't Hooft loop in SU(2) lattice gauge theories

E. Tomboulis
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544
(Received 6 October 1980)

We examine the behavior of the 't Hooft loop (magnetic disorder parameter) for weak coupling ($\beta \to \infty$) in the standard SU(2) lattice gauge theory. We find that it exhibits length-law falloff due to the presence of dynamical Z_2 monopoles. These dynamical monopoles are a consequence of the nonconservation of Z_2 flux, and the possibility of flux spreading in the non-Abelian theory. The implications for quark confinement are discussed.

I. INTRODUCTION

The introduction of the lattice by Wilson[1] provides a nonperturbative framework for the formulation and study of field theory. Lattice gauge theories have been shown[1,2] to exhibit confinement of static sources in the strong-coupling regime. One hopes that, in the case of non-Abelian gauge groups, this confining phase persists for all couplings. By letting the bare coupling vanish as the lattice spacing goes to zero holding physical masses fixed, one could recover the continuous asymptotically free theory, while preserving confinement at large distances. However, no proof of confinement for arbitrarily weak bare coupling yet exists, although numerical computation[3] and semiclassical techniques[4] support the scenario of no phase transition.

The Wilson loop[1,5] is the natural order parameter for distinguishing the phases of a pure gauge theory. 't Hooft introduced[6] a magnetic disorder parameter which is generally referred to as the 't Hooft loop. It satisfies simple commutation relations with the Wilson loop. From these commutation relations 't Hooft tried to argue that this new parameter could be used as an alternative criterion for confinement. However, it turns out that it is dual to the Wilson loop only in a limited sense (see Sec. III), and knowledge of its behavior cannot directly produce the corresponding behavior of the Wilson loop. Nevertheless, it is computationally easier to handle than the Wilson loop, and a valuable tool in investigating possible phases.

In physical terms, the 't Hooft loop represents an external source of (color) magnetic flux. Therefore, naively one would expect that in a confining phase, where the magnetic flux is defocused, the loop exhibits perimeter-law behavior. This is indeed true for strong coupling. However, the 't Hooft commutation relations allow the possibility of simultaneous area-law behavior for both the Wilson and 't Hooft loops.[6] In fact, Mack and Petkova[7] considered a modified SU(2) lattice gauge theory, for which they proved that the 't Hooft loop has area-law behavior in the weak-coupling regime. Nevertheless, one still expects confinement for weak coupling in this model.[7,8]

In this paper we investigate the behavior of the 't Hooft loop for weak coupling in the standard (Wilson) SU(2) lattice gauge theory. In Sec. II we rewrite the theory in terms of new variables which allow us to isolate the presence of dynamical monopoles[9] in the theory. These monopoles are sources of Z_2-color magnetic flux. The 't Hooft operator is precisely defined and rewritten in the new variables, and its physical meaning discussed. In Sec. III we further transform the theory by performing a dual transformation with respect to the center Z_2 of the group SU(2).[7,9,10] In Sec. IV we examine the behavior of the 't Hooft loop for weak coupling. We find perimeter-law behavior due to screening by the dynamical Z_2 monopoles. Finally, Sec. V contains a discussion of the result and of the implications for quark confinement.

II. DYNAMICAL Z_2 MONOPOLES

We work on a hypercubic lattice $\Lambda \subset Z^d$ $(d=3,4)$ containing r-cells $[c_r]$, $r=0,\ldots,d$, i.e., sites $[s]$, bonds $[b]$, plaquettes $[p]$, cubes $[c]$, and hypercubes $[h]$. An r-cell on Λ corresponds to a $(d-r)$-cell on the dual lattice Λ^*. Each r-cell c_r is assigned a standard orientation; $-c_r$ denotes the opposite orientation. Boundary and coboundary operators are denoted by ∂ and $\hat{\partial}$, respectively. With appropriate orientations, $\partial^2=0$ and $\hat{\partial}^2=0$. An arbitrary set of r-cells X is said to be closed if $\partial X=0$, and co-closed if $\hat{\partial}X=0$.[11]

The standard SU(2) lattice gauge theory[1,2] is defined in terms of the bond variables $U[b]\in \mathrm{SU}(2)$, $U[-b]\equiv U[b]^{-1}$. The measure is given by

$$d\mu(U)\equiv \frac{1}{Z}\prod_{b\in\Lambda} dU[b]\exp[L(U)]\,,$$
$$Z\equiv\int\prod_{b\in\Lambda} dU[b]\exp[L(U)]\,, \tag{1}$$

where $L(U)$ is the standard (Wilson) Euclidean action

$$L(U)=\sum_p \beta\,\mathrm{tr}U[\partial p] \tag{2}$$

and $\beta=1/g^2$ with g the standard bare gauge coupling constant. $U[\partial p]$ denotes the product of bond variables around the perimeter of the plaquette p. $dU[b]$ is the normalized invariant Haar measure on SU(2). The expectation values of functions $F(U)$ of the $U[b]$'s are then given by

$$\langle F\rangle=\int d\mu(U)F(U)\,. \tag{3}$$

For our considerations it is necessary to separate the Z_2 part of the theory from the $SU(2)/Z_2$ part. Variables in Z_2 [the center of SU(2)] will always be denoted by Greek letters, e.g., $\gamma,\sigma,\tau,\ldots$; they take the values ±1. When integrating over Z_2 we always use the normalized Haar measure:

$$\int d\gamma(\cdots)\equiv\tfrac{1}{2}\sum_{\gamma=\pm1}(\cdots),\quad \gamma\in Z_2\,. \tag{4}$$

If $\gamma[c_r]$ is a Z_2 variable defined as r-cells, then $\gamma[-c_r]\equiv\gamma[c_r]^{-1}$. Note, however, that for any Z_2 variable γ, we have $\gamma=\gamma^{-1}$, so in this sense orientations are actually irrelevant. If X is a set of r-cells, then we use the notation $\gamma[X]\equiv\prod_{c_r\in X}\gamma[c_r]$.

To effect the above-mentioned separation, insert $1=\int\prod_b d\gamma[b]$ and shift $U[b]\to U[b]\gamma[b]$, $\gamma\in Z_2$, in the partition function Z in (1). We obtain

$$Z=\int\prod_b dU[b]\prod_b d\gamma[b]\prod_p d\sigma[p]\times\prod_p\delta(\gamma[\partial p]\eta[p]\sigma^{-1}[p])\exp\Big(\sum_p K_p\sigma[p]\Big),$$

where we wrote

$$\beta\,\mathrm{tr}U[\partial p]\equiv\beta\,|\mathrm{tr}U[\partial p]|\,\eta[p]\,,$$
$$K_p\equiv\beta\,|\mathrm{tr}U[\partial p]|\,,\quad \eta[p]=\pm1\in Z_2\,.$$

The δ function over Z_2 is defined by

$$\delta(\tau)=\begin{cases}2 & \text{if } \tau=1\\ 0 & \text{if } \tau=-1\end{cases},\quad \int d\tau\,\delta(\tau)=1 \tag{5}$$

and has a character expansion

$$\delta(\tau)=2\int d\sigma\,\chi_\sigma(\tau)\,, \tag{6}$$

where

$$\chi_\sigma(\tau)=\chi_\tau(\sigma)=\begin{cases}1 & \text{if } \sigma=1\,,\\ \tau & \text{if } \sigma=-1\end{cases} \tag{7}$$

are the Z_2 characters.

Using (6) in the above expression for Z yields

$$Z=\text{const}\times\int\prod_b dU[b]\prod_b d\gamma[b]\prod_p d\sigma[p]\prod_p d\tau[p]\prod_p\chi_{\tau[p]}[\sigma^{-1}[p]\eta[p]]\chi_{\tau[p]}[\gamma[\partial p]]\exp\Big(\sum_p K_p\sigma[p]\Big)$$
$$=\text{const}\times\int\prod_b dU[b]\prod_p d\sigma[p]\prod_p d\tau[p]\prod_p\chi_{\tau[p]}[\sigma^{-1}[p]\eta[p]]\prod_b\big(\tfrac{1}{2}\delta[\tau[\hat{\partial} b]]\big)\exp\Big(\sum_p K_p\sigma[p]\Big).$$

Performing the constrained τ integration

$$Z=\text{const}\times\int\prod_b dU[b]\prod_p d\sigma[p]\prod_c\big(\tfrac{1}{2}\delta[\eta^{-1}[\partial c]\sigma[\partial c]]\big)\exp\Big(\sum_p K_p\sigma[p]\Big). \tag{8}$$

Now, the integrand in (8) is invariant under $U[b]\to U[b]\gamma[b]$, for any $\gamma[b]\in Z_2$, and thus $U[b]$ may be any representative of the cosets $\hat{U}[b]\in SU(2)/Z_2$. Indeed,

$$K_p\equiv\beta\,|\mathrm{tr}U[\partial p]|=K_p(\hat{U})\geq0 \tag{9}$$

depends only on $\hat{U}$'s, as explicitly indicated, and also

$$\eta[\partial c]=\prod_{p\in\partial c}\eta[p]=\prod_{p\in\partial c}\mathrm{sign}\,\mathrm{tr}U[\partial p]\equiv f_c(\hat{U}) \tag{10}$$

can easily be checked to actually depend only on $\hat{U}$'s, as explicitly indicated. Therefore, we may replace the U integration by integration over the coset variables $\hat{U}$, and obtain the new form of the measure

$$d\mu(\hat{U},\sigma,\xi)=\frac{1}{Z}\prod_c d\xi[c]\prod_b d\hat{U}[b]\prod_p d\sigma[p]\times\prod_c\delta[f_c^{-1}(\hat{U})\xi[c]]\prod_c\big(\tfrac{1}{2}\delta[\xi^{-1}[c]\sigma[\partial c]]\big)\times\exp\Big(\sum_p K_p(\hat{U})\sigma[p]\Big), \tag{11}$$

where $K_p(\hat{U})$, and $f_c(\hat{U})$ are defined as functions of $\hat{U}$ by (9) and (10). The new expression for the partition function Z follows from $\int d\mu(\hat{U},\sigma)\equiv 1$. The ζ integration is trivial and has been introduced for later convenience.

Equation (11) was first obtained in Ref. 9 by a somewhat different method. The derivation presented here, involving straightforward manipulation of the original expressions (1), is, we believe, more direct and easily generalized to SU(N) with any N.

The form (11) of the measure has a rather transparent physical interpretation. It describes a Z_2 gauge field $\sigma[p]$ with (nonnegative) "fluctuating" coupling constants determined by the $\hat{U}$'s, and interacting with currents $\zeta[c]$. These currents, dynamically produced by the $\hat{U}$ part of the theory, are sources of Z_2 magnetic flux, i.e., they are Z_2-(Dirac) monopole currents. Note that in four dimensions the currents $\zeta[c]=f_c(\hat{U})$ form co-closed sets of cubes (closed world lines on the dual lattice) and are conserved as a consequence of their definition (10), i.e.,

$$\prod_{c\in\partial h} f_c(\hat{U})=1 \tag{12}$$

for every hypercube $h\in\Lambda$. In three dimensions the magnetic sources reside on single cubes (single sites on the dual lattice). Cubes are, of course, trivially co-closed for $d=3$.

We now introduce the 't Hooft operator $B[C^*]$[6]—also called "magnetic-disorder parameter"—which is simply an external (classical) source of Z_2-magnetic flux. Its usual definition in terms of the original measure (1) is

$$\langle B[C^*]\rangle=\frac{1}{Z}\int\prod_b dU[b]\exp\Big(\beta\sum_p \mathrm{tr}U[\partial p](-1)^{E_{S^*}[p]}\Big). \tag{13}$$

Here C^* is a co-closed set of cubes that forms the coboundary of the set of plaquettes S^*, i.e., $C^*=\hat{\partial}S^*$. E_{S^*}, the characteristic function on the set S^*, is defined by $E_{S^*}[p]=\pm 1$ if $\pm p\in S^*$, 0 otherwise. It is easily checked that B is independent of any particular S^*; it depends only on the coboundary C^*. Indeed, if we are given two choices S^* and S'^* with $C^*=\hat{\partial}S^*$, $C^*=\partial S'^*$, then $\hat{\partial}(S^*-S'^*)=0$. This implies that $(S^*-S'^*)=\hat{\partial}L^*$ for some set of bonds L^*. The change of variables $U[b]\to(-1)U[b]$ for all $b\in L^*$ in (13) then moves S^* to S'^*. S^* may be thought of then as a Dirac sheet (Dirac string in $d=3$). In terms of the new measure (11) we have

$$\langle B[C^*]\rangle=\frac{1}{Z}\int \prod_c d\zeta[c]\prod_p d\sigma[p]\prod_b d\hat{U}[b]\prod_c \delta[f_c^{-1}(\hat{U})\zeta[c]]\prod_c\left(\tfrac{1}{2}\delta[\zeta^{-1}[c]\sigma[\partial c]]\right)\exp\Big(\sum_p K_p(\hat{U})\sigma[p](-1)^{E_{S_*}[p]}\Big) \tag{14a}$$

In this form the physical interpretation of $B[C^*]$ becomes rather obvious. The simple change of variables $\sigma[p]\to(-1)\sigma[p]$ for all $p\in S^*$ leads to

$$\langle B[C^*]\rangle=\frac{1}{Z}\int \prod_c d\zeta[c]\prod_b d\hat{U}[b]\prod_p d\sigma[p]\prod_c \delta[f_c^{-1}(\hat{U})\zeta[c]]\prod_c\left(\tfrac{1}{2}\delta[\zeta^{\mathrm{ext}-1}[c]\zeta^{-1}[c]\sigma[\partial c]]\right)\exp\Big(\sum_p K_p(\hat{U})\sigma[p]\Big), \tag{14b}$$

where we defined

$$\zeta^{\mathrm{ext}}[c]=\begin{cases}(-1) & \text{if } c\in C^*,\\ (+1) & \text{if } c\notin C^*.\end{cases}$$

The fact that B depends only on C^*, where it acts as an external source of Z_2-magnetic flux, is now made manifest: the effect of B appears explicitly as simply an external shift in the Z_2-monopole current residing on C^*.

$B[C^*]$ is generally referred to as the "'t Hooft loop" although C^* is a loop of cubes (closed curve on Λ^*) only for $d=4$. In three dimensions C^* is just two separated (oppositely oriented) cubes (one may be taken to be at "infinity").

The Mack-Petkova (MP) model[7] is defined by modifying the measure of the standard theory (1), (11) by the imposition of a constraint. This constraint, in terms of our new variables, is the requirement that $\zeta[c]=+1$ on all cubes, i.e., the model can be defined by multiplying (11) by $\{\prod_c\delta[\zeta[c]]\}$. Physically, this amounts to excluding all dynamical monopoles from the theory. Mack and Petkova note that this makes very little difference in the vacuum state as $\beta\to\infty$. By an application of "chessboard" estimates[12] they prove[7,9] that

$$\Big\langle\prod_{c\in G}\theta[-\zeta[c]]\Big\rangle\leq C(\beta)^{|G|}, \qquad C(\beta)=\mathrm{const}\times e^{-\mathrm{const}\times\beta}\to 0,\quad \beta\to\infty \tag{15}$$

for any collection G of $|G|$ cubes. Hence, the probability that the constraint is violated on any given cube goes to zero. However, the situation can be a great deal different in the presence of operators that couple to the degrees of freedom affected by the constraint. In particular, $\langle B[C^*]\rangle$

was shown[7] by Mack and Petkova to exhibit area-law behavior as $\beta\to\infty$ in their modified model. Whereas in the standard theory, where dynamical monopole excitation is allowed, one would expect screening of the external magnetic source and consequently perimeter-law behavior. This is the question we examine in Sec. IV.

III. DUALITY TRANSFORMATION

In the above we isolated the presence of dynamical monopoles by separating the Z_2 from the $SU(2)/Z_2$ part of the theory. We now proceed to derive still another form of the measure by performing a duality transformation, i.e., a Fourier transform with respect to Z_2, on $\sigma[p]$ in (11). The basic duality transformation with respect to the center for arbitrary SU(N), but without separating out the monopoles, has been given by Ukawa, Windey, and Guth.[10] The SU(2) case, with monopoles separated out, has been presented in Ref. 9. We will therefore be brief.

Starting from (11), expand the terms involving $\sigma[p]$ in a character expansion, i.e.,

$$\delta\left[\zeta^{-1}[c]\prod_{p\in\partial c}\sigma[p]\right]=2\int d\omega[c]\chi_{\omega[c]}\left(\zeta^{-1}[c]\prod_{p\in\partial c}\sigma[p]\right), \tag{16}$$

$$e^{K_p(\hat{U})\sigma[p]}=2\int d\alpha[p]e^{\hat{\mathcal{L}}_p[\alpha[p],\hat{U}]}\chi_{\alpha[p]}[\sigma[p]]. \tag{17}$$

The explicit expression for the dual Lagrangian $\hat{\mathcal{L}}_p$ is easily obtained by inverting (17) (Refs. 7, 9, 10, 13):

$$\hat{\mathcal{L}}_p[\gamma[p],\hat{U}]=\hat{M}_p(\hat{U})+\hat{K}_p(\hat{U})\gamma[p], \tag{18}$$

where

$$\hat{M}_p(\hat{U})\equiv\tfrac{1}{2}\ln[\tfrac{1}{2}\sinh 2K_p(\hat{U})], \tag{19}$$

$$\hat{K}_p(\hat{U})\equiv\tfrac{1}{2}\ln[\coth K_p(\hat{U})]. \tag{20}$$

Inserting (16) and (17) in (11), the σ integration can be performed. This results into δ functions that allow the $\alpha[p]$ integration to be performed also. The final result for the dual transformed measure is

$$d\hat{\mu}(\hat{U},\omega,\zeta)=\frac{1}{Z}\int\prod_c d\zeta[c]\prod_b d\hat{U}[b]\prod_c d\omega[c]\prod_c\delta[f_c^{-1}(\hat{U})\zeta[c]]\prod_c\chi_{\omega[c]}[\zeta[c]]\exp\left(\sum_p\{\hat{M}_p(\hat{U})+\hat{K}_p(\hat{U})\omega[\hat{\partial}p]\}\right). \tag{21}$$

The dual form of Z follows from $\int d\hat{\mu}(\hat{U},\omega)=1$.

The ω subsystem defines a Z_2-gauge theory interacting with external sources $\zeta[c]$, and fluctuating non-negative couplings $\hat{K}_p$ determined by the $\hat{U}$'s. Note that, $\zeta[c]$ now couple as "electric" sources to the ω system. On the dual lattice Λ^* the cube variables $\omega[c]$ appear as site variables for $d=3$ (dual Z_2-spin system), and as bond variables for $d=4$ (self-dual Z_2-gauge system).

The same treatment can be given for the expectation values of observables. Starting from (14a) or, more simply (14b), and repeating the analogous steps we obtain

$$\langle B[C^*]\rangle=\int d\hat{\mu}(\hat{U},\omega,\zeta)\prod_{c\in C^*}\omega[c]. \tag{22}$$

The form of the result (22) is as expected. The duality transformation interchanges Z_2-magnetic and Z_2-electric flux for the Z_2-gauge subsystem. Hence, the 't Hooft operator, a magnetic-flux source in the original variables, appears as an external electric source in the dual variables, i.e., as a "Z_2-Wilson loop" which is precisely the form of (22). Equation (22) shows that $B[C^*]$ is expressible solely in terms of Z_2 variables, and that it is dual only to that part of the standard Wilson loop that couples to the center of the group. In that sense it is an Abelian operator, not a fully non-Abelian order parameter such as the Wilson loop. As a consequence the behavior of the Wilson loop may not be directly deduced from a knowledge of the behavior of the 't Hooft loop.

In principle it should be possible to define a fully non-Abelian magnetic disorder parameter as an external source of arbitrary color-magnetic flux. It would appear as an ordinary Wilson loop in the fully dual version of the theory obtained by Fourier transforming with respect to the whole group and not just its center. Although such a procedure is perfectly well defined on the lattice, the actual carrying out of the Fourier transform appears technically very difficult. An attempt to define a fully non-Abelian magnetic disorder parameter in the naive continuum limit has been made by Mandelstam.[14] Considerable additional obstacles of definition of composite operators and renormalization must be overcome before a well-defined construction can be given in the continuum limit, especially in the case of pure gauge theory (i.e., no Higgs fields).

IV. THE BEHAVIOR OF $\langle B[C^*]\rangle$ FOR $\beta\to\infty$

Let us trivially rewrite (22) out in full as follows:

$$\langle B[C^*]\rangle = Z^{-1}\int\prod_c d\xi[c]\int\prod_b d\hat{U}[b]\prod_c \delta[f_c^{-1}(\hat{U})\xi[c]]\exp\left[\sum_p \hat{M}_p(\hat{U})\right]\left\{\int\prod_c d\omega[c]\prod_c \chi_{\xi[c]}[\omega[c]]\prod_{c\in C^*}\omega[c]\times\exp\left(\sum_p \hat{K}_p(\hat{U})\omega[\hat{\partial}p]\right)\right\}. \quad (23)$$

We introduce the measure for expectation values in the Z_2 subsystem

$$d\hat{\mu}_{\hat{K}(\hat{U})}\equiv\frac{1}{Z_{\hat{K}(\hat{U})}}\prod_c d\omega[c]\exp\left(\sum_p \hat{K}_p(\hat{U})\omega[\hat{\partial}p]\right), \quad (24a)$$

$$Z_{\hat{K}(\hat{U})}\equiv\int\prod_c d\omega[c]\exp\left(\sum_p \hat{K}_p(\hat{U})\omega[\hat{\partial}p]\right), \quad (24b)$$

$$\langle F\rangle_{\hat{K}(\hat{U})}\equiv\int d\hat{\mu}_{\hat{K}(\hat{U})}F[\omega] \quad (24c)$$

and rewrite (23) as

$$\langle B[C^*]\rangle = Z^{-1}\int\prod_c d\xi[c]\int\prod_b d\hat{U}[b]\prod_c\delta[f_c^{-1}(\hat{U})\xi[c]]\exp\left(\sum_p\hat{M}_p(\hat{U})\right)Z_{\hat{K}(\hat{U})}\left\langle\prod_c\chi_{\xi[c]}[\omega[c]]\prod_{c\in C^*}\omega[c]\right\rangle_{\hat{K}(\hat{U})}. \quad (25)$$

We rewrite Z in (25) in the same way:

$$Z=\int\prod_c d\xi[c]\int\prod_b d\hat{U}[b]\prod_c\delta[f_c^{-1}(\hat{U})\xi[c]]\exp\left(\sum_p\hat{M}_p(\hat{U})\right)Z_{\hat{K}(\hat{U})}\left\langle\prod_c\chi_{\xi[c]}[\omega[c]]\right\rangle_{\hat{K}(\hat{U})}. \quad (26)$$

Note that, for any given ξ configuration in the integrand of (25) or (26), we have

$$\prod_c\chi_{\xi[c]}[\omega[c]]=\prod_{c\in R}\omega[c], \quad (27)$$

where $\{R\}=\{c\,|\,\xi[c]=-1\}$. Also, from (20)

$$\left.\begin{array}{l}\hat{K}_p(\hat{U})\geq 0\\ \hat{K}_p(\hat{U})\geq\hat{K}^0\end{array}\right\}\text{ for all }\hat{U}, \quad (28)$$

where

$$\hat{K}^0\equiv\tfrac{1}{2}\ln\coth 2\beta \quad (29)$$

is the value of $\hat{K}_p(\hat{U})$ with $U[\partial p]$'s at their classical vacuum value. $\hat{K}^0\simeq e^{-4\beta}$ for $\beta\to\infty$.

Using (28) and applying Griffiths inequalities[15] we have

$$\left\langle\prod_{c\in G}\omega[c]\right\rangle_{\hat{K}(\hat{U})}\geq 0, \quad (30)$$

$$\left\langle\prod_{c\in G}\omega[c]\right\rangle_{\hat{K}(\hat{U})}\geq\left\langle\prod_{c\in G}\omega[c]\right\rangle_{\hat{K}^0} \quad (31)$$

for any arbitrary set of cubes G. It follows from (30) that the integrands in (25) and (26) are always nonnegative. Hence, with the notation $\langle\ \rangle_0\equiv Z\langle\ \rangle$, (31) gives the rigorous inequalities

$$\langle B[C^*]\rangle_0\geq\langle B[C^*]\rangle'_0\equiv\int\prod_c d\xi[c]\int\prod_b d\hat{U}[b]\prod_c\delta[f_c^{-1}(\hat{U})\xi[c]]\exp\left(\sum_p\hat{M}_p(\hat{U})\right)Z_{\hat{K}(\hat{U})}\left\langle\prod_c\chi_{\xi[c]}[\omega[c]]\prod_{c\in C^*}\omega[c]\right\rangle_{\hat{K}^0} \quad (32)$$

and

$$Z \geqslant Z' \equiv \int \prod_c d\xi[c] \int \prod_b d\hat{U}[b] \prod_c \delta[f_c^{-1}(\hat{U})\xi[c]] \exp\Big(\sum_p \hat{M}_p(\hat{U})\Big) Z_{\hat{K}(\hat{U})} \Big\langle \prod_c \chi_{\xi[c]}[\omega[c]] \Big\rangle_{\hat{K}^0}. \tag{33}$$

Owing to lack of sufficient control over the $\hat{U}$ integration we cannot rigorously conclude that

$$\langle B[C^*] \rangle \geqslant \langle B[C^*] \rangle',$$

where

$$\langle B[C^*] \rangle' \equiv \langle B[C^*] \rangle_0' / Z', \tag{34}$$

i.e., a sort of generalized Griffiths inequality, although such a bound is quite likely to hold. However, it is sufficient for our purposes that, as we argue below, $\langle B[C^*] \rangle'$ differs from $\langle B[C^*] \rangle$ only by exponential corrections. $\langle B[C^*] \rangle'$ differs from $\langle B[C^*] \rangle$ by omission of the $\hat{U}$ dependence of $\hat{K}_p(\hat{U})$ in $\langle\ \rangle_{\hat{K}(\hat{U})}$ in (25) and (26). If we temporarily accept this omission, and then imagine doing the $\hat{U}$ integration, the result will be a probability distribution for the currents $\xi[c]$. Equation (26) then would have the form of the partition function of a Z_2-gauge theory coupled to a scalar Higgs field—possibly with complicated (nonminimal) interactions—of unit charge; and (25) would be the expectation of a Z_2-Wilson loop for this system.

Now, as $\beta \to \infty$, $\hat{K}_p(\hat{U}) \to \hat{K}^0 \to 0$ for almost all $\hat{U}$ in the integrands in (25) and (26). Hence, we very nearly have the situation just described, but we have to be careful about the small set of exceptional $\hat{U}$'s. As $\beta \to \infty$, the measure of the $\hat{U}$ integration actually suppresses this set, and we will show that indeed the probability that $\hat{K}_p(\hat{U})$ differs from $\hat{K}^0$ on any plaquette vanishes exponentially. This result holds in the presence of any arbitrary operator that depends on the ω's.

Let us write

$$\begin{aligned}\Big\langle \prod_{c\in G} \omega[c] \Big\rangle_{\hat{K}(\hat{U})} &= \frac{\int \prod_c d\omega[c] \prod_{c\in G} \omega[c] \prod_p \exp\{[\hat{K}_p(\hat{U}) - \hat{K}^0]\omega[\hat{\partial}p]\} \exp\Big(\sum_p \hat{K}^0 \omega[\hat{\partial}p]\Big)}{\int \prod_c d\omega[c] \prod_p \exp\{[\hat{K}_p(\hat{U}) - \hat{K}^0]\omega[\hat{\partial}p]\} \exp\Big(\sum_p \hat{K}^0 \omega[\hat{\partial}p]\Big)} \\ &= \frac{\sum_{Q\subset\Lambda} \int \prod_c d\omega[c]\omega[G] \Big\{ \prod_{p\in Q} \omega[\hat{\partial}p] \tanh[\hat{K}_p - \hat{K}^0] \Big\} \exp\Big(\sum_p \hat{K}^0 \omega[\hat{\partial}p]\Big)}{\sum_{Q\subset\Lambda} \int \prod_c d\omega[c] \Big\{ \prod_{p\in Q} \omega[\hat{\partial}_p] \tanh[\hat{K}_p - \hat{K}^0] \Big\} \exp\Big(\sum_p \hat{K}^0 \omega[\hat{\partial}p]\Big)}\,,\end{aligned} \tag{35}$$

where the sum is over all sets of plaquettes Q in Λ, and treat all contributions from $Q \neq \phi$ as "small perturbations." If the contribution of the exceptional $\hat{U}$'s in the $\hat{U}$ integration does not make these perturbations large, then the replacement $\hat{K}_p(\hat{U}) \to \hat{K}^0$ [i.e., keeping only the $Q = \phi$ terms in (35)], in the numerator and denominator (i.e., in Z) in (25), will give the dominant contributions. To estimate the magnitude of the errors involved, let us examine any typical contribution to the corrections, e.g., terms of the form

$$\begin{aligned} I_Q &\equiv \Big[\int \prod_c d\xi[c] \prod_b d\hat{U}[b] \prod_c \delta(f_c^{-1}(\hat{U})\xi[c]) \exp\Big(\sum_p \hat{M}_p(\hat{U})\Big) Z_{\hat{K}(\hat{U})} \prod_{p\in Q} \tanh[\hat{K}_p(\hat{U}) - \hat{K}^0] \\ &\qquad \times \Big\langle \prod_c \chi_{\xi[c]}[\omega]\omega[C^*] \prod_{p\in Q} \omega[\hat{\partial}p] \Big\rangle_{\hat{K}^0} \Big] \\ &\qquad \times \Big[\int \prod_c d\xi[c] \prod_b d\hat{U}[b] \prod_c \delta[f_c^{-1}(\hat{U})\xi[c]] \exp\Big(\sum_p \hat{M}_p(\hat{U})\Big) Z_{\hat{K}(\hat{U})} \Big]^{-1} \\ &\leqslant \frac{\int \prod_b d\hat{U}[b] \prod_c d\omega'[c] \exp\Big(\sum_p \{\hat{M}_p(\hat{U}) + \hat{K}_p(\hat{U})\omega'[\hat{\partial}p]\}\Big) \prod_{p\in Q} \tanh[\hat{K}_p(\hat{U}) - \hat{K}^0]}{\int \prod_b d\hat{U}[b] \prod_c d\omega'[c] \exp\Big(\sum_p \{\hat{M}_p(\hat{U}) + \hat{K}_p(\hat{U})\omega'[\hat{\partial}p]\}\Big)} \\ &\equiv \frac{1}{Z''} \Big(\int \prod_b d\hat{U}[b] \prod_c d\omega'[c] \exp\Big(\sum_p \{\hat{M}_p(\hat{U}) + \hat{K}_p(\hat{U})\omega'[\hat{\partial}p]\}\Big) \prod_{p\in Q} \tanh[\hat{K}_p(\hat{U}) - \hat{K}^0] \Big) \\ &\equiv \Big\langle \prod_{p\in Q} \tanh[\hat{K}_p(\hat{U}) - \hat{K}^0] \Big\rangle''. \end{aligned} \tag{36}$$

Here we used $\langle\omega[G]\rangle_{\hat{K}^0}\le 1$ and the definition (24b), and we defined a new expectation value $\langle\ \rangle''$ as indicated. Now, from (28),

$$\begin{aligned}\tanh[\hat{K}_p(\hat{U})-\hat{K}^0]&=\left\{\frac{1-\exp\{-2[\hat{K}_p(\hat{U})-\hat{K}^0]\}}{1+\exp\{-2[\hat{K}_p(\hat{U})-\hat{K}^0]\}}\right\},\\ &\le(1-\exp\{-2[\hat{K}_p(\hat{U})-\hat{K}^0]\}),\\ &\le\{1-\exp[-2\hat{K}_p(\hat{U})]\}\equiv F[\hat{U}[\partial p]].\end{aligned}\tag{37}$$

Choose any two-dimensional "horizontal plane" of plaquettes P_h, and let $Q_h=Q\cap P_h$. Then, since $F[\hat{U}[\partial p]]\le 1$,

$$I_Q\le\left\langle\prod_{p\in Q_h}F[\hat{U}[\partial p]]\right\rangle''.\tag{38}$$

The measure [see (36)] of this expectation value allows the use of the chessboard-estimates theorems.[12] They give

$$I_Q\le\left\langle\prod_{p\in Q_h}F[\hat{U}[\partial p]]\right\rangle''\le\left(\left\langle\prod_{p\in P_h}F[\hat{U}[\partial p]]\right\rangle''\right)^{|Q_h|/|P_h|}.\tag{39}$$

$|Q_h|$, $|P_h|$ are the number of plaquettes in Q_h and P_h, respectively. Finally, the expectation value $\langle\prod_{p\in P_h}F\rangle''$ is bounded by the usual crude estimate of entropies and free energies.[12] Z'' is bounded from below by restricting the variables $\omega'[c]$ to 1, and the $\hat{U}[b]$ integrations to a region where $|\mathrm{tr}U[\partial p]|\ge 2e^{-\delta}$. The volume of this region is denoted by τ_δ. The numerator in $\langle\ \rangle''$ is bounded from above by replacing the integrand by its maximum. Substituting into (39), and using the fact that we can always choose P_h so that $|Q_h|/|P_h|\ge|Q|/|P|$, one obtains

$$I_Q\le C(\beta)^{|Q|},\tag{40}$$

where

$$C(\beta)=\min_\delta\left(\frac{\text{const}}{\tau_\delta}\exp[-2\beta s(\delta)]\right),\tag{41}$$

with $s(\delta)>0$ for sufficiently small δ. $|Q|$ and $|P|$ are the number of plaquettes in Q and Λ, respectively. Hence $C(\beta)\to 0$ as $\beta\to\infty$. Equation (40) implies that, in the numerator and denominator in (25), the probability that $\hat{K}(\hat{U})$ differs from $\hat{K}^0$ on any set of plaquettes vanishes exponentially as $\beta\to\infty$. It follows that $\langle B[C^*]\rangle$ differs from $\langle B[C^*]\rangle'$ by exponentially small terms.

From now on we will consider $\langle B[C^*]\rangle'$ instead of $\langle B[C^*]\rangle$ for $\beta\to\infty$. As mentioned above it can actually be expected to be a lower bound on $\langle B[C^*]\rangle$, although we have no rigorous proof of this. From (32) and (24b)

$$\langle B[C^*]\rangle_0'=\int\prod_c d\zeta[c]\left\langle\prod_c\chi_{\zeta[c]}[\omega[c]]\prod_{C^*}\omega[c]\right\rangle_{\hat{K}^0}\times\left(\int\prod_b d\hat{U}[b]\prod_c d\omega'[c]\prod_c\delta[f_c^{-1}(\hat{U})\zeta[c]]\exp\left\{\sum_p\{\hat{M}_p(\hat{U})+\hat{K}_p(\hat{U})\omega'[\hat{\partial}p]\}\right\}\right).$$

Performing an inverse dual transform on $\omega'[c]$ we obtain

$$\langle B[C^*]\rangle_0'=\int\prod_c d\zeta[c]\left\langle\prod_c\chi_{\zeta[c]}[\omega[c]]\prod_{C^*}\omega[c]\right\rangle_{\hat{K}^0}\times\left[\int\prod_b d\hat{U}[b]\prod_p d\sigma'[p]\prod_c\delta[f_c^{-1}(\hat{U})\zeta[c]]\prod_c\left(\tfrac{1}{2}\delta\left(\prod_{\partial c}\sigma'[p]\right)\right)\exp\left(\sum_p K_p(\hat{U})\sigma'[p]\right)\right],$$

and solving the constraint on $\sigma'[p]$ by writing $\sigma'[p]=\prod_{b\in\partial p}\gamma[b]$:

$$\langle B[C^*]\rangle_0'=\int\prod_c d\zeta[c]\left\langle\prod_c\chi_{\zeta[c]}[\omega[c]]\prod_{C^*}\omega[c]\right\rangle_{\hat{K}^0}\left[\int\prod_b d\hat{U}[b]\prod_b d\gamma[b]\prod_c\delta[f_c^{-1}(\hat{U})\zeta[c]]\times\exp\left(\sum_p K_p(\hat{U})\gamma[\partial p]\right)\right].\tag{42}$$

Similarly (33) is written as

$$Z'=\int\prod_c d\zeta[c]\left\langle\prod_c\chi_{\zeta[c]}[\omega[c]]\right\rangle_{\hat{K}^0}\left[\int\prod_b d\hat{U}[b]\prod_b d\gamma[b]\prod_c\delta[f_c^{-1}(\hat{U})\zeta[c]]\exp\left(\sum_p K_p(\hat{U})\gamma[\partial p]\right)\right].\tag{43}$$

The expression in large square brackets in (42) and (43) can be viewed as giving the probability distribution $J[\zeta]$ for the ζ's. Our aim below will be to obtain an estimate for this distribution as $\beta\to\infty$. To this end consider, for a given ζ configuration, the set R defined as in (27): $\{R\}=\{c\,|\,\zeta[c]=-1\}$. Choose an (open) sublattice $\mathcal{R}$ containing R as follows. Let Q be the set of plaquettes that have bonds in common with $\partial c, c\in R$, i.e.,

$$Q=\{p\,|\,\partial p\cap\partial p'\ne\phi,\ p'\in\partial c',\ c'\in R\}\tag{44a}$$

Then

$$\mathcal{R}=\{c\,|\,c\in\hat{\partial}p, p\in Q\}.\tag{44b}$$

With $\mathcal{R}_c$ denoting the complement of $\mathcal{R}$, and $\tilde{\mathcal{R}}_c$ the closure of the compliment of $\mathcal{R}$, we then have

$$J[\zeta]\equiv\left[\int\prod_b d\hat U[b]\prod_b d\gamma[b]\prod_c \delta[f_c^{-1}(\hat U)\zeta[c]]\exp\left(\sum_p K_p(\hat U)\gamma[\partial p]\right)\right]$$
$$=\int\prod_{b\in\tilde{\mathfrak{R}}_c} d\hat U[b]\prod_{b\in\tilde{\mathfrak{R}}_c} d\gamma[b]\prod_{c\in\mathfrak{R}_c}\delta[f_c(\hat U)]\prod_{p\in\tilde{\mathfrak{R}}_c}\exp\{K_p(\hat U)\gamma[\partial p]\}\left(\int\prod_{b\notin\tilde{\mathfrak{R}}_c} d\hat U[b]\prod_{b\notin\tilde{\mathfrak{R}}_c} d\gamma[b]\right.$$
$$\left.\times\prod_{c\in\mathfrak{R}}\delta[f_c^{-1}(\hat U)\zeta[c]]\prod_{p\notin\tilde{\mathfrak{R}}_c}\exp\{K_p(\hat U)\gamma[\partial p]\}\right). \quad (45)$$

This can be written in the form

$$J[\zeta]=\int\prod_b d\hat U[b]\prod_b d\gamma[b]\prod_c d\rho[c]\prod_c \delta[\rho[c]^{E_{\mathfrak{R}}[c]}f_c(\hat U)]\exp\left(\sum_p K_p(\hat U)\gamma[\partial p]\right)w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}[\zeta], \quad (46)$$

where

$$w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}\equiv\left[\int\prod_{b\notin\tilde{\mathfrak{R}}_c} d\hat U[b]\prod_{b\notin\tilde{\mathfrak{R}}_c} d\gamma[b]\prod_{\substack{c\in\mathfrak{R}\\ c\notin R}}\delta[f_c(\hat U)]\prod_{c\in R}\delta[-f_c(\hat U)]\prod_{p\notin\tilde{\mathfrak{R}}_c}\exp\{K_p(\hat U)\gamma[\partial p]\}\right]$$
$$\times\left[\int\prod_{b\notin\tilde{\mathfrak{R}}_c} d\hat U[b]\prod_{b\notin\tilde{\mathfrak{R}}_c} d\gamma[b]\prod_{c\in\mathfrak{R}} d\rho[c]\prod_{c\in\mathfrak{R}}\delta[\rho[c]f_c(\hat U)]\prod_{p\notin\tilde{\mathfrak{R}}_c}\exp\{K_p(\hat U)\gamma[\partial p]\}\right]^{-1}. \quad (47)$$

As the subscripts indicate, $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}[\zeta]$ depends on the choice of $\mathfrak{R}$, on the fixed values of $\hat U$ on $\partial\mathfrak{R}$, and on $\zeta[c]$ in $\mathfrak{R}$. Clearly, due to the continuous nature of the group, there is at least some range of boundary conditions $\hat U_{\partial\mathfrak{R}}$ of finite measure for which $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}$ is nonzero, i.e., $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}>0$. For the moment let us assume that this is so, and proceed to bound $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}$.

We may bound the numerator from below by evaluating it for only a limited set of configurations. For given boundary conditions $\hat U_{\partial\mathfrak{R}}$, consider a particular configuration $\{\hat U'[b], b\notin\tilde{\mathfrak{R}}_c\}$ which satisfies the δ functions and hence contributes to nonzero $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}$. Denote the corresponding $K_p(\hat U)$ by $K_p(\hat U',\hat U_{\partial\mathfrak{R}})$ for p in $\mathfrak{R}$. Restrict the $\hat U[b]$ variables to a region in the group such that

$$K_p(\hat U',\hat U_{\partial\mathfrak{R}})\geq K_p(\hat U)\geq K_p(\hat U',\hat U_{\partial\mathfrak{R}})e^{-\delta}.$$

Let τ_δ be the volume of the corresponding group subset when $\hat U[\partial p]$, $p\in\mathfrak{R}$, satisfy this restriction; $\tau_\delta<1$. Then

$$[\text{numerator in (47)}]\geq\int\prod_{b\notin\tilde{\mathfrak{R}}_c} d\gamma[b]\prod_{p\notin\tilde{\mathfrak{R}}_c}(\tau_\delta\exp\{K_p(\hat U',\hat U_{\partial\mathfrak{R}})e^{-\delta}\gamma[\partial p]\}).$$

By our assumption $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}>0$, we may, by choosing sufficiently small δ, take τ_δ to be independent of $\hat U_{\partial\mathfrak{R}},\mathfrak{R}$. Since $\gamma[\partial p]=\pm1$, $0\leq K_p(\hat U)\leq 2\beta$, we can write for the minimum over any possible $\{\hat U'[b]\}$ and all $\hat U_{\partial\mathfrak{R}}$

$$[\text{numerator in (47)}]\geq(\tau_\delta e^{-2\beta})^{r|R|},$$

where r is a dimension-dependent constant, and $|R|$ is the number of cubes in R. Similarly the denominator is bounded from above by replacing the integrand by its maximum:

$$[\text{denominator in (47)}]\leq(e^{2\beta})^{r|R|}.$$

Hence, we obtain the crude bound

$$w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}[\zeta]\geq(\tau_\delta^r e^{-4r\beta})^{|R|} \quad (48)$$

for all $\hat U_{\partial\mathfrak{R}}$. On the other hand $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}$ is bounded from above by

$$w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}[\zeta]\leq 1. \quad (49)$$

$w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}[\zeta]=1$ makes $\zeta[c]=-1$ equally probable to $\zeta[c]=+1$ on any given c. It could be attained if there are zero-action configurations for the production of monopoles as $\beta\to\infty$.

Equations (48) and (49) hold for those $\hat U_{\partial\mathfrak{R}}$ that give a nonzero $w_{\mathfrak{R},\hat U_{\partial\mathfrak{R}}}$. When we substitute in (46) we have to worry about the coupling of this condition on $\hat U_{\partial\mathfrak{R}}$ to the "outside" integrations. Ignoring this coupling leads to a pure length behavior for the monopole distribution (46). The possibility of spreading the flux in a pure gauge theory means that actually the effect of this coupling results in at most power corrections to such length behavior. This is of course already true in a massless perturbative expansion. The nonperturbative excitations of the $\hat U$'s can only improve the situation since they should be responsible for the expected mass gap of the theory. However, about this last point we have nothing to say in this paper. As will be evident in the following, it is sufficient for our purposes that, to within possible power corrections, (46) gives a length-law probability distribution for the currents $\zeta[c]$. We will, therefore, utilize (48) and (49) in (46) assuming that they hold for almost all $\hat U_{\partial\mathfrak{R}}$.

(48) and (49) give bounds on the probability distribution of the ζ's that can be represented by

$$z'(\beta)\prod_h\{\tfrac{1}{2}\delta[\xi[\partial h]]\}\geq J[\xi]$$
$$\geq z(\beta)\prod_h\{\tfrac{1}{2}\delta[\xi[\partial h]]\}\prod_c\frac{(e^{D(\beta)}+\xi[c]e^{-D(\beta)})}{(e^{D(\beta)}+e^{-D(\beta)})}. \tag{50}$$

Here we defined

$$z(\beta)\equiv\int\prod_b d\hat{U}[b]\prod_b d\gamma[b]\prod_c\delta[f_c(\hat{U})]\exp\Big(\sum_p K_p(\hat{U})\gamma[\partial p]\Big), \tag{51}$$

$$(\tau_6^r e^{-4r\beta})=\text{const}\times e^{-\text{const}\times\beta}\equiv\tanh D(\beta) \tag{52}$$

and $z'(\beta)$ is given by the same expression as $z(\beta)$ but with the δ functions omitted. Note the δ functions that enforce the co-closure (at $d=4$) of any ξ configuration, which was automatic in (46) and (47) because of the definition of $f_c(\hat{U})$ [see (12)].

We will choose that form of probability distribution which, after integrating over all $\xi[c]$, will result into the lowest value for $\langle B[C^*]\rangle'$. This form is the lower bound on $J[\xi]$ given in (50). This is indeed a consequence of the Griffiths inequalities[15] for the effective Higgs-Z_2 gauge system resulting from replacing $J[\xi]$ by the bounding distributions in (50) (see below); lowest $\langle B[C^*]\rangle'$ is given by the lowest possible Higgs modulus value. In terms of the physics of the problem, $D(\beta)\to 0$ as $\beta\to\infty$ means that monopole production is suppressed, and hence screening of an external magnetic source is also suppressed. Indeed if we were to take $D=0$ exactly, then all monopoles would be rigorously excluded, and all screening eliminated (the MP model). On the other hand, taking $D\to\infty$ corresponds to the upper bound in (50), where monopoles are freely produced, and screening becomes very strong: $\langle B[C^*]\rangle'\to 1$.

From (42) and (50) we have

$$\langle B[C^*]\rangle_0'\geq z(\beta)\left(\frac{1}{(e^{D(\beta)}+e^{-D(\beta)})}\right)^{|C|}\int\prod_c d\xi[c]\prod_h\{\tfrac{1}{2}\delta[\xi[\partial h]]\}\prod_c(e^{D(\beta)}+\xi[c]e^{-D(\beta)})$$
$$\times\Big\langle\prod_c\chi_{\xi[c]}[\omega[c]]\omega[C^*]\Big\rangle_{\hat{K}^0}, \tag{53}$$

where $|C|$ is the number of cubes in Λ. We now perform a dual transformation on $\xi[c]$. Using the Fourier transforms

$$e^{D(\beta)}+\xi[c]e^{-D(\beta)}=2\int d\xi[c]\,e^{D[\beta]\xi[c]}\chi_{\xi[c]}[\zeta[c]], \tag{54}$$

$$\tfrac{1}{2}\delta[\zeta[\partial h]]=\int\prod_h d\phi[h]\chi_{\phi[h]}[\zeta[\partial h]], \tag{55}$$

we have

$$\int\prod_c d\xi[c]\prod_h\{\tfrac{1}{2}\delta[\xi[\partial h]]\}\prod_c\{e^{D(\beta)}+\xi[c]e^{-D(\beta)}\}\times\prod_c\chi_{\xi[c]}[\omega[c]]$$
$$=\int\prod_h d\phi[h]\times\prod_c d\xi[c]\prod_c e^{D(\beta)\xi[c]}\times\prod_c\delta[\phi[\hat{\partial}c]\omega[c]\xi[c]]$$
$$=\int\prod_h d\phi[h]\prod_c e^{D(\beta)\phi[\hat{\partial}c]\omega[c]}. \tag{56}$$

This represents a Higgs field $\phi[h]$ of fixed modulus $D(\beta)^{1/2}$ minimally coupled to $\omega[c]$. Substituting (56) in (53) we obtain

$$\langle B[C^*]\rangle_0'\geq z(\beta)\left[\frac{1}{(e^{D(\beta)}+e^{-D(\beta)})}\right]^{|C|}\frac{1}{Z_{\hat{K}^0}}\times\int\prod_h d\phi[h]\prod_c d\omega[c]\prod_{c\in C^*}\omega[c]\times\exp\Big\{D(\beta)\sum_c\phi[\hat{\partial}c]\omega[c]+\hat{K}^0\sum_p\omega[\hat{\partial}p]\Big\}. \tag{57}$$

In completely analogous fashion we obtain for Z' from (43)

$$Z'\geq z(\beta)\left[\frac{1}{(e^{D(\beta)}+e^{-D(\beta)})}\right]^{|C|}\frac{1}{Z_{\hat{K}^0}}\times\Big\{\int\prod_h d\phi[h]\prod_c d\omega[c]\times\exp\Big(D[\beta]\sum_c\phi[\hat{\partial}c]\omega[c]+\hat{K}^0\sum_p\omega[\hat{\partial}p]\Big)\Big\}. \tag{58}$$

The expression in curly brackets is precisely the partition function $Z_{\hat{K}^0,D}$ of the Higgs-Z_2 gauge system

$$Z_{\hat{K}^0,D} \equiv \int \prod_h d\phi[h] \prod_c d\omega[c] \times \exp\left(D(\beta) \sum_c \phi[\hat{\partial} c]\,\omega[c] + \hat{K}^0 \sum_p \omega[\hat{\partial} p] \right). \quad (59)$$

For $\langle B[C^*]\rangle'$ then, as explained above, we have

$$\langle B[C^*]\rangle' \geq \frac{1}{Z_{\hat{K}^0,D}} \left\{ \int \prod_h d\phi[h] \prod_c d\omega[c] \prod_{c\in C^*} \omega[c] \times \exp\left(D(\beta) \sum_c \phi[\hat{\partial} c]\,\omega[c] + \hat{K}^0(\beta) \sum_p \omega[\hat{\partial} p] \right) \right\}. \quad (60)$$

This is the expression for a Wilson loop C^* in the Higgs-Z_2 gauge system. $D(\beta)$ and $\hat{K}^0(\beta)$ both tend to zero as $\beta \to \infty$, and hence we can evaluate (60) in a high-temperature expansion. One finds

$$\langle B[C^*]\rangle' \geq (\tanh D(\beta))^{|C^*|} + \cdots + (\tanh \hat{K}^0(\beta))^{|S^*|} + \cdots . \quad (61)$$

$|C^*|$ is the number of cubes in C^*, and $|S^*|$ the number of plaquettes in the minimal area whose coboundary is C^*, i.e., $\hat{\delta} S^* = C^*$. The ellipses indicate the presence of other terms besides those written out explicitly. The $|C^*|$ term is the perimeter-law term, the $|S^*|$ term is the area-law term. For asymptotically large loops the perimeter-decaying term always dominates. Thus the long-distance behavior is given by perimeter law

$$\langle B[C^*]\rangle' \geq \sim \exp[-\alpha(\beta)|C^*|], \quad (62)$$

with

$$\alpha(\beta) = -\ln \tanh D(\beta). \quad (63)$$

Then, using (52)

$$\alpha \simeq 4r\beta, \quad \beta \to \infty. \quad (64)$$

Physically, the perimeter-law dependence describes the fact that a pair of dynamical monopoles can always pop out of the vacuum and shield any external source.[16] This happens for any finite value of $D(\beta)$, the fixed length of the Higgs field describing the monopoles. Indeed, perimeter dependence for any nonzero D is a consequence of the Griffiths inequalities.[15] For $D = 0$ (exclusion of monopoles) we have a sudden crossover to area law.

This situation is a special case of the general structure of the phase diagram of lattice gauge theories coupled to (fixed-length) Higgs fields, which has been studied[16,17] both for discrete and continuous gauge groups [Z_N, U(1), SU(N)]. This phase diagram depends crucially on whether the Higgs field is in the fundamental representation of the gauge group or not. In our case the "Higgs field" is necessarily in the fundamental representation since the Z_2 group has only one nontrivial representation. If we temporarily ignore the β dependence of D and $\hat{K}^0$ and treat them as free, independent parameters, the phase diagram of the effective Higgs-Z_2 gauge system in (60), taken from the results of Fradkin and Shenker, appears in Fig. 1. The main feature is the shaded analyticity region which continuously connects the "confined" to the "Higgs" regime. Note that it has finite width at both ends. It is the region called the "total screening phase" by Banks and Rabinovici. Now, in (60) D and $\hat{K}^0$ are functions of β given by (29) and (52) for $\beta \to \infty$, i.e., our region of interest is the lower left-hand corner in Fig. 1. It is described by Eqs. (61)–(64) of the high-temperature expansion. The upper left-hand corner of Fig. 1, i.e., $\hat{K}^0$ small, $D \to \infty$, corresponds to the effective Higgs-Z_2 gauge theory we would have obtained, had we used the *upper* bound in (50) as our limiting distribution for $J[\zeta]$. In this case, following the steps analogous to (53)–(60), we obtain

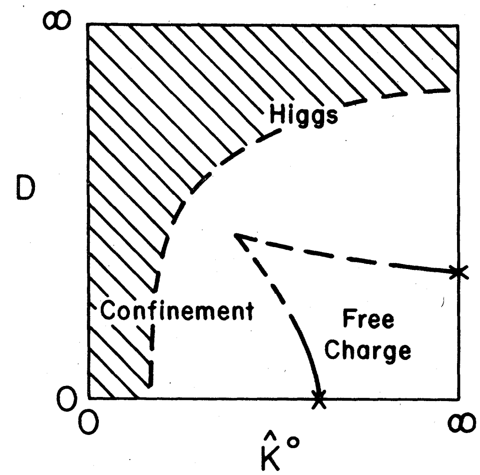

FIG. 1. Phase diagram of the effective Higgs-Z_2 gauge system [Eq. (59)]. The shaded region represents the analyticity region which continuously (i.e., without any phase boundary) connects the "Higgs" ($D, \hat{K}^0$ large) and "confinement" ($D, \hat{K}^0$ small) regimes. Note the finite width of the region at both ends ($\hat{K}^0 = \infty$, and $D = 0$). The figure is taken from the paper by Fradkin and Shenker (Ref. 16).

$$\langle B[C^*]\rangle' \le \lim_{D\to\infty} \frac{1}{Z_{\hat{K}^0,D}} \left\{ \int \prod_h d\phi[h] \prod_c d\omega[c] \prod_{c\in C^*} \omega[c] \exp\left[D \sum_c \phi[\hat{\partial} c]\omega[c] + \hat{K}^0(\beta) \sum_p \omega[\hat{\partial} p]\right]\right\}. \qquad (65)$$

This upper bound describes the extreme situation in which the effective mass of the dynamical monopoles goes to zero, and screening becomes arbitrarily strong: $\langle B[C^*]\rangle' \to 1$. As noted before, this could occur if there exist zero-action configurations for the creation of the dynamical monopoles.

What the phase diagram of Fig. 1 shows is that, for small $\hat{K}^0$, we have perimeter law for all $D>0$. In other words, in our region of interest, $\beta\to\infty$, $\hat{K}^0 = \hat{K}^0(\beta) \to 0$, *any* distribution we may choose for representing $J[\xi]$ between the two limits in (50) will result in a perimeter falloff. In this connection note that the coefficient of β in (63) and (64) is rather large, so that $\langle B[C^*]\rangle' \to 0$ rapidly for $\beta\to\infty$. This reflects the crudeness of the lower bound in (50). However, we want to argue that this behavior is representative for all possible distributions that do not approach the upper bound in (50) arbitrarily close, albeit with a possibly much smaller coefficient in front of β in (63) and (64). To see this go back to the exact definition of $J[\xi]$ in (43), and ask for the probability that $K_p(\hat{U}) < (2-\delta)$, $\delta>0$, as $\beta\to\infty$ in the measure defined by Z'. By simple chessboard estimate this can be shown to go to zero as $e^{-a(\delta)\beta}$, $a(\delta)>0$. This means that if we are to have $f_c(\hat{U})$ equal to -1 on one cube, and $+1$ on a neighboring cube, we will end up with the behavior (62)–(64)—unless this can be accomplished by having at the same time $K_p(\hat{U}) = 2$ everywhere. The latter possibility requires the existence of zero-action configurations producing monopoles. It is amusing to observe that such zero-action configurations producing Z_2-flux can be found[18] in the case of the magnetic-flux free energy, also introduced by 't Hooft.[19] (The magnetic-flux free energy introduces topologically stable sourceless flux into the lattice by imposing twisted periodic boundary conditions.) However, in the case of the 't Hooft loop and the monopoles needed to cancel its flux, one would always expect some finite "end effects," and precisely zero-action configurations appear unlikely.

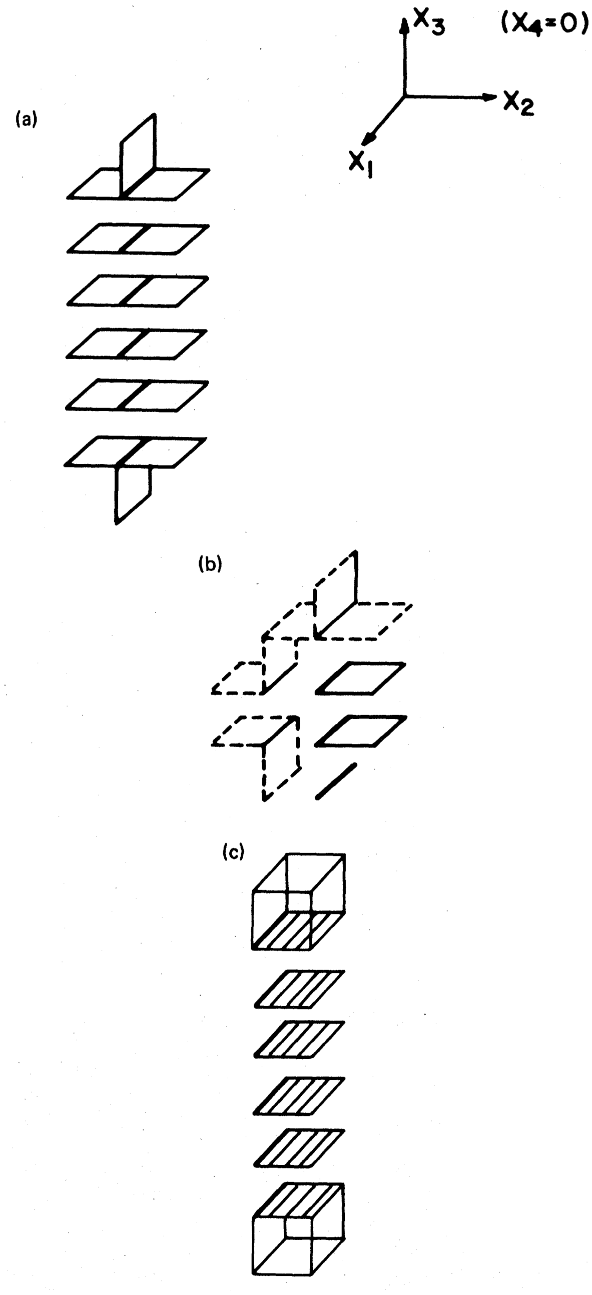

FIG. 2. (a) The set of bonds L (heavy lines) and the plaquettes in its coboundary $\hat{\partial}L$. The shaded plaquettes are the set $Q \subset \hat{\partial}L$. $d=3$, or three-dimensional section at $d=4$. (b) Flux spreading. By adjusting appropriate bonds (light lines) the flux on certain plaquettes is nearly canceled, and the remaining flux spills over onto further neighboring plaquettes (broken lines). (c) Monopole-antimonopole pair (cubes) connected by a line of flux on plaquettes Q.

V. DISCUSSION

We have investigated the behavior of the expectation of the 't Hooft operator as $\beta\to\infty$ in the standard SU(2) lattice gauge theory. We argued that $\langle B[C^*]\rangle$ exhibits perimeter-law behavior as a result of the presence of dynamical monopoles of Z_2 flux. These configurations, which can screen an external Z_2-magnetic source, express the possibility of nonconservation of Z_2 flux in a SU(2) theory. They are a consequence of the continuous nature of the group that allows flux spreading. Consider a region of the lattice where initially we have $U[b]=1$ for all the bonds. Choose a set of neighboring bonds L as in Fig. 2(a), and set $U[b]=-1$ if $b\in L$. This will induce flux $\tau=-1$ on any plaquette $p\in\hat{\partial}L$. Let Q be all $p\in\hat{\partial}L$ which protrude one unit in a given direction, say the x^2 direction in Fig. 2(a). Perturb the bonds on the boundaries of all other plaquettes in $\hat{\partial}L$, i.e., $b\in\partial p$, $p\in\hat{\partial}L\setminus Q$, so that the flux on these plaquettes is approximately canceled, i.e., $U[\partial p]\approx\tau^{\delta}$ is some SU(2) element very close to 1. This is clearly possible due to the continuous nature of the group. A flux $\sim\tau^{(-1+\delta)}$ now spills over on further neighboring plaquettes. By continuing the process [Fig. 2(b)] over a sufficiently large volume, one can clearly arrange things so that one ends up with $U[\partial p]\approx 1$ everywhere except on Q where $U[\partial p]=-1$. We then have the situation depicted in Fig. 2(c): a monopole-antimonopole pair connected by a line of Z_2 flux on $p\in Q$. This is precisely the type of configuration needed to cancel the flux of a 't Hooft loop. In the presence of an external source like $B[C^*]$ one would have flux approximately equal to one on *all* plaquettes, so that cancellation can be obtained for arbitrarily small coupling. Note in this connection that, precisely because of the continuity of the non-Abelian group, an infinity of individual "neighboring" configurations of $U[b]$'s can actually describe a certain "monopole on a given cube." In a Z_2 gauge theory, and in the MP model where Z_2-flux conservation is enforced as a constraint, such configurations cannot occur, and area law for $\langle B[C^*]\rangle$ follows.

However, in the MP model, as opposed to a Z_2 gauge theory, one can still consider the spreading of conserved flux configurations (closed sheets at $d=4$, lines at $d=3$) by transverse flux spreading. This could result in the disordering of the Wilson loop for arbitrarily weak coupling.[7] Therefore, one can expect area law for the Wilson loop in the MP model just like in the standard model. Simultaneous area law for both the Wilson and 't Hooft loops is, of course, allowed by the 't Hooft commutation relations. A concise, lucid discussion of the expected behavior of the various observables as a consequence of the flux spreading possible in an SU(N) theory has been given by Yaffe.[8]

Perimeter law for the 't Hooft loop is then not necessary for quark confinement. It is sufficient? In view of the fact that the same process of flux spreading allowing configurations that could disorder the Wilson loop for arbitrarily small coupling is also responsible for monopoles screening the 't Hooft loop, an affirmative answer is likely. Unfortunately, an actual proof of this is not easy.

A possible framework for approaching this question has been provided by Mack and Petkova.[20] They derive a simple, "kinematical" bound on the Wilson loop based on the behavior of "thick vortex containers" winding around the loop. These containers are sublattices with fixed boundary conditions and external Z_2 flux running through them. The bound provides a precise formulation of the folklore idea of "vortices" disordering a loop. It shows that if, for sufficiently thick containers, the external flux can spread and the free energy of the container becomes nearly independent of it, then area law for the Wilson loop follows. The hard dynamical question is to actually prove such behavior for the containers.[21] Now, the containers can be represented as 't Hooft loops in the limit where the loop length goes to zero while its flux is constrained to lie in nonsimply connected, finite sublattices. Showing perimeter-law behavior for such 't Hooft loops would be sufficient for obtaining confinement. This may not be easy to do rigorously since dealing with the effects of the sublattice boundary conditions amounts to eventually facing up to the issue of the existence of a mass gap. However, the results of this paper can be used to provide, if not a rigorous, at least a quite plausible argument for confinement as $\beta\to\infty$. This will be the subject of a future paper.

ACKNOWLEDGMENTS

I would like to thank A. Ukawa, L. Yaffe, P. Windey, and D. Gross for many discussions on lattice gauge theories. This research was partially supported by the National Science Foundation under Grant No. PHY78-01221.

[1]K. G. Wilson, Phys. Rev. D 10, 2445 (1974).

[2]J. B. Kogut and L. Susskind, Phys. Rev. D 11, 395 (1975); R. Balian, J. M. Drouffe, and C. Itzykson, *ibid.* 10, 3376 (1974); K. Osterwalder and E. Seiler, Ann. Phys. (N. Y.) 110, 440 (1978). For a review of lattice gauge theories, see J. M. Drouffe and C. Itzykson, Phys. Rep. 38C, 133 (1978); J. B. Kogut, Rev. Mod. Phys. 51, 659 (1979).

[3]M. Creutz, Phys. Rev. D 21, 2308 (1980); K. G. Wilson, Cornell University report, 1980 (unpublished); J. B.

Kogut, R. B. Pearson, and J. Shigemitsu, Phys. Rev. Lett. 43, 484 (1979); J. B. Kogut and J. Shigemitsu, *ibid.* 45, 410 (1980).

[4]C. G. Callan, R. Dashen, and D. J. Gross, Phys. Rev. Lett. 44, 435 (1980).

[5]F. G. Wegner, J. Math. Phys. 12, 2259 (1971).

[6]G. 't Hooft, Nucl. Phys. B138, 1 (1978).

[7]G. Mack and V. B. Petkova, Ann. Phys. (N. Y.) 123, 442 (1979).

[8]L. G. Yaffe, Phys. Rev. D 21, 1574 (1980).

[9]G. Mack and V. B. Petkova, DESY report, 1979 (unpublished).

[10]A. Ukawa, P. Windey, and A. H. Guth, Phys. Rev. D 21, 1013 (1980).

[11]The simple topological concepts and notations we use are standard, and have been summarized in several papers. See, e.g., Refs. 7, 8, 10, and A. H. Guth, Phys. Rev. D 21, 2291 (1980). For complete mathematical treatments, see P. S. Aleksandrov, *Combinational Topology* (Graylock, New York, 1956); J. G. Hocking and G. S. Young, *Topology* (Addison-Wesley, Reading, Mass. 1961).

[12]J. Fröhlich, R. Israel, E. H. Lieb, and B. Simon, Commun. Math. Phys. 62, 1 (1978); J. Stat. Phys. 22, 297 (1980); B. Simon, in *Statphys*. 13, Proceedings of 13th IUPAP Conference on Statistical Physics, Haifa, 1977 (Annals of the Israel Physical Society, Haifa, 1978, Vol. 2, Chap. 22). For gauge theories, D. Brydges, J. Fröhlich, and E. Seiler, Ann. Phys. (N. Y.) 121, 217 (1979); also G. Mack and V. B. Petkova, Ref. 7; L. G. Yaffe, Ref. 8 (Appendix A).

[13]R. Balian, J. M. Drouffe, and C. Itzykson, Phys. Rev. D 11, 2098 (1975); C. P. Korthals Altes, Nucl. Phys. B142, 315 (1978); T. Yoneya, *ibid.* B144, 195 (1978).

[14]S. Mandelstam, Phys. Rev. D 19, 2391 (1979).

[15]R. B. Griffiths, J. Math. Phys. 8, 478 (1967); 8, 484 (1967); D. G. Kelly and S. Sherman, *ibid.* 9, 466 (1968); J. Ginibre, Commun. Math. Phys. 16, 310 (1970); G. F. De Angelis, D. de Falco, and F. Guerra, *ibid.* 57, 201 (1977).

[16]E. Fradkin and S. H. Shenker, Phys. Rev. D 19, 3682 (1979); T. Banks and E. Rabinovici, Nucl. Phys. B160, 349 (1979).

[17]K. Osterwalder and E. Seiler, in Ref. 2; G. F. De Angelis, D. de Falco, and F. Guerra, in Ref. 15.

[18]G. Mack, private communication from D. Gross.

[19]G. 't Hooft, Nucl. Phys. B153, 141 (1979).

[20]G. Mack and V. B. Petkova, Ann. Phys. (N. Y.) 125, 117 (1980).

[21]One may try to approach this question by means of some approximation scheme. For sufficiently thin containers semiclassical techniques should be valid. One can derive an effective Z_2 gauge theory representing containers with magnetic flux spread in their interior. The coupling is renormalized as a function of the container thickness. By including appropriate entropy effects it is possible to obtain "condensation" at some critical distance. A calculation for the string tension of a quark pair in this effective theory produces a rather sharp crossover from a typical "strong coupling" logarithmic behavior to a "weak coupling behavior," where the tension falls very rapidly (essentially exponentially) to zero (E. Tomboulis, unpublished). It is somewhat remarkable that such basically simple calculations can produce this picture. A quite similar approximation attempt, also involving proposals for a Monte Carlo calculation of container free energies, is advocated by G. Mack, Phys. Rev. Lett. 45, 1378 (1980).