like to estimate is the limiting value,

$$\langle \Phi \rangle = \lim_{n_{sw} \rightarrow \infty} \lim_{V \rightarrow \infty} \langle \Phi_{n_{sw},V}[U] \rangle, \quad (27)$$

again with the order of limits as shown. In the infinite volume limit, we expect, on general statistical grounds, that

$$\langle \Phi_{n_{sw},\infty}[U] \rangle \approx \langle \Phi \rangle + \frac{\text{const}}{\sqrt{n_{sw}}}. \quad (28)$$

In the unbroken phase, with $\langle \Phi \rangle = 0$, this behavior would also hold at finite volume. In the broken phase, however, we expect $\langle \Phi_{n_{sw},V}[U] \rangle \approx \langle \Phi_{n_{sw},\infty}[U] \rangle$ to only hold for n_{sw} smaller than the lifetime $T_{\text{meta}}(V)$ of the metastable state, and then to go to zero as n_{sw} increases beyond $T_{\text{meta}}(V)$. So on a finite volume we must use (28) to extrapolate, from a set of values $\{\langle \Phi_{n_{sw},V}[U] \rangle\}$ computed at $n_{sw} < T_{\text{meta}}(V)$ to the $n_{sw} \rightarrow \infty$ limit, checking that $T_{\text{meta}}(V)$, where the linear extrapolation breaks down, increases with lattice volume V , and that the extrapolated estimate for $\langle \Phi \rangle$ converges as V increases.

To pin down the point of transition, it is also helpful to introduce a gauge-invariant quantity which functions as a susceptibility:

$$\chi = V \left\langle \left| \frac{1}{V} \sum_x \text{Tr}[\varphi(x; U) \{ \phi(x) - \bar{\phi}(x; U) \}] \right|^2 \right\rangle, \quad (29)$$

where we have defined a gauge covariant, unimodular field

$$\varphi(x; U) = \frac{\bar{\phi}(x; U)}{|\bar{\phi}(x; U)|}. \quad (30)$$

The transition point, at fixed β , is identified with the value of γ where χ is maximized.

In the unbroken phase, $\varphi(x, U)$ is 0/0, strictly speaking, and χ has to be defined again in a fifth-time formalism. Let $\phi(x, t_5)$ denote the Higgs field configuration obtained after t_5 update sweeps in the spin system simulation at fixed U and lattice volume V . Then define

$$\begin{aligned} \bar{\phi}_{n_{sw},V}(x; U) &= \frac{1}{n_{sw}} \sum_{t_5=1}^{n_{sw}} \phi(x, t_5) \\ \varphi_{n_{sw},V}(x; U) &= \frac{\bar{\phi}_{n_{sw},V}(x; U)}{|\bar{\phi}_{n_{sw},V}(x; U)|}, \end{aligned} \quad (31)$$

and construct, in terms of these quantities,

$$\begin{aligned} \chi &= V \left\langle \left| \frac{1}{V} \sum_x \frac{1}{n_s} \sum_{t_5=n_{sw}+1}^{n_{sw}+n_s} \text{Tr}[\varphi_{n_{sw},V}(x; U) \right. \right. \\ &\quad \left. \left. \times \{ \phi(x, t_5) - \bar{\phi}_{n_{sw},V}(x; U) \}] \right|^2 \right\rangle. \end{aligned} \quad (32)$$

Apart from the finite lattice volume V , this definition involves a choice of n_{sw} for defining a gauge covariant field $\varphi_{n_{sw},V}(x; U)$, and a choice of $n_s \ll n_{sw}$ for the estimate of susceptibility. The parameter n_{sw} is chosen to be large enough to avoid substantial statistical errors, but small enough so that, in the broken phase, we do not have $\bar{\phi}_{n_{sw},V}(x; U)$ much smaller than the limit in (12) just due to the formal absence of symmetry breaking in a finite volume. Likewise, the choice of n_s balances the requirement of small statistical errors (n_s large), with a condition that $\bar{\phi}_{n_{sw},V}(x; U)$ and $\bar{\phi}_{n_{sw}+n_s,V}(x; U)$ do not differ appreciably. In practice, we have used $n_{sw} = 900$ and $n_s = 100$ in computing χ .

We have found that χ defined in this way is very useful in practice for locating the transition point, but we do not have a rigorous argument for why this works so well. The proper definition of the transition point is that $\langle \Phi \rangle$ is zero, in the appropriate limits, below the transition point, and is non-zero above that point. We have found that this condition is satisfied by the transition point suggested by the peak in χ , in every case we have examined.

C. Landau gauge

We must check whether the gauge-invariant, symmetry-breaking criterion $\Phi[U] > 0$ is a Landau gauge criterion [2,18,19] in disguise.

When $\Phi[U] > 0$ in the appropriate limits, it means that the Higgs field fluctuates preferentially around one of a set of field configurations, related by $SU(2)_{\text{global}}$ transformations, in the infinite volume limit. It is natural to suppose that, in a given background gauge field, $\phi(x)$ fluctuates around the configuration which minimizes the Higgs action S_H , and which, therefore, maximizes

$$\sum_{x,\mu} \frac{1}{2} \text{Tr}[\phi^\dagger(x) U_\mu(x) \phi(x + \hat{\mu})]. \quad (33)$$

That is, after all, the starting assumption of any perturbative expansion.

Let $\phi_{\max}(x)$ be this maximizing configuration. Then $g(x) = \phi_{\max}(x)$ is a gauge transformation which takes $U_\mu(x)$ into Landau gauge. It follows that if $U_\mu(x)$ is already in Landau gauge, then $g(x) = \phi_{\max}(x)$ is the gauge transformation which preserves the Landau gauge condition, and this is the remnant symmetry of Landau gauge, namely the transformations $g(x) = g$ which are independent of position. So we might expect, in the broken phase of the spin system (8) with U in Landau gauge, that $\phi(x)$ fluctuates around one of the maximizing configurations $\phi(x) = g \in \text{SU}(2)$.

The order parameter proposed in [2,19] was devised to detect the breaking of remnant symmetry in Landau gauge. We define, in lattice volume V ,

$$\Omega_V[U] = \left| \frac{1}{V} \sum_x \phi_L(x) \right|^2, \quad (34)$$

where the subscript L in $\phi_L(x)$ indicates that $\phi(x)$ is computed Landau gauge. Note that in (34) the modulus is taken after the spatial average, whereas in the definition of $\Phi[U]$ the modulus is taken prior to the sum over position. We can pin down the transition point from the peak in susceptibility,

$$\chi_L = V(\langle \Omega_V^2 \rangle - \langle \Omega_V \rangle^2). \quad (35)$$

Since $\Phi[U]$ is gauge invariant, it can always be evaluated in Landau gauge, and if $\Phi[U] > 0$, it means that $\phi(x)$ fluctuates around some fixed configuration. One would imagine that this configuration would be a fixed group element, constant in spacetime, in which case $\Omega[U]$ is also nonzero, and there is then no real difference between the two criteria. The flaw in the argument is that there exist many Gribov copies in Landau gauge, and if U is fixed to one of them, then there exists a gauge transformation $g'(x)$ to some other copy, and, therefore, $\phi_{\max}(x) = g'(x)$ is also a local maximum of S_H . It may be that $\phi(x, t_5)$ fluctuates around a $\phi_{\max}(x)$ of this kind, whose spatial average vanishes. In that case, it is possible that both $\langle \Phi \rangle > 0$ and $\langle \Omega \rangle = 0$ hold simultaneously for some range of couplings, a possibility which we now show is confirmed by the data.

III. NUMERICAL RESULTS

There have been many numerical studies of the phase structure of the SU(2) gauge-Higgs model described by (2), and these have found a transition line in the $\beta - \gamma$ plane terminating at a finite $\beta \approx 2$. In early studies [20], this was considered to be a line of first order transition, but according to the most recent work [21] it is only a region of sharp crossover behavior, up to at least $\beta = 2.725$. It is possible that true first order transitions appear at $\beta > 2.725$. This is all in accordance with the Osterweiler-Seiler theorem [3]. No ‘‘Coulomb’’ region, corresponding to a

$1/R$ potential between static sources, has been found in the phase diagram, although in principle such a region is not ruled out *a priori* [4]. String breaking in the confinement-like region of SU(2) gauge-Higgs models has been reported in [22].

Our procedure is as follows: After thermalization (up to 4000 updating sweeps on a 20^4 lattice), we take data after every 100 updating sweeps. In each data-taking sweep, we begin by saving the lattice configuration, fixing to Landau gauge, and computing $\Omega(U)$ in (34), which is used to compute the Landau gauge susceptibility χ_L (35). The lattice is then restored to the saved configuration. This is followed by a Monte Carlo within a Monte Carlo; meaning that we hold the gauge link variables fixed, and update only the Higgs field from $t_5 = 1$ to $t_5 = n_{sw}$ sweeps. Denote the Higgs field at point x and the Higgs-only update sweep t_5 as $\phi(x, t_5)$. We compute $\Phi_{n_{sw}, V}$, $\bar{\phi}_{n_{sw}, V}(x; U)$, $\varphi_{n_{sw}, V}(x; U)$ according to Eqs. (26) and (31) respectively. Our simulations were carried out on volumes $8^4, 12^4, 16^4, 20^4$, at each of $n_{sw} = 100N^2$, $N = 2, 3, \dots, 12$. We then carried out the Higgs-only updates for a further $n_s = 100$ sweeps, to calculate χ in Eq. (32). Finally, the lattice is restored to the saved configuration. For the largest 20^4 lattice we collected 80 data sets.³

We begin with a display of the susceptibilities χ, χ_L vs γ at $\beta = 1.2$ in Fig. 1. It is known, from [20] and from [2], that there is no thermodynamic transition in γ , or even a sharp crossover, at this fixed value of β . At $\beta = 1.2$ there seems, however, to be a gauge-invariant, symmetry-breaking transition at $\gamma = 1.28$, and the Landau transition is at $\gamma = 1.4$. In Fig. 2, we plot the corresponding order parameter $\langle \Phi \rangle$ vs $1/\sqrt{n_{sw}}$, at various lattice volumes, below (Fig. 2(a)) and above (Fig. 2(b)) the transition point, at $\gamma = 1.2$ and $\gamma = 1.35$ respectively, and we see that the order parameter behaves as expected, falling to zero as $n_{sw} \rightarrow \infty$ below the transition. Above the transition the data indicates that $\langle \Phi \rangle \rightarrow 0$ in this limit *at fixed volume*, but it can also be seen that the onset of the drop towards zero increases with lattice volume, consistent with $\langle \Phi \rangle > 0$ in the appropriate pair of $V \rightarrow \infty, n_{sw} \rightarrow \infty$ limits. Likewise, the order parameter $\langle \Omega \rangle$ shown in Fig. 3 for the Landau transition just below ($\gamma = 1.35$) and just above ($\gamma = 1.5$) the transition behave as expected, falling to zero with $1/\sqrt{V}$ below the transition, and converging to a nonzero constant at large V above the transition. The point to notice here is that at $\beta = 1.2, \gamma = 1.35$ we have exactly the situation noted last section; i.e., there is a region in the phase diagram where $\langle \Phi \rangle > 0$ and $\langle \Omega \rangle = 0$. From this type of data, we conclude that the gauge-invariant criterion for SU(2)_{global} symmetry breaking is not the same as the Landau gauge criterion.

³Error bars were computed from a simple standard deviation of the mean; we did not check autocorrelations in this study.

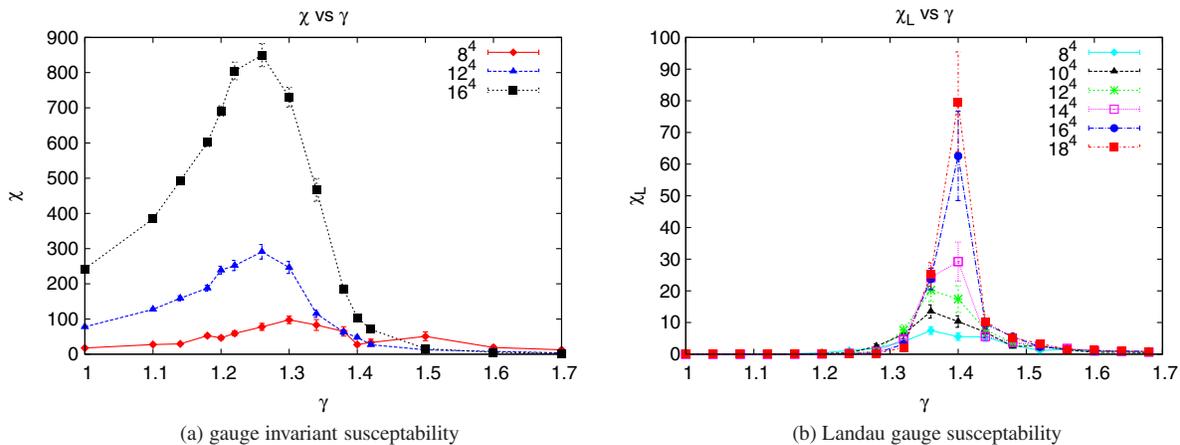


FIG. 1. (a) Gauge-invariant susceptibility χ vs γ , and (b) Landau gauge susceptibility χ_L vs γ , both at $\beta = 1.2$ and various lattice volumes. Note that the peaks in these two susceptibilities occur at different places, i.e., at $\gamma = 1.28$ for the gauge-invariant transition, and at $\gamma = 1.4$ for the Landau gauge transition.

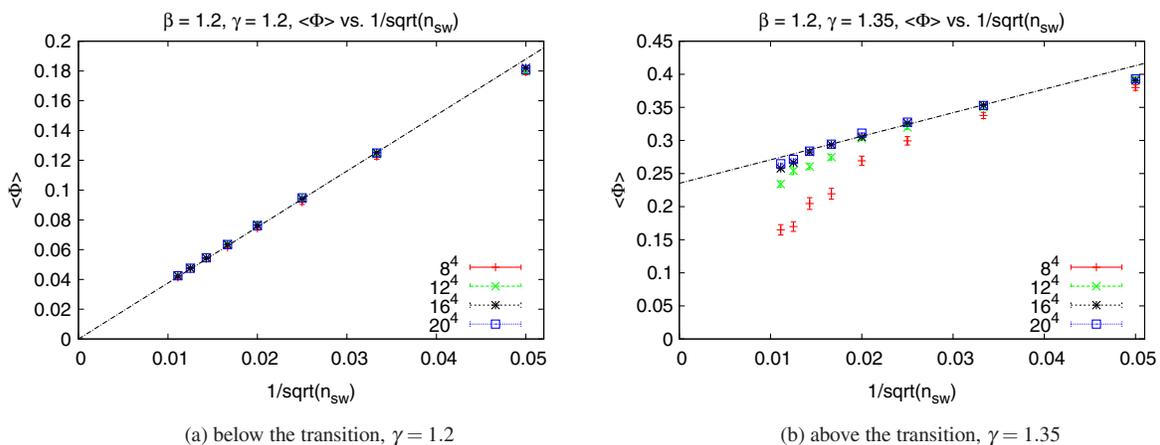


FIG. 2. Gauge invariant order parameter Φ vs $1/\sqrt{n_{sw}}$, where n_{sw} are the number of sweeps carried out on the matter field at fixed gauge field. The data is for $\beta = 1.2$ at lattice volumes 8^4 , 12^4 , 16^4 , 20^4 . (a) below the transition, at $\gamma = 1.2$; (b) above the transition, at $\gamma = 1.35$. Note the convergence, in subfigure (b), to a straight line with nonzero intercept on the y axis, as lattice volume increases.

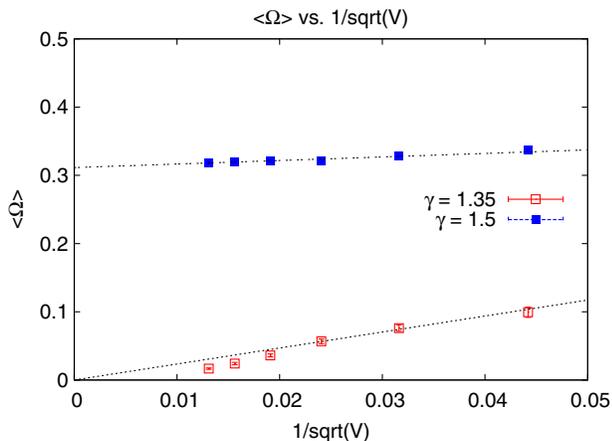


FIG. 3. Landau gauge order parameter Ω vs $1/\sqrt{V}$, where V is the lattice volume, at $\beta = 1.2$. Data is shown below the transition, at $\gamma = 1.35$, and above the transition, at $\gamma = 1.5$.

From the peaks in χ and χ_L we can locate the transition lines for the gauge-invariant, symmetry-breaking transition, and for the Landau gauge transition, and these are shown in Fig. 4. The Landau gauge transition line was previously found in [2], and our present result for that line agrees with the older calculation. The gauge-invariant, symmetry-breaking transition line is a new result.

At the larger $\beta \geq 2.0$ values, where there is a relatively sharp thermodynamic crossover in γ , we can find the crossover point from a peak in the plaquette susceptibility. Let

$$E_P = \frac{1}{6L^4} \sum_{\text{plaq}} \frac{1}{2} \text{Tr}[U_\mu(x)U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x)]$$

$$E_H = \frac{1}{4L^4} \sum_{x,\mu} \frac{1}{2} \text{Tr}[\phi^\dagger(x)U_\mu(x)\phi(x+\hat{\mu})], \quad (36)$$

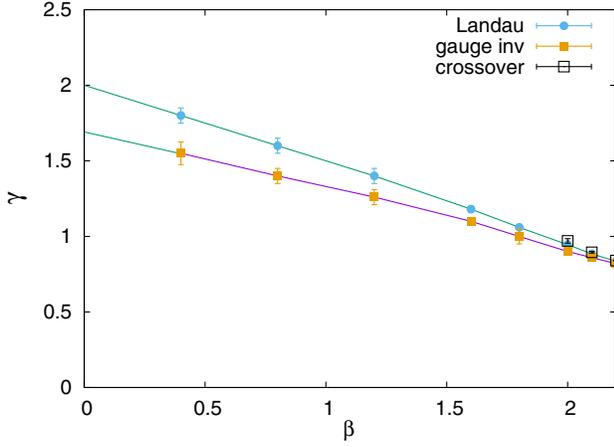


FIG. 4. Transition line (square points) for the gauge-invariant global SU(2) symmetry described in the text. The transition line for remnant gauge symmetry breaking in Landau gauge (circles) is shown for comparison, along with points at $\beta \geq 2.0$ (open squares) where we find a sharp thermodynamic crossover.

be the plaquette energy density and average Higgs energy density, respectively. Then the plaquette susceptibility is

$$\chi_P = \frac{\partial \langle E_P \rangle}{\partial \gamma} = 4L^4 (\langle E_P E_H \rangle - \langle E_P \rangle \langle E_H \rangle). \quad (37)$$

In this susceptibility, the location of the peaks, which lie on the thermodynamic crossover line (at $\beta \geq 2.0$) originally found in [20], are also displayed in Fig. 4.

IV. U(1) SYMMETRY BREAKING IN SU(3) GAUGE-HIGGS THEORY

The $SU(2)_{\text{global}}$ symmetry in SU(2) gauge-Higgs theory is in some sense accidental and there is, in the general case, no $SU(N)_{\text{global}}$ symmetry in an SU(N) gauge-Higgs theory. This is simply because the mapping of a Higgs multiplet to a group element, as in (1), does not generalize to SU(N) theories. There does exist, however, a global U(1) symmetry in SU($N > 2$) gauge-Higgs theories, with

$$S_H[U, \phi] = -\gamma \sum_{x, \mu} \text{Re}[\phi^\dagger(x) U_\mu(x) \phi(x + \hat{\mu})], \quad (38)$$

and where the unimodular Higgs field transforms in the fundamental representation of SU(N). This action is invariant, as pointed out by Maas *et al.* [7], under the U(1) transformations

$$\phi(x) \rightarrow e^{i\theta} \phi(x), \quad (39)$$

and our point is that this global symmetry, like any global symmetry, can be spontaneously broken.

The order parameter for the spontaneous symmetry breaking of the global symmetry (39) in the spin

system (8) is essentially identical to the $\langle \Phi \rangle$ order parameter defined in Sec. II, changing only the definition of the gauge-invariant modulus

$$|\bar{\phi}(x; U)| = \sqrt{\bar{\phi}^\dagger(x; U) \bar{\phi}(x; U)}, \quad (40)$$

where a dot product of color indices, rather than a trace, is implied. As before, $\langle \Phi \rangle = 0$ means that the global symmetry is unbroken, while $\langle \Phi \rangle > 0$ implies spontaneous breaking of the global U(1) symmetry.

We have computed the transition line in the $\beta - \gamma$ coupling plane for SU(3) gauge-Higgs theory, with the action consisting of the Wilson action

$$S_W = -\beta \sum_{\text{plaq}} \frac{1}{3} \text{ReTr}[U_\mu(x) U_\nu(x + \hat{\mu}) U_\mu^\dagger(x + \hat{\nu}) U_\nu^\dagger(x)] \quad (41)$$

plus S_H in (38), and a unimodular Higgs field. The numerical ‘‘Monte Carlo within a Monte Carlo’’ procedure is essentially the same as the one described in the previous section for the SU(2) case. The only difference is that updates of link and scalar field variables are carried out via the Metropolis algorithm. The first step is to generate an SU(3) matrix close to the identity element, and this is done by generating stochastically three SU(2) matrices, which are each embedded in a 3×3 matrix, as described in section IV.2.3 of Ref. [16]. Each SU(2) submatrix is generated by the heat bath method, with $A = \alpha \mathbb{1}$ in (23). The product of these three matrices is an SU(3) matrix G . When updating a link variable, we generate a trial link variable $U_{\text{try}} = G U_\mu(x)$, and then compute the change in the action $\Delta S = \Delta S_W + \Delta S_H$ generated by the trial link, which is then accepted or rejected according to the Metropolis algorithm. Similarly, when updating the scalar field, which is a unimodular 3-vector, we generate a trial variable $\phi_{\text{try}} = G \phi(x)$, compute the change in the Higgs action ΔS_H , and accept or reject by the Metropolis rule. We adjust the spread of the (stochastically generated) SU(2) matrices away from the identity matrix by adjusting the parameter α . This parameter is assigned different values for the link and scalar field updates, in order to obtain an acceptance rate in the Metropolis algorithm of about 50%.

The transition points are located by computing the susceptibility χ of (29) at fixed β over a range of γ and lattice volumes, and identifying the transition point as the location of the peak, as shown in Fig. 5 at $\beta = 3.0$. In this case, the transition is at $\beta \approx 1.85$. We also check that $\langle \Phi \rangle \rightarrow 0$ as $n_{\text{sw}} \rightarrow \infty$ below the transition, while $\langle \Phi \rangle$ extrapolates to a nonzero value above the transition. This is illustrated at $\beta = 3.0$ in Fig. 6. The transition line in the $\beta - \gamma$ coupling plane, for $0 < \beta < 5.6$ is shown in Fig. 7.

In the case of compact U(1) gauge-Higgs theory, with a single-charged scalar field, the additional symmetry is also

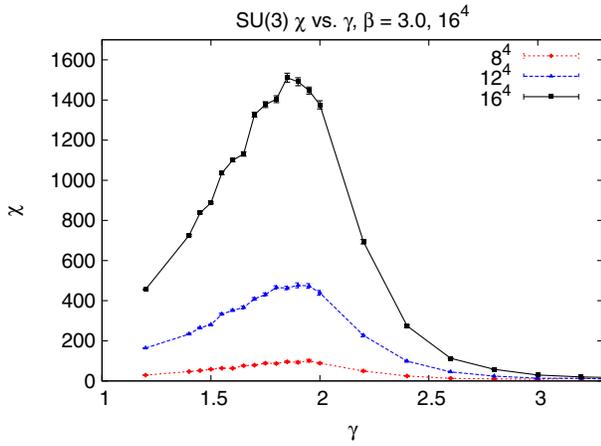


FIG. 5. Gauge-invariant susceptibility χ vs γ for SU(3) gauge-Higgs theory at $\beta = 3.0$, and lattice volumes 8^4 , 12^4 , 16^4 .

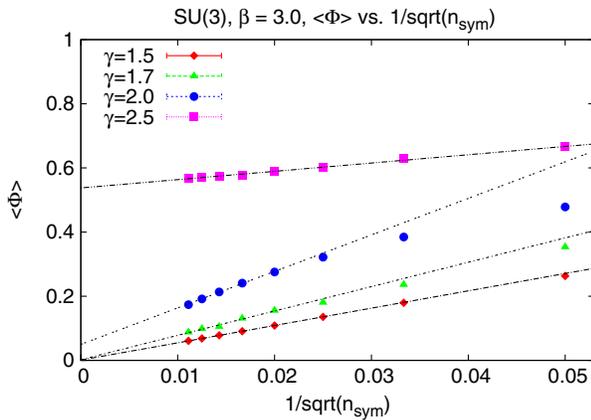


FIG. 6. Gauge invariant SU(3) order parameter Φ vs $1/\sqrt{n_{\text{sym}}}$ at $\beta = 3.0$ on a 16^4 lattice volume. Below the transition at $\gamma = 1.85$, the data extrapolates to zero as $n_{\text{sw}} \rightarrow \infty$. Above the transition, the data extrapolates to nonzero values.

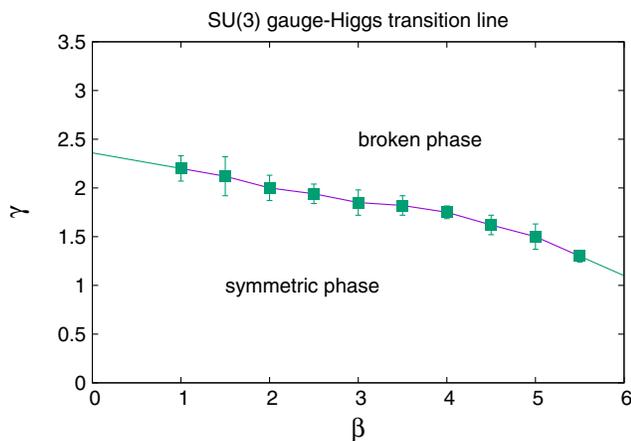


FIG. 7. Gauge-invariant transition line for global U(1) symmetry breaking in SU(3) gauge-Higgs theory.

global U(1), and it is that symmetry which can be spontaneously broken. This point seems to have been made previously, in connection with superconductivity, by Greiter [23]. It would be interesting to explore the phase diagram of compact scalar QED with a singly charged matter field, which contains a massless as well as confinementlike and Higgs regions, in connection with the broken vs unbroken realization of the global U(1) symmetry. We leave this for a future investigation.

V. SEPARATION-OF-CHARGE CONFINEMENT

Given that there exists a gauge-invariant global symmetry in gauge-Higgs theory which is realized in either a broken or unbroken phase, the obvious question is what physical property distinguishes these two phases. Our conjecture is that the symmetry-breaking transition corresponds to a transition between two different types of confinement, which in a previous article [8] we have referred to as C- and S_c -confinement.

In any SU(N) gauge theory with matter fields in the fundamental representation, such as QCD or gauge-Higgs theories, Wilson loops have a perimeter-law falloff asymptotically, and Polyakov lines have a nonzero vacuum expectation value. So what does it mean to say that such theories (QCD in particular) are confining? Historically, confinement was taken to mean the absence of free quarks in the QCD spectrum, and more generally confinement is defined as the color neutrality of the asymptotic particle spectrum. We will refer to this property as “C-confinement.” As we have already noted in the Introduction, gauge-Higgs theories in the Higgs regime, where there are no linearly rising Regge trajectories, no string formation and no string breaking, and only Yukawa forces, are also confining by this definition.

On the other hand, in a pure SU(N) gauge theory, there is a different and stronger meaning that can be assigned to the word “confinement”, which goes beyond C-confinement. Of course the spectrum of pure gauge theories consists of only color neutral objects, i.e., glueballs. But such theories also have the property that the static quark potential rises linearly or, equivalently, that large planar Wilson loops have an area-law falloff. It is reasonable to ask if there is any way to generalize this property to gauge theories with matter in the fundamental representation, and such a generalization was proposed in ref. [8]. It begins by noting that the Wilson area-law criterion in a pure gauge theory is equivalent to the following statement, which we call “separation of charge” confinement or “ S_c ” confinement. We consider a class of functionals $V(\mathbf{x}, \mathbf{y}; A)$ of the gauge field which transform bi-covariantly under a gauge transformation $g(x)$, i.e.,

$$V^{ab}(\mathbf{x}, \mathbf{y}; A) \rightarrow g^{ac}(\mathbf{x}, t) V^{cd}(\mathbf{x}, \mathbf{y}; A) g^{\dagger db}(\mathbf{y}, t), \quad (42)$$

and then define

S_c-confinement

Let $E_V(R)$, with $R = |\mathbf{x} - \mathbf{y}|$ be the energy (above the vacuum energy) of a state

$$\Psi_V \equiv \bar{q}^{a+}(\mathbf{x})V^{ab}(\mathbf{x}, \mathbf{y}; A)q^{b+}(\mathbf{y})\Psi_0, \quad (43)$$

where q^+, \bar{q}^+ represent creation operators for static quark/antiquark color charges, and Ψ_0 is the vacuum state. *S_c-confinement* means that there exists an asymptotically linear function $E_0(R)$, i.e.,

$$\lim_{R \rightarrow \infty} \frac{dE_0}{dR} = \sigma > 0, \quad (44)$$

such that

$$E_V(R) \geq E_0(R) \quad (45)$$

for any choice of bi-covariant $V(\mathbf{x}, \mathbf{y}; A)$.

In a pure gauge theory, *S_c-confinement* is equivalent to the Wilson area law criterion, with $E_0(R)$ the static quark potential and σ the asymptotic string tension.

Our proposal in [8] is that *S_c-confinement* should also be regarded as the confinement criterion in gauge + matter theories. The crucial element is that the bi-covariant operators $V^{ab}(\mathbf{x}, \mathbf{y}; A)$ must depend only on the gauge field A at a fixed time, and not on the matter fields. Excluding matter fields from $V^{ab}(\mathbf{x}, \mathbf{y}; A)$ means that we are dealing with a subclass of physical states Ψ_V which really correspond to two separated color charges, rather than two separated color neutral objects. The question that is addressed by this exclusion is whether (i) a nonconfining static quark potential is due exclusively to string-breaking effects by matter fields, or whether instead (ii) a nonconfining state can be constructed without any appeal to string breaking. Case (i) is *S_c-confinement*, case (ii) is C-confinement. The distinction is that in *S_c-confinement*, gauge-invariant physical states containing isolated color charges are associated with an energy proportional to the separation, and this cost in energy can only be eliminated by a string breaking process which essentially neutralizes the formerly isolated color charges by binding them to other particles. In C-confinement, the energy of states with separated charges needs not rise linearly, even without the intervention of a string-breaking process.

In an *S_c* confining theory, states Ψ_V are inevitably metastable for large charge separation, evolving (in Euclidean time) into two color neutral objects by string-breaking. But the point is that a string-broken state is not a state of separated color charge; color-electric gauge fields do not emanate from color neutral objects. The idea underlying *S_c* confinement is to focus on the subclass of states, metastable or not, which *do* correspond to separated

color charges, and these must be sources, because of the Gauss law, of some extended gauge field.

In [8], we showed that *S_c-confinement* does not exist everywhere in the $\beta - \gamma$ plane of SU(2) gauge-Higgs theory, by constructing V operators which do not satisfy the *S_c-confinement* criterion for sufficiently large γ . But this leaves open the question of whether the *S_c* condition is satisfied *anywhere* in the gauge-Higgs phase diagram, apart from the pure-gauge theory at $\gamma = 0$. In the next section, we will show that *S_c-confinement* exists in some $\gamma > 0$ region of the phase diagram, and this in turn implies the existence of a transition line between the C- and *S_c-confinement* phases, which we may speculate is identical to the symmetry-breaking transition discussed in the previous sections. The V operators introduced in [8] have found C-confinement only in some region above the gauge-invariant transition line shown in Fig. 4. Our conjecture is that there is no V operator which will find C-confinement below that symmetry-breaking transition line.

VI. S_c-CONFINEMENT AT STRONG COUPLINGS

We will show in this section, using strong-coupling expansions and a theorem from linear algebra, that *S_c-confinement* exists in the SU(2) gauge-Higgs system of Eq. (2) if the following conditions are satisfied:

$$\tilde{\gamma} \ll \tilde{\beta} \ll 1, \quad \gamma \ll \frac{1}{10}, \quad (46)$$

where we have defined

$$\tilde{\beta} \equiv \frac{\beta}{4}, \quad \tilde{\gamma} \equiv \frac{\gamma}{4}. \quad (47)$$

It should be stressed that this is an “if” but not an “only if” statement; it may be that *S_c-confinement* exists even if these conditions are not satisfied.

In order to introduce static quark-antiquark sources at points \mathbf{x}, \mathbf{y} , we include the hopping terms

$$\mu \sum_t \{ \bar{q}(\mathbf{x}, t+1)U_0(\mathbf{x}, t)q(\mathbf{x}, t) + \bar{q}(\mathbf{y}, t+1)U_0(\mathbf{y}, t)q(\mathbf{y}, t) + \text{H.c.} \} \quad (48)$$

in the gauge-Higgs action. The central idea is to show that Higgs part of the action is negligible in the expression

$$W_V(T) = \langle \Psi_V | e^{-HT} | \Psi_V \rangle, \quad (49)$$

providing the conditions (46) hold, and T is small enough. This implies that the energy expectation value, which is the logarithmic time derivative of $W_V(T)$, will conform to the *S_c* confinement criterion. As a trivial example, which nonetheless illustrates the general idea, let \mathbf{x}, \mathbf{y} be points separated by a distance L along the x axis, and let the

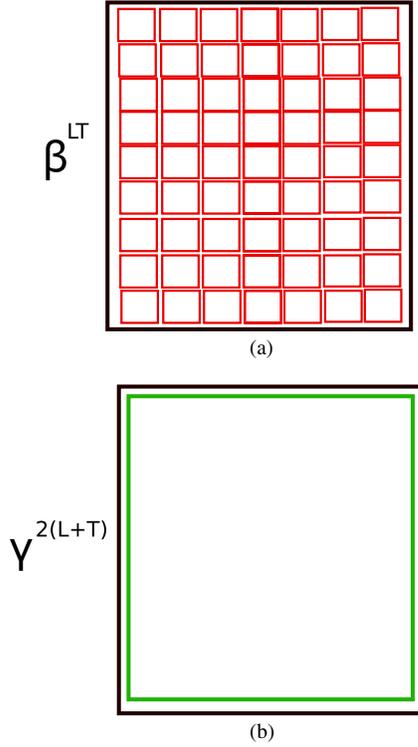


FIG. 8. Diagrammatic contributions to a rectangular $L \times T$ Wilson loop. (a) leading order in β ; (b) leading order in γ .

operator $V(\mathbf{x}, \mathbf{y}, A)$ be the Wilson line running along the x axis between these two points. Then $W_V(T)$ is proportional to the expectation value $W(L, T)$ of a rectangular Wilson loop of sides of length L and T . The strong-coupling diagrams to leading order in β alone, and in γ alone, are shown in Fig. 8, and their contribution to $W(L, T)$ is

$$2\tilde{\beta}^{LT} + 2\tilde{\gamma}^{2(L+T)}. \quad (50)$$

It is easy to see that for $L \gg T$ the γ contribution is negligible compared to the β contribution providing

$$T \ll 2 \frac{\log \tilde{\gamma}}{\log \tilde{\beta}}, \quad (51)$$

and in this limit the lattice version of the logarithmic time derivative reveals a linearly rising energy expectation value

$$E \approx -\log \left[\frac{W(L, T)}{W(L, T-1)} \right] = (-\log \tilde{\beta})L. \quad (52)$$

Conversely, at times $T \gg 2 \log \tilde{\gamma} / \log \tilde{\beta}$, it is the β contribution that is negligible, and the energy

$$E \approx -2 \log(\tilde{\gamma}) \quad (53)$$

is independent of separation L . In other words, around time $T = 2 \log \tilde{\gamma} / \log \tilde{\beta}$ the string breaks, and the static charges are screened by scalar particles.

String-breaking is generic in gauge-Higgs theories, but the point which is illustrated in this simple example is that for $\gamma \ll \beta$ this process takes time, which means that the energy E_V of the state Ψ_V , which corresponds to the logarithmic time derivative (52) at $T = 1$, obeys the S_c confinement criterion.⁴ This fact is unsurprising for $V = A$ Wilson line. The question is whether that same result is obtained for *arbitrary* choices of $V(\mathbf{x}, \mathbf{y}, A)$ which, we recall, can depend only on gauge link variables on a time slice, and not on the matter field. We, therefore, consider the more general expression

$$W_V(L, T) = \langle \Psi_0 | (\bar{q}^c(\mathbf{x}) V^{cd}(\mathbf{x}, \mathbf{y}; A) q^d(\mathbf{y}))_{t=T}^\dagger \times e^{-HT} (\bar{q}^a(\mathbf{x}) V^{ab}(\mathbf{x}, \mathbf{y}; A) q^b(\mathbf{y}))_{t=0} | \Psi_0 \rangle. \quad (54)$$

After integrating out the static quark fields, we have

$$W_V(L, T) = \mu^{2T} \frac{1}{Z} \int DUD\phi \text{Tr}[V(\mathbf{x}, \mathbf{y}; A)_{t=0} P(\mathbf{y}, T) \times V^\dagger(\mathbf{x}, \mathbf{y}; A)_{t=T} P^\dagger(\mathbf{x}, T)] e^{-S}, \quad (55)$$

where

$$P(\mathbf{x}, T) = \prod_{t=0}^{T-1} U_0(\mathbf{x}, t) \quad (56)$$

is a timelike Wilson line. This expression for $W_V(L, T)$ can be written as

$$W_V(L, T) = \mu^{2T} \int DU_1 DU_2 V^\dagger{}^{cd}(U_2) \times \mathcal{M}_T^{cd, ab}(U_2, U_1) V^{ab}(U_1) = \mu^{2T} (V | \mathcal{M}_T | V), \quad (57)$$

where $V(U) = V(\mathbf{x}, \mathbf{y}, A)$, and

$$\mathcal{M}_T^{cd, ab}(U_2, U_1) = \frac{1}{Z} \int DUD\phi \left\{ \prod_z \prod_{k=1}^3 \delta[U_k(\mathbf{z}, 0) - U_{k,1}(\mathbf{z})] \times \delta[U_k(\mathbf{z}, T) - U_{k,2}(\mathbf{z})] \right\} P^{bc}(\mathbf{y}, T) P^{\dagger da}(\mathbf{x}, T) e^{-S}. \quad (58)$$

It is useful to introduce a basis for $V(U)$. First define upper-lower index notation for a link variable in representation j

⁴In fact, we see from (50) that $E_V \propto L$ at large L even at $\beta = 0$. In that case, we have string breaking for any $T > 1$.

$$[U_\mu^{(j)}(x)]_{ab} = [U_\mu^{(j)}(x)]_a^b, \quad (59)$$

$$[U_\mu^{(j)}(x)]_{ab}^\dagger = [U_\mu^{(j)}(x)]^b_a. \quad (60)$$

We define a cluster \mathcal{C} , in a three-dimensional time slice, as

- (1) a set of spacelike links Λ , connected in the sense that there is a path on the lattice between any two links in Λ which is contained entirely in Λ ;
- (2) an $SU(2)$ representation $j(l) = \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$ at each link $l \in \Lambda$;
- (3) a set of vertices \mathcal{V} , and a ‘‘color connection array’’ B at each vertex. A vertex is a site shared by two or more (up to six) links in Λ . The upper and lower indices of the corresponding connection array $B_{a_1 \dots a_n}^{b_1 \dots b_m}(x)$ are contracted with the lower and upper indices of link variables (in various representations) which transform at that site, such that the product is a gauge singlet at the vertex x .

In an $SU(N > 2)$ gauge theory, it would also be necessary to specify an orientation, i.e., a choice of U_μ or U_μ^\dagger at each link. This choice is not strictly necessary for the $SU(2)$ group because of the pseudoreality property of $SU(2)$ group representations, and inclusion of all orientations would constitute an over-complete basis. However, any given cluster may be represented more compactly, meaning with a simpler set of connection arrays, by using a particular choice of orientations. The simplest example of a connection matrix at site x is one which connects the color indices of a single $j = \frac{1}{2}$ ‘‘ingoing’’ link with one $j = \frac{1}{2}$ ‘‘outgoing’’ link

$$[U_\mu(x - \hat{\mu})]_a^b B_b^c(x) [U_\nu(x)]_c^d(x), \quad B_b^c(x) = \frac{1}{\sqrt{2}} \delta_b^c. \quad (61)$$

If both links were represented as ingoing, we would have

$$[U_\mu(x - \hat{\mu})]_a^b B_{bc}(x) [U_\nu(x)]_c^d(x), \quad B_{bc}(x) = \frac{1}{\sqrt{2}} \varepsilon_{bc}. \quad (62)$$

These combinations could in principle occur in equivalent representations of the same cluster.

In general, under a transformation $g \in SU(2)$, we have that, for states in representation j ,

$$\begin{aligned} \varphi(j)_a &\rightarrow \varphi'(j)_a = G(j)_a^b \varphi(j)_b \\ \psi(j)^a &\rightarrow \psi'(j)^a = G(j)^a_b \psi(j)^b. \end{aligned} \quad (63)$$

where $G(j)$ is the gauge transformation corresponding to g in representation j , and suppose that we have a set of φ, ψ transforming in this way. Then the connection array has the property that

$$\begin{aligned} \varphi_{1a_1}(j_1) \varphi_{2a_2}(j_2) \dots \varphi_{na_n}(j_n) \psi_1^{b_1}(j'_1) \psi_2^{b_2} \dots \psi_m^{b_m}(j'_m) \\ \times B_{b_1 b_2 \dots b_m}^{a_1 a_2 \dots a_n}(\{j, j'\}) \end{aligned} \quad (64)$$

transforms like a singlet. In general, there may be more than one singlet in the decomposition of a product of representations, so we distinguish among them (suppressing the dependence on the representations $\{j, j'\}$) by an additional index κ in $B_{b_1 b_2 \dots b_m}^{a_1 a_2 \dots a_n}(x, \kappa)$, with B normalized such that

$$B_{b_1 b_2 \dots b_m}^{a_1 a_2 \dots a_n}(x, \kappa) B_{a_1 a_2 \dots a_n}^{b_1 b_2 \dots b_m}(x, \kappa') = \delta_{\kappa \kappa'}. \quad (65)$$

Then the gauge-invariant functional $U(\mathcal{C})$ defined on the cluster \mathcal{C} is

$$U(\mathcal{C}) = \mathcal{N} \prod_{l=(x,k) \in L} \sqrt{2j_l + 1} U_k^{(j_l)}(x) \prod_{x \in \mathcal{V}} B_{b_1 \dots b_m}^{a_1 \dots a_n}(x, \kappa_x), \quad (66)$$

where \mathcal{N} is some overall normalization constant, and it is understood that the lower(upper) indices of B contract with the upper(lower) indices of links entering (leaving) a site. With these definitions, taking account of (65) and

$$\int dU [U_\mu^{(j)}(x)]_a^b [U_\mu^{(j)}(x)]_c^d = \frac{\delta_{jj'}}{2j+1} \delta_a^c \delta_b^d, \quad (67)$$

we have

$$(U(\mathcal{C}')|U(\mathcal{C})) = \delta_{\mathcal{C}\mathcal{C}'}. \quad (68)$$

The simplest cluster is a Wilson loop $W_j(\mathcal{C})$ in representation j .

A bicovariant function $U^{ab}(\mathcal{C}_{xy})$ is a function of links on a cluster \mathcal{C}_{xy} which transforms like $V^{ab}(x, y; A)$. This means that in a cluster \mathcal{C}_{xy} there is either one single link entering or leaving sites x, y , or else a connection array at x, y forms a fundamental representation out of links attached to that site, rather than a singlet. The simplest example is a Wilson line running between sites x and y .

We now define the matrix in a cluster basis,

$$M_T(\mathcal{C}_{2xy}, \mathcal{C}_{1xy}) = \quad (69)$$

$$\int DU_2 DU_1 U^{\dagger cd}(\mathcal{C}_{2xy}) \mathcal{M}_T^{cd, ab}(U_2, U_1) U^{ab}(\mathcal{C}_{1xy}), \quad (70)$$

and the strategy is to estimate the largest eigenvalues of this matrix. The logarithmic lattice time derivative of the largest eigenvalue $\lambda_{\max}(T)$ is simply $-\log[\lambda_{\max}(T)/\lambda_{\max}(T-1)]$, and the case of $T = 1$ is of particular interest, since in this case $\lambda_{\max}(1)$ is the largest eigenvalue of the transfer matrix restricted to the subspace of Ψ_V states. Then

$$E_V \geq E_{\min} = -\log \lambda_{\max}(1), \quad (71)$$

and we have S_c confinement if E_{\min} is bounded from below by a function which rises linearly with $L_0 = |x - y|$. In an S_c confining theory, a Ψ_V state is metastable and evolves in Euclidean time to a lower energy state consisting of two color neutral objects via the usual string breaking process. We stress again that stability is in no way a condition for S_c confinement. String breaking takes place when T is large enough such that $-\log[\lambda_{\max}(T)/\lambda_{\max}(T-1)]$ ceases to rise linearly with separation L_0 .

A. Simple contours and the Gershgorin theorem

We begin with the simplest possible bicovariant clusters, i.e., Wilson lines in the fundamental ($j = \frac{1}{2}$) representation, running along contours C_1, C_2 between points x and y . Within the strong coupling expansion, with $\tilde{\gamma} \ll \tilde{\beta} \ll 1$ and small T , it can be seen that the *leading* contributions to each matrix element $M(C_2, C_1)$ are either “confining”, in the sense that their logarithmic time derivatives are greater than or equal to $-\log(\beta)L'$, where $L' \geq |x - y|$, or else they are “screening”, in the sense that the logarithmic time derivative is a constant independent of $L_0 = |x - y|$. The diagonal elements $M(C_2, C_1)$ with $C_2 = C_1$, and also “near-diagonal” elements, are of the confining type, while most off-diagonal elements are of the screening type. Since the off-diagonal screening matrix elements are far more numerous than confining matrix elements, the question is whether these screening elements can overwhelm the confining elements, and result in a ground state with nonconfining behavior.

The gauge-Higgs action consists of an “ E^2 -term,” which is the sum over timelike plaquettes in the Wilson action, a “ B^2 -term”,⁵ which is the sum over spacelike plaquettes in the Wilson action, and a Higgs-term, which is the term proportional to γ . It simplifies the discussion to initially ignore the B^2 term. Then the leading contributions, under the condition that $T \ll \log \gamma / \log \beta$, can be grouped into terms that depend only on β , which are confining, that depend only on γ , which are screening, and mixed terms that depend on both β and γ , i.e.,

$$M_T(C_2, C_1) = M_\beta(C_1, C_1)\delta_{C_2, C_1} + M_\gamma(C_2, C_1) + M_{\text{mix}}(C_2, C_1), \quad (72)$$

where the leading contributions in M_β, M_γ go as

$$M_\beta(C, C) = \tilde{\beta}^{L(C)T} \quad (73)$$

$$M_\gamma(C_2, C_1) = \tilde{\gamma}^{L(C_1)+L(C_2)+2T}, \quad (74)$$

If loops C_1, C_2 are large in the sense that

$$\beta^{L(C_1)T} \quad \text{and} \quad \beta^{L(C_2)T} \ll \beta^{L_0 T} \gamma^{2L_0} \quad (75)$$

then the leading contribution to M_{mix} is

$$M_{\text{mix}}(C_2, C_1) = \tilde{\beta}^{L_0 T} \tilde{\gamma}^{L(C_1)+L(C_2)+2L_0+2T}, \quad (76)$$

otherwise

$$M_{\text{mix}}(C_2, C_1) = \tilde{\beta}^{L'T} \tilde{\gamma}^{P(C_1, C_2)}, \quad (77)$$

where L_0 in (76) is the minimal distance between points x and y , L' in (77) is the smaller of $L(C_1)$ and $L(C_2)$, and $P(C_1, C_2)$ is the perimeter of the area enclosed by curves C_1, C_2 (which is $\leq L(C_1) + L(C_2)$). Examples of the diagrammatic representation of these matrix elements are shown in Fig. 9(a) for (73), Fig. 9(b) for (74), Fig. 11(b) for (76) and Fig. 10(a) for (77). The M_{mix} confining matrix element is negligible compared to the corresponding M_γ matrix element for $L_0 \gg T$, $P(C_1, C_2) \approx L(C_1) + L(C_2)$, and when

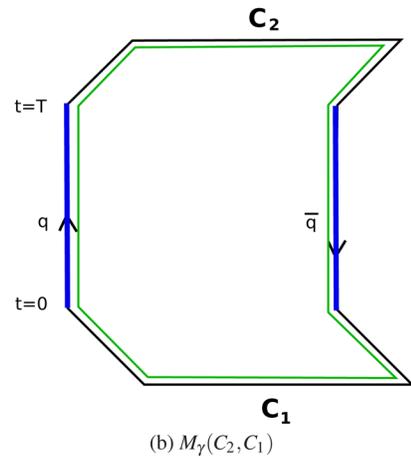
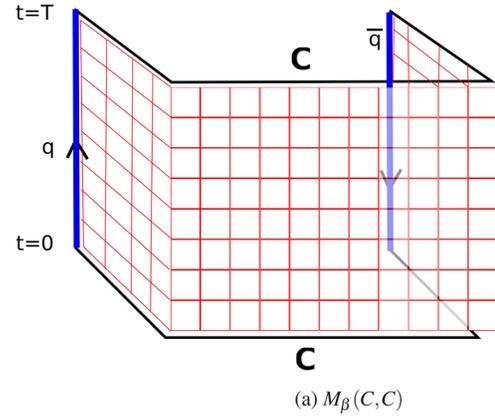


FIG. 9. (a) Diagonal contributions to $M_\beta(C_2, C_1)$. (b) Screening contributions to off-diagonal terms in $M_\gamma(C_2, C_1)$.

⁵Not to be confused, of course, with the B connection arrays.

$$T \ll \frac{\log \tilde{\gamma}}{\log \tilde{\beta}}. \quad (78)$$

Now consider the matrix

$$\tilde{M}(C_2, C_1) = M_\beta(C_1, C_1)\delta_{C_2, C_1} + M_\gamma(C_2, C_1). \quad (79)$$

This is a matrix of enormous dimensionality, and only the diagonal terms are confining. The rest, i.e., the vast majority, are screening. So the question is whether the eigenvalues of this matrix are of the confining or screening type. This question can be answered with the help of the following theorem (see e.g., [24]):

The Gershgorin Circle Theorem

Let A be a complex $n \times n$ matrix, with matrix elements A_{ij} , and let

$$r_i = \sum_{j \neq i} |A_{ij}| \quad (80)$$

be the sum of the magnitudes of the off-diagonal entries in the i -th row. Let D_i be a closed disk of radius r_i , centered at A_{ii} , in the complex plane. These, for $i = 1, 2, \dots, n$, are known as ‘‘Gershgorin disks’’. The theorem states that every eigenvalue of A must lie within at least one of the Gershgorin disks. In particular, let an eigenvalue λ correspond to an eigenvector \mathbf{u} with $u_i = 1$ and $|u_j| < 1$ for all $j \neq i$. Then

$$|\lambda - A_{ii}| \leq r_i. \quad (81)$$

With this motivation we compute an upper bound for r_C :

$$\begin{aligned} r_C &= \sum_{C_1 \neq C} \tilde{M}(C, C_1) \\ &\approx \tilde{\gamma}^{L(C)+2T} \sum_{C_1 \neq C} \tilde{\gamma}^{L(C_1)} < \tilde{\gamma}^{L(C)+2T} \sum_{L=L_0}^{\infty} \tilde{\gamma}^L N(L), \end{aligned} \quad (82)$$

where $N(L)$ is the number of open contours with endpoints \mathbf{x}, \mathbf{y} of length L . Without the second endpoint restriction, $N(L) = 5^L$ since at each step there are five possible directions to go without backtracking. This will serve as an upper limit

$$r_C < \tilde{\gamma}^{L(C)+2T} \sum_{L=L_0}^{\infty} (5\tilde{\gamma})^L < \tilde{\gamma}^{L(C)+2T} \frac{(5\tilde{\gamma})^{L_0}}{1 - 5\tilde{\gamma}}, \quad (83)$$

and this radius should be compared with the diagonal term $M_\beta(C, C)$ in (73). Assuming the conditions (46) are satisfied, and that T is small enough so that (78) also holds, then the radius r_C of the Gershgorin disk D_C is negligible compared to the diagonal term $\tilde{M}(C, C)$. Since also $M_{\text{mix}}(C_2, C_1) \ll M_\gamma(C_2, C_1)$, it follows that the

difference between the largest eigenvalues of M_T and the corresponding eigenvalues of M_β are, by the Gershgorin theorem, negligible. This in turn implies S_c confinement.

Adding back the B^2 term in the action does not change this conclusion. The effect of the B^2 term is to introduce subleading dressings of the sheet of plaquettes, and to allow for new contributions to off-diagonal elements of the confining type. These types of contributions are both illustrated in Fig. 10(b). There are certainly contours C_1, C_2 such that $M_{\text{mix}}(C_2, C_1)$ is larger than $M_\beta(C_2, C_1)$, as would be the case for the contributions shown in Fig. 11 if the combined area $A(C_1, C_2)$ bounded by C_1 and C_2 in a plane is such that

$$\tilde{\beta}^{A(C_1, C_2)} < \tilde{\gamma}^{L(C_1)+L(C_2)}. \quad (84)$$

But inspection of such terms (compare, e.g., Fig. 11(b) with Fig. 9(b)) shows that, under the assumption (78), they are always very much smaller than the corresponding matrix elements of $M_\gamma(C_2, C_1)$ which, we have already argued from the Gershgorin theorem, can be neglected. So apart from negligible terms, and assuming the conditions (46)

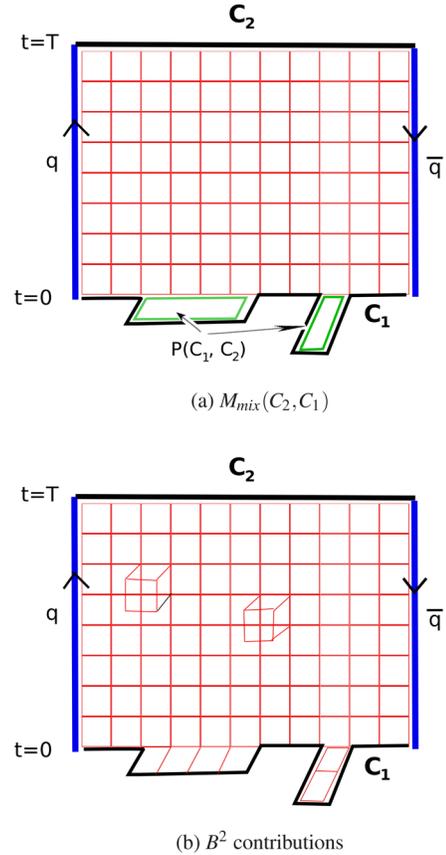
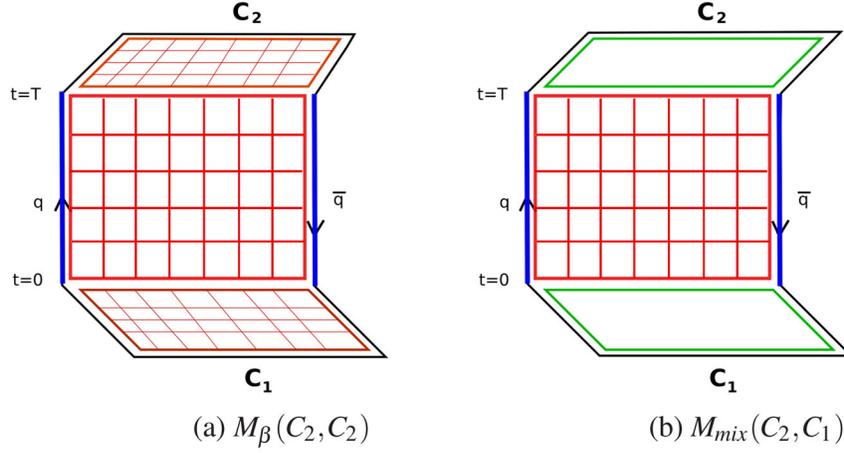


FIG. 10. (a) Off-diagonal ‘‘confinement-type’’ contributions to M_{mix} . (b) off-diagonal contributions to M_β . These necessarily include B^2 terms.

FIG. 11. A comparison of off-diagonal confinement-type contributions in (a) M_β and (b) M_{mix} .

and (78), all of the leading terms in $M_T(C_2, C_1)$ are in $M_\beta(C_1, C_2)$; i.e., the terms of a pure gauge theory. Then the eigenvalue spectrum of M_T , for times T up to the limit in (78), will not differ much from that of a pure gauge theory.

It follows that up to this limit in T , the spectrum of static $q\bar{q}$ states, obtained by evolving states of the form Ψ_V shown in (43) for Euclidean time T , will have energies bounded by a linear potential, with a string tension which is closely approximated by that of a pure gauge theory. Beyond that time, screening will take over. This is S_c -confinement, but the argument assumes that we restrict the basis states to simple open contours.

B. Summation over clusters

So the next step is to enlarge the basis and, motivated by the Gershgorin theorem, we consider the sum over general clusters in the screening terms

$$r_C = \sum_{C_1 \neq C} |M_\gamma(C, C_1)|. \quad (85)$$

It simplifies matters at this point to go to unitary gauge $\phi(x) = 1$. Since all group representations are in play, we will need the SU(2) character expansion

$$\exp\left[\frac{1}{2}\gamma\text{Tr}[U]\right] = \sum_j c_j(\gamma)\chi_j[U], \quad (86)$$

where

$$c_j(\gamma) = 2(2j+1) \frac{I_{2j+1}(\gamma)}{\gamma} \approx \frac{2j+1}{(2j+1)!} \left(\frac{\gamma}{2}\right)^{2j}. \quad (87)$$

We then have for the off-diagonal terms in M_γ , to lowest order in γ ,

$$M_\gamma(C, C_1) = \tilde{\gamma}^{L(C)+2T} \int DU \prod_{l \in \Lambda} \sqrt{2j_l + 1} U^{(j_l)}(l) \\ \times \prod_{x \in \mathcal{V}} B_{b_1 \dots b_m}^{a_1 \dots a_n}(x, \kappa_x) \prod_{l' \in \Lambda} \sum_{j'_l} c_{j'_l} \text{Tr}[U^{(j'_l)}(l')]. \quad (88)$$

Carrying out the U integrations and using (87),

$$M_\gamma(C, C_1) = \tilde{\gamma}^{L(C)+2T} \prod_{l \in \Lambda} \sum_{j_l} \frac{\sqrt{2j_l + 1}}{(2j_l + 1)!} \left(\frac{\gamma}{2}\right)^{2j_l} \\ \times \left(\prod_{x \in \mathcal{V}} B_{b_1 \dots b_m}^{a_1 \dots a_n}(x, \kappa_x) \right)_{\text{contracted}} \quad (89)$$

Each upper (lower) index in a B array at vertex site x is associated with a lower (upper) index on a link attached to that site. The meaning of $(\prod B)_{\text{contracted}}$ is that each upper (lower) index in $B(x)$ associated with a particular link is contracted with the corresponding lower (upper) index associated with the same link in a B array at a neighboring vertex; i.e., the pattern is

$$B_{\dots}^{a \dots}(x) B_{\dots}^{\dots}(x + \hat{k}) \quad \text{or} \quad B_{\dots}^{\dots}(x) B_{\dots}^{a \dots}(x + \hat{k}). \quad (90)$$

Unfortunately we have no general formula for $(\prod B)_{\text{contracted}}$, but we can argue for a rough upper bound based on the normalization condition (65). Suppose all array elements have about the same magnitude. Then (65) requires that

$$|B_{b_1 b_2 \dots b_m}^{a_1 a_2 \dots a_n}(x, \{j, j'\}, \kappa)| \sim \prod_{i=1}^n \frac{1}{\sqrt{2j_i + 1}} \prod_{k=1}^m \frac{1}{\sqrt{2j'_k + 1}}, \quad (91)$$

where the products are over links entering and leaving vertex x . In that case, it is easy to see that $(\prod B)_{\text{contracted}} \sim 1$. In fact, the assumption of equal magnitude array elements results in

a large overestimate, as can be seen for the case that \mathcal{C}_1 is a simple open contour composed of links in representation j , in which case

$$B_{s_2}^{s_1} B_{s_3}^{s_2} \dots B_{s_{L+1}}^{s_L} = \frac{1}{(2j+1)^{L/2}} \delta_{s_{L+1}}^{s_1}. \quad (92)$$

The reason for the overestimate is that because $B_b^a = \frac{1}{\sqrt{2j+1}} \delta_b^a$, the magnitude of the ‘‘average’’ array element is really $(2j+1)^{-3/2}$, rather than $(2j+1)^{-1}$.

In a little more generality, consider the product of a set of B arrays with a total of N upper indices and an equal number of lower indices, and, initially, no sum over indices. Each index is associated with an $SU(2)$ representation j , with the index running from 1 to $2j+1$. Consider choosing each index at random, within its allowed range, and let

$$\begin{aligned} \bar{B} &= \prod_{i=1}^n \frac{1}{2j_i+1} \prod_{k=1}^m \frac{1}{2j'_k+1} \sum_{a_1=1}^{2j_1+1} \sum_{a_2=1}^{2j_2+1} \dots \sum_{a_n=1}^{2j_n+1} \\ &\quad \sum_{b_1=1}^{2j'_1+1} \sum_{b_2=1}^{2j'_2+1} \dots \sum_{b_m=1}^{2j'_m+1} |B_{b_1 b_2 \dots b_m}^{a_1 a_2 \dots a_n}| \end{aligned} \quad (93)$$

be the average of the moduli of the array elements in a given B array. Then the expectation value of the modulus of the product of this random choice of (real-valued) array elements is simply the product of average values, i.e.,

$$\begin{aligned} &\langle |B_{\dots}(x_1) B_{\dots}(x_2) \dots B_{\dots}(x_n)| \rangle \\ &= \langle |B_{\dots}(x_1)| |B_{\dots}(x_2)| \dots |B_{\dots}(x_n)| \rangle \\ &= \bar{B}(x_1) \bar{B}(x_2) \dots \bar{B}(x_n). \end{aligned} \quad (94)$$

Now suppose, in the first line of (94), that we pair each upper index with a lower index such that the paired indices belong to different B 's, and assign the same value to each paired index. This reduces the number of indices which can be chosen randomly from $2N$ to N , but under a random choice of the remaining N index values the expectation value of any $|B_{\dots}(x)|$ appearing in the product is again $\bar{B}(x)$. The pairing restriction introduces a weak correlation among the different B 's in the product, but if we ignore this correlation then the expectation value of the modulus of the product with paired indices is still (94). If we denote the values of the i -th set of paired indices as a_i , a_i , and then sum over those values, we then have the estimate

$$\begin{aligned} &\sum_{a_1=1}^{2j_1+1} \sum_{a_2=1}^{2j_2+1} \dots \sum_{a_N=1}^{2j_N+1} B_{\dots}(x_1) B_{\dots}(x_2) \dots B_{\dots}(x_n) \\ &\approx \left(\prod_{i=1}^N (2j_i+1) \right) \bar{B}(x_1) \bar{B}(x_2) \dots \bar{B}(x_n). \end{aligned} \quad (95)$$

To support the validity of this approximation we return to the simplest case, (92). Since the average of the each array is $(2j+1)^{-3/2}$, and summation over each contracted index gives a factor of $(2j+1)$, the approximation delivers the correct overall factor $(2j+1)^{-L/2}$.

Now, under the constraint (65), it is easy to show that the average value \bar{B} of any array

$$B_{b_1 b_2 \dots b_m}^{a_1 a_2 \dots a_n}, \quad a_i = 1, 2, \dots, 2j_i + 1, \quad b_i = 1, 2, \dots, 2j'_i + 1 \quad (96)$$

is maximized when all array elements are identical, and equal to the right-hand side of (91). This then leads us to the upper bound

$$\left| \left(\prod B \right)_{\text{contracted}} \right| \leq 1, \quad (97)$$

and from that we obtain

$$|M_\gamma(C, \mathcal{C}_1)| < \tilde{\gamma}^{L(C)+2T} \prod_{l \in \Lambda} \frac{\sqrt{j_l+1}}{(2j_l+1)!} \left(\frac{\gamma}{2} \right)^{2j_l}. \quad (98)$$

Now summing over all possible clusters, we have from the Gershgorin theorem

$$r_C < \tilde{\gamma}^{L(C)+2T} \sum_{\Lambda} \prod_{l \in \Lambda} \sum_{j_l} \frac{\sqrt{j_l+1}}{(2j_l+1)!} \left(\frac{\gamma}{2} \right)^{2j_l} \prod_{x \in \mathcal{V}} n_s(x, \{j, j'\}), \quad (99)$$

where $n_s(x, \{j, j'\})$ is the number of singlets (which may be zero, if the set does not form a cluster) at vertex site x , and this number depends on the representations $\{j, j'\}$ of links entering/leaving the site. An upper bound on the number of singlets that can be formed is the number of orthogonal states that can be formed at x , i.e.,

$$n_s < \prod_i (2j_i + 1), \quad (100)$$

where the product is over each link attached to the site. Then absorbing two factors of $2j_l+1$ (from each end of link l) into the product over links,

$$r_C < \tilde{\gamma}^{L(C)+2T} \sum_{\Lambda} \prod_{l \in \Lambda} \sum_{j_l} \frac{(2j_l+1)^{5/2}}{(2j_l+1)!} \left(\frac{\gamma}{2} \right)^{2j_l}. \quad (101)$$

and we note that

$$\begin{aligned}
\sum_{j_l=\frac{1}{2}, 1, \frac{3}{2}, 2, \dots} \frac{(2j_l+1)^{5/2}}{(2j_l+1)!} \left(\frac{\gamma}{2}\right)^{2j_l} &< \sum_{n=1}^{\infty} \frac{(n+1)^3}{(n+1)!} \left(\frac{\gamma}{2}\right)^n \\
&= \frac{1}{4}(e^{\gamma/2}\gamma^2 + 6e^{\gamma/2}\gamma + 4e^{\gamma/2} - 4) \\
&= 2\gamma + O(\gamma^2). \quad (102)
\end{aligned}$$

Consequently, proceeding along the lines of the previous subsection,

$$r_C < \tilde{\gamma}^{L(C)+2T} \sum_{L=L_0}^{\infty} 5^L (2\gamma)^L < \tilde{\gamma}^{L(C)+2T} \frac{(10\gamma)^{L_0}}{1-10\gamma}, \quad (103)$$

where we recall that L_0 is the minimal distance on the lattice between sites x, y . So this is the bound on the sum of off-diagonal terms. The diagonal term from M_β , to leading order in β , is

$$M_\beta(C, C) = \left(\frac{\beta}{4}\right)^{L(C)T}, \quad (104)$$

which means that $r_C \ll M_\beta(C, C)$ providing that $\tilde{\gamma} \ll \tilde{\beta} \ll 1$ with T small enough so that condition (78) is satisfied, while convergence of the sum in (103) also requires

$$\gamma \ll \frac{1}{10}. \quad (105)$$

Of course one can go on to consider the more general case

$$r_{C_2} = \sum_{C_1 \neq C_2} |M_T(C_2, C_1)|, \quad (106)$$

and, by the previous analysis, establish an upper bound

$$r_{C_2} < \tilde{\gamma}^{2T} \prod_{l \in \Lambda_2} \frac{\sqrt{j_l+1}}{(2j_l+1)!} \left(\frac{\gamma}{2}\right)^{2j_l} \frac{(10\gamma)^{L_0}}{1-10\gamma}, \quad (107)$$

with a diagonal term

$$M_\beta(C_2, C_2) < \left(\frac{\beta}{4}\right)^{L(C_2)T}. \quad (108)$$

But we are interested in the largest eigenvalue of M_T . Let C_{0xy} be the contour of minimal distance L_0 between x and y . Then the largest diagonal term $M_T(C, C)$ corresponds to $C = C_{0xy}$, while in general $r_C \ll M_T(C_{0xy}, C_{0xy})$ for any C . It follows that the largest eigenvalue is within $r_{C_{0xy}}$ of $M_T(C_{0xy}, C_{0xy})$, which for small T is a small correction to the pure gauge value. This implies that the logarithmic time derivative of the largest eigenvalue satisfies the S_c confinement condition.

C. Summation over disconnected clusters

Once we include the B^2 term we must also consider matrix elements such as $M_T(C, \{C_1, C_2, \dots\})$, where C_{1xy}, C_2, \dots are disjoint clusters having no links in common. Only connected diagrams contributing to M_T are relevant. Some simple examples of relevant strong coupling diagrams connecting such disjoint clusters are shown in Fig. 12(a), which is a contribution to M_{mix} , and Fig. 12(b), which is a contribution to M_β . We are mainly concerned with contributions of the type M_{mix} , since off-diagonal contributions in M_β are known to be compatible with S_c confinement in pure gauge theories. The goal is to place an upper bound on the contribution to the Gershgorin disk radius r_C due to the sum

$$\sum_{\{C\}_{\text{dis}} \neq C} |M_{\text{mix}}(C, \{C\}_{\text{dis}}; \beta, \gamma)|, \quad (109)$$

where the sum is over all sets of clusters $\{C_{1xy}, C_2, \dots\}$ which are disconnected in the sense that they have no links in common, and the dependence of M_{mix} on both couplings is indicated here explicitly.

Let us define a ‘‘ribbon’’ to be a one-plaquette wide strip of plaquettes bounded by Higgs lines on either side, both originating from an expansion of the action in powers of β, γ , beginning at a link l , and terminating on a cluster, as shown in Fig. 13(a). Keeping l fixed, we consider summing

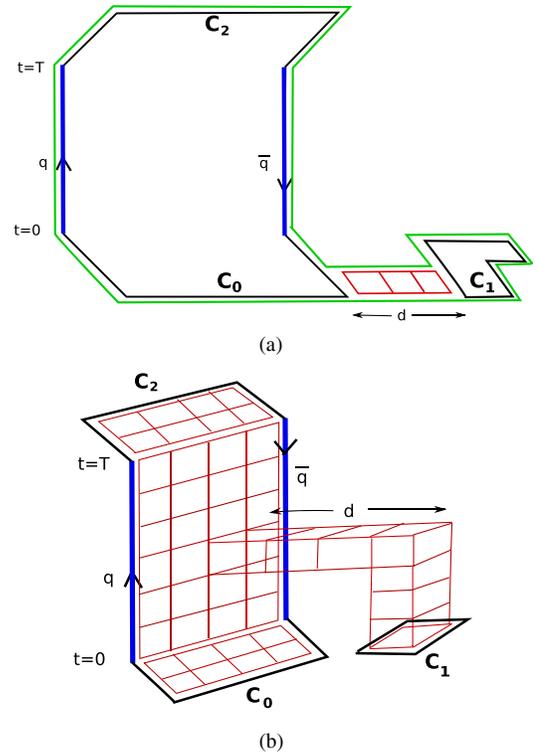


FIG. 12. Connected diagrams for disconnected clusters in the initial state. (a) screening contributions; (b) confining contributions.

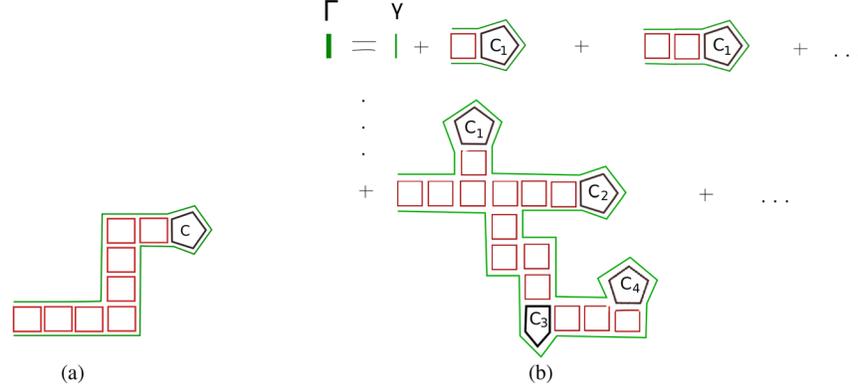


FIG. 13. (a) A “ribbon” diagram. This is a chain, one lattice spacing in width, consisting of a series of plaquettes (red) taken from the expansion of the Wilson action, bordered by links taken from the expansion of the Higgs action (green), and terminating in a cluster C (black). (b) Schematic representation of a sample of diagrams which are implicitly summed in Eq. (114).

over all ribbons, and for each ribbon summing over the clusters at the endpoint. Assuming for simplicity that the Higgs lines and plaquettes are in the fundamental representation (and this is not an important restriction, since the sum over representations is rapidly convergent), and that the ribbon is of length d , the ribbon is associated (after integration over gauge and Higgs fields) with a weight

$$\left(\frac{\beta}{4}\right)^d \left(\frac{\gamma}{4}\right)^{2d} 3F(C), \quad (110)$$

where $F(C)$ is the contribution from the cluster. The factor of 3 comes from the fact that the cluster may be attached at either of the three sides of the final plaquette in the ribbon. Summing over all ribbons and all clusters, we have

$$\begin{aligned} Q(\beta, \gamma) &< \sum_{d=1}^{\infty} \left(\frac{\beta}{4}\right)^d \left(\frac{\gamma}{4}\right)^{2d} 9^d \mathcal{F}(\beta, \gamma) \\ &< \frac{2^{-6} 9 \gamma^2 \beta}{1 - 2^{-6} 9 \gamma^2 \beta} \mathcal{F}(\beta, \gamma), \end{aligned} \quad (111)$$

where

$$\mathcal{F}(\beta, \gamma) = 3 \sum_C F(C), \quad (112)$$

and the inequality in (111) follows from the fact that there is some overcounting on the right-hand side, since a self-avoiding constraint on the sum has not been imposed. The factor of 9^d derives from the fact that in building a ribbon plaquette by plaquette, then at the n th plaquette there are three links at which to join the $n + 1$ th plaquette, which may be any one of the three plaquettes adjoining that link which does not backtrack on the ribbon. The smallest cluster is composed of four links. If we sum up only γ -dependent contributions, then by previous methods we obtain an upper bound

$$\mathcal{F}(\gamma \text{ only}) < 3 \frac{(10\gamma)^4}{1 - 10\gamma}. \quad (113)$$

For small clusters there may also be significant β contributions, but these can be neglected for large clusters in which plaquettes would have to span large areas.

Now define the “dressed link” factor Γ to be the solution of

$$\Gamma = \gamma + Q(\beta, \Gamma). \quad (114)$$

Diagrammatically, Γ is a sum of all treelike structures connecting disconnected clusters, as indicated schematically in Fig. 13(b). A bound on the contributions to r_C from such treelike structures is obtained from the bound (103) by replacing the γ factors that arise in the sum over initial states by the dressed link factor Γ , i.e.,

$$\tilde{\gamma}^{L(C)+2T} \frac{(10\Gamma)^{L_0}}{1 - 10\Gamma}. \quad (115)$$

The reason is that this replacement accounts for the sum over all treelike arrangements of disconnected clusters which terminate on links of the “trunk” \mathcal{C}_{xy} , which is also summed over.

Disconnected diagrams can also be joined by tubelike structures, such as the one appearing in Fig. 12(b). An estimate, at $T = 1$, of the sum of tubes at constant time leaving a given plaquette and terminating in a cluster is

$$\begin{aligned} P(\beta, \gamma) &< \sum_{d=1}^{\infty} \left(\frac{\beta}{4}\right)^{4d} 15^d \mathcal{F}(\beta, \gamma) \\ &< 15 \left(\frac{\beta}{4}\right)^4 \frac{\mathcal{F}(\beta, \gamma)}{1 - 15\left(\frac{\beta}{4}\right)^4}. \end{aligned} \quad (116)$$

where again the right-hand side is actually an overestimate, due to ignoring a self-avoidance constraint. The factor of 15 is the number of ways that a tube of a given length can

be extended at one end by one cube, and convergence requires $\beta \ll 4/15^{1/4} \approx 2$. The tube can be extended, without backtracking, at one of five plaquettes at the end of the tube, in one of three possible steps (counting backward/forward in one of the possible directions) orthogonal to the plaquette. Then in complete analogy to (114) we define a “dressed plaquette” factor

$$\bar{\beta} = \beta + P(\bar{\beta}, \gamma). \quad (117)$$

Taking account of both tube and ribbon structures, we obtain the simultaneous equations

$$\Gamma = \gamma + Q(\bar{\beta}, \Gamma), \quad \bar{\beta} = \beta + P(\bar{\beta}, \Gamma). \quad (118)$$

The prescription for including tree diagrams joining disconnected clusters, in order to get a bound on the right-hand side of (109), is to replace the couplings β, γ in matrix elements for a single cluster in the initial state by the dressed factors $\bar{\beta}, \Gamma$. But in the end, for $\tilde{\gamma} \ll \bar{\beta} \ll 1$, the solution of (118) is simply $\Gamma \approx \gamma, \bar{\beta} \approx \beta$, and the previous upper bound on r_C is not much affected.

D. $\beta \ll \lambda \ll 1/10$

As β is reduced, the time interval required for string breaking is also reduced. For $\gamma \ll 1/10$, however, Ψ_V is always S_c confining. It is sufficient to consider the limiting case of $\beta \rightarrow 0$. For V a simple Wilson line, we have from Eq. (50) that

$$E_V = -\log [W(L, 1)] = (-2 \log \tilde{\gamma}^2)L - \log(2\tilde{\gamma}^2). \quad (119)$$

which is obviously bounded by a linear potential. In the more general case, $M_T(C_2, C_1) = M_\gamma(C_2, C_1)$. Then, using the Gershgorin Theorem and (103), the largest eigenvalue of M_T for $T = 1$ is bounded from above by

$$\lambda_{\max}(1) < 2\tilde{\gamma}^{2L_0+2} + \tilde{\gamma}^{L_0(C)+2} \frac{(10\gamma)^{L_0}}{1 - 10\gamma} \quad (120)$$

and the logarithm gives a lower bound to the energy, which again increases linearly with separation L_0 . Beyond $T = 1$ we have string breaking. To refine the estimate of the string-breaking time, one may go to a time-asymmetric lattice, with the lattice spacing a_t in the time direction much smaller than the lattice spacing a in the space directions. This asymmetry is accompanied by an increase in the lattice coupling β_t associated with the timelike plaquettes. Eventually β_t exceeds γ , and the preceding analysis can be applied.

E. Summary

The strong-coupling argument presented in this section is a bit lengthy, and despite the length it does not rise to the level of rigor required of a formal proof. But the central idea

is simple, and it boils down to this: In the absence of matter loops, the energies of the Ψ_V states are given by pure gauge theory, which we already know to be S_c confining. Inclusion of matter loops will eventually cause string breaking and a consequential loss of the linear potential in the course of Euclidean time evolution, but this event occurs only after the system has evolved for some finite time period. If the strong-coupling conditions (46) are satisfied and the Euclidean time T obeys the bound (78), then the multiplicity of screening contributions is outweighed by their exponential suppression in powers of γ , and the energy of a time-evolved Ψ_V state, obtained from the lattice logarithmic time derivative, is approximately that of the pure gauge theory. Even for $\beta \ll \gamma \ll 1/10$, the S_c condition is satisfied at $T = 1$. Hence S_c -confinement exists in some region of the $\beta - \gamma$ phase diagram, and given the known result [8] that S_c -confinement does not exist everywhere in the phase plane, it follows that there is somewhere a transition line between the stronger separation-of-charge property in the confinement region and the weaker color-neutrality property in the Higgs region.

VII. SYMMETRY BREAKING AND THE S_c -TO-C TRANSITION

We conjecture that the transition from S_c to C confinement coincides with the gauge-invariant, symmetry-breaking transitions seen in Figs. 4 and 7. The first question to ask is whether existing data on the location of C confinement, in the SU(2) gauge-Higgs theory, already rule this out.

In Ref. [8], we considered three possible choices of Ψ_V states: the Dirac state (a non-Abelian generalization of charged states in an Abelian theory), a “pseudomatter” state based on eigenmodes of the covariant Laplacian operator, and a “fat link” Wilson line state derived from a familiar method of noise reduction in lattice gauge theory. An S_c -to-C confinement transition was found for the first two states, but not for the third, which was everywhere S_c confining. But it must be understood that if a region is S_c confining, this behavior must be obtained not just for one choice of Ψ_V state, but for *all* such states. In other words, S_c confining behavior in a particular Ψ_V (e.g., the fat link state) in some region is a necessary but not sufficient condition for S_c confinement in that region. If even one Ψ_V exhibits C confining behavior in a region, then that region is C confining. Put another way, C confining behavior found for one state Ψ_V in some region is a sufficient but not a necessary condition for C confinement in the region.

The transition from S_c to C confinement in the Dirac state corresponds, as explained in ref. [8], to the spontaneous breaking of a remnant gauge symmetry, global on each time slice, that exists in Coulomb gauge, and the location of that remnant symmetry breaking in SU(2) gauge-Higgs theory was found in ref. [2]. It is certain that the region in the $\beta - \gamma$ phase diagram above the remnant

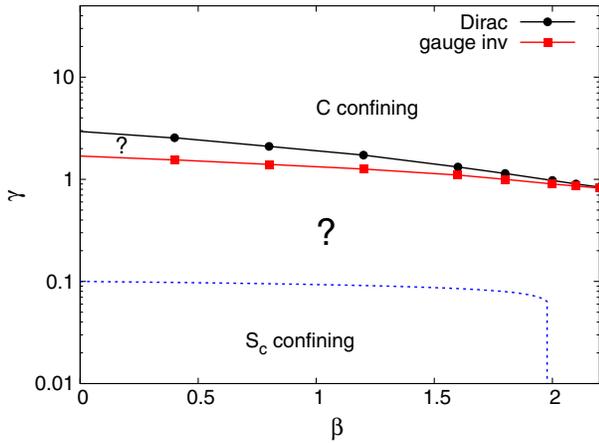


FIG. 14. For SU(2) gauge-Higgs theory, C confinement exists above the line denoted “Dirac,” and S_c -confinement exists in a strong-coupling region, as well as along the line at $\beta = 0$. The location of C and S_c confinement in the rest of the phase diagram is uncertain. Our conjecture is that the S_c -to-C confinement transition line coincides with the gauge-invariant, symmetry-breaking line, denoted “gauge inv” in the figure. Error bars on data points are, on this scale, smaller than the symbol size.

symmetry breaking line is C confining. But whether the region below this line is S_c confining is uncertain, at least until we come to the region of strong couplings, where the analysis of the previous section shows the existence of S_c confinement.

The situation at the moment is illustrated in Fig. 14. C confinement is known to exist above the Dirac line shown, but we do not know how far it extends below that line. S_c confinement exists inside a strong-coupling region, whose boundary is indicated somewhat schematically in Fig. 14, but we do not know how far it extends outside the region of convergence of the strong-coupling expansion. If one can find a V operator such that Ψ_V has C confining behavior anywhere below the gauge-invariant, symmetry-breaking line, also displayed in Fig. 14, then our conjecture about the coincidence of the S_c -to-C and symmetry breaking transitions is wrong. Only two points for the pseudomatter transition were obtained in [8], and one of these points (at $\beta = 1.2$) lies at a gamma value which is slightly below the corresponding Dirac operator transition.⁶ That point is still above the gauge-invariant transition, however. This means that at least some of the C confining region lies below the Dirac transition, and the conjecture is that the entire region between the Dirac line and the gauge-invariant transition line is C confining, while the region below the gauge-invariant transition line is S_c confining.

So the existing data is at least consistent with our conjecture. To proceed further, some effort must be devoted

⁶The other point, at $\beta = 2.2$, coincides with both the Dirac and gauge-invariant transitions.

to inventing and testing more operators which might falsify (or, alternatively, support) this proposal. A first step would be to test operators, already studied in [8] for SU(2) gauge-Higgs theory, in the SU(3) gauge-Higgs case. We hope to report on these efforts at a later time.

VIII. CONCLUSIONS

There exist global symmetries in the Higgs sector of gauge-Higgs theories which are independent of any gauge choice, and these symmetries can break spontaneously in the sense explained in Sec. II, where we also explain the absence of Goldstone excitations in the full theory. We have constructed gauge-invariant order parameters which can detect the spontaneous breaking of such symmetries. There are two obvious questions, both relating to the nature of this transition. First, given the result of Osterwalder and Seiler [3], this transition cannot correspond to a thermodynamic transition everywhere along the transition line. So is it possible to speak of a phase transition which does not correspond to a nonanalyticity in the free energy? In fact, there are examples of such transitions, namely the Kertesz transition line [25] found in Ising and Potts models in an external magnetic field.⁷ But the next question is what is the physical difference between the symmetric and broken phases in a gauge-Higgs theory. If there is no physical difference and no singular behavior in the free energy, then this transition is physically meaningless. However, we believe there is a natural candidate for the physical difference between the two phases, and that is the distinction between separation-of-charge (S_c) confinement and color (C) confinement. In Sec. VI, we have shown that S_c confinement must exist somewhere in the $\beta - \gamma$ coupling plane, and given the fact [8] that S_c confinement does not exist throughout the plane, there must be a transition between these physically distinct phases.

So we will conclude this article by repeating the conjecture, made in the previous section, that the gauge-invariant global symmetry-breaking transition that we have located in gauge-Higgs theory coincides with the transition that must exist between the S_c and C confinement. If so, this transition separates two phases that can be meaningfully distinguished as confinement vs Higgs, in which a global symmetry is either unbroken, or spontaneously broken in sense explained above.

ACKNOWLEDGMENTS

We thank Axel Maas for helpful discussions. J.G.’s research is supported by the U.S. Department of Energy under Grant No. DE-SC0013682.

⁷The possible relevance of this example in the context of pure gauge theories at finite temperature, and in gauge-Higgs models, has been discussed by a number of authors [19,26].

- [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. (N.Y.)* **110**, 440 (1978).
- [4] E. H. Fradkin and S. H. Shenker, *Phys. Rev. D* **19**, 3682 (1979).
- [5] J. Frohlich, G. Morchio, and F. Strocchi, *Nucl. Phys.* **B190**, 553 (1981); *Phys. Lett.* **97B**, 249 (1980).
- [6] G. 't Hooft, *NATO Sci. Ser. B* **59**, 117 (1980).
- [7] A. Maas and P. Törek, *Phys. Rev. D* **95**, 014501 (2017); A. Maas, R. Sondenheimer, and P. Törek, [arXiv:1709.07477](https://arxiv.org/abs/1709.07477); in Proceedings of Lattice 2017, Granada, Spain, *Eurphys. J. Web Conf.* **175**, 08002 (2018).
- [8] J. Greensite and K. Matsuyama, *Phys. Rev. D* **96**, 094510 (2017).
- [9] T. Kugo and I. Ojima, *Prog. Theor. Phys. Suppl.* **66**, 1 (1979).
- [10] M. Stingl, *Phys. Rev. D* **34**, 3863 (1986); **36**, 651 (1987).
- [11] K. Fredenhagen and M. Marcu, *Phys. Rev. Lett.* **56**, 223 (1986).
- [12] S. Willenbrock, [arXiv:hep-ph/0410370](https://arxiv.org/abs/hep-ph/0410370).
- [13] S. Weinberg, *The Quantum Theory of Fields*, Vol. 2 (Cambridge University Press, Cambridge, England, 2013).
- [14] A. Maas, [arXiv:1712.04721](https://arxiv.org/abs/1712.04721).
- [15] J. Ranft, J. Kripfganz, and G. Ranft, *Phys. Rev. D* **28**, 360 (1983).
- [16] C. Gatttringer and C. B. Lang, *Lect. Notes Phys.* **788**, 1 (2010).
- [17] M. Creutz, *Phys. Rev. D* **21**, 2308 (1980).
- [18] T. Kennedy and C. King, *Commun. Math. Phys.* **104**, 327 (1986).
- [19] K. Langfeld, in *Proceedings: Strong and Electroweak Matter 2002*, edited by M. Schmidt (World Scientific, Singapore, 2003).
- [20] C. B. Lang, C. Rebbi, and M. Virasoro, *Phys. Lett.* **104B**, 294 (1981); I. Campos, *Nucl. Phys.* **B514**, 336 (1998); W. Langguth, I. Montvay, and P. Weisz, *Nucl. Phys.* **B277**, 11 (1986); J. Jersak, C. B. Lang, T. Neuhaus, and G. Vones, *Phys. Rev. D* **32**, 2761 (1985).
- [21] C. Bonati, G. Cossu, M. D'Elia, and A. Di Giacomo, *Nucl. Phys.* **B828**, 390 (2010).
- [22] F. Knechtli, R. Sommer, *Phys. Lett. B* **440**, 345 (1998); **454**, 399(E) (1999); O. Philipsen and H. Wittig, *Phys. Rev. Lett.* **81**, 4056 (1998); **83**, 2684(E) (1999).
- [23] M. Greiter, *Ann. Phys. (N.Y.)* **319**, 217 (2005).
- [24] G. H. Golub and C. F. Van Loan, *Matrix Computations* (3rd edition) (Johns Hopkins University Press, Baltimore, 1996).
- [25] J. Kertesz, *Physica (Amsterdam)* **161A**, 58 (1989).
- [26] S. Fortunato and H. Satz, *Phys. Lett. B* **475**, 311 (2000); S. Wenzel, E. Bittner, W. Janke, A. M. J. Schakel, and A. Schiller, *Phys. Rev. Lett.* **95**, 051601 (2005); R. Bertle, M. Faber, J. Greensite, and S. Olejnik, *Phys. Rev. D* **69**, 014007 (2004).