PHYSICAL REVIEW D VOLUME 22, NUMBER 10 15 NOVEMBER 1980

Monte Carlo calculations and a model of the phase structure for gauge theories on discrete subgroups of SU(2)

D. Petcher and D. H. Weingarten
Physics Department, Indiana University, Bloomington, Indiana 47405
(Received 27 May 1980)

A Monte Carlo technique is used to evaluate path integrals for lattice gauge theories with gauge groups given by the SU(2) images of the rotational symmetries of a tetrahedron, a cube, and an icosahedron, called $\overline{T}$, $\overline{O}$, and $\overline{I}$, respectively. The coupling-constant dependence of the mean action per plaquette provides evidence for two phases, in each of these theories, separated by a first-order phase transition. The critical gauge coupling constant moves toward zero as the order of the group is increased. The mean plaquette action, the expectations of square gauge loops, and an estimate of the string tension for the largest group $\overline{I}$ agree with Creutz's results for SU(2) over a wide range of coupling constants. A model for the phase transitions in $\overline{T}$, $\overline{O}$, $\overline{I}$, and Z_N is discussed which predicts critical coupling constants close to the observed values and suggests, as expected, that these transitions are a special property of gauge theories over discrete groups.

I. INTRODUCTION

The lattice formulation of quantum chromodynamics (QCD) (Ref. 1) provides a definition of the theory's vacuum expectation values which is free of the inherent ambiguities of a definition using perturbation theory. The main problems of lattice QCD are to determine whether the vacuum expectation values have limits as the lattice spacing goes to zero and the box volume goes to infinity, and, if limits do exist, to find a means of calculating their values for moderate and small momenta. For a simplified version of QCD with quarks removed and the gauge group SU(3) replaced by SU(2), some progress toward a resolution of these problems has been reported recently by Creutz using a Monte Carlo calculation,[2] and by Wilson, using a Monte Carlo formulation of the renormalization-group method.[3]

In the present article we give Monte Carlo calculations for lattice gauge theories simplified still further by replacing SU(2) by one of its non-Abelian discrete subgroups. The subgroups of SU(2) which we consider are those which are homomorphic to the rotational symmetries of a tetrahedron, a cube, and an icosahedron. These groups will be called $\overline{T}$, $\overline{O}$, and $\overline{I}$, respectively.[4] The group of an octahedron is the same as that of a cube and the dodecahedron's group is the same as that of an icosahedron. $\overline{I}$ is the best approximation to SU(2) which can be obtained with a discrete group.

Our motivation for undertaking this project was largely curiosity. An unexpected result which may be of some practical significance, however, is that, although calculations for $\overline{T}$, $\overline{O}$, and $\overline{I}$ can be done more rapidly than for SU(2), the results we obtain agree with those for SU(2) surprisingly well. In particular, our results for $\overline{I}$ agree with Creutz's for SU(2) over a wide range of bare coupling g_0 extending from the region in which the strong-coupling expansion is reliable well into the region in which the weak-coupling expansion works. Thus to study the nature of the transition between the two regions it may be more efficient to use $\overline{I}$ rather than SU(2) itself. On the other hand, the agreement between our results and those of Creutz tends to confirm both.

We begin in Sec. II by discussing $\overline{T}$, $\overline{O}$, and $\overline{I}$ in slightly greater detail and reviewing the standard definition of a lattice gauge theory for any of these groups. Then in Sec. III we present our Monte Carlo calculations of various vacuum expectation values. We work primarily on a periodic lattice with dimensions $8\times8\times8\times8$. The main quantities we evaluate are the expectation values of square gauge loops with sides of length 1 for $\overline{T}$ and $\overline{O}$, and lengths 2, 3, 4, and 5 for $\overline{I}$ in addition. From the results for $\overline{I}$ we extract the string tension.

The essential difference between the results we obtain and those for SU(2) itself[2] is that, while SU(2) shows no sign of a phase transition, and apparently confines static quarks for all values of the bare coupling, the discrete groups do have a phase transition. The phase transition seems to be first order. For values of $\beta=4g_0^{-2}$ below the critical point β_c, a phase occurs which confines static quarks and yields expectation values agreeing with SU(2). For β above β_c, on the other hand, static quarks are not confined and the results we obtain disagree with those of SU(2). The value of β_c rises as the group becomes larger.

It is natural to wonder, of course, whether the transition found in the discrete groups might not also be present in SU(2) but somehow have been missed. If such a transition did occur in SU(2),

 © 1980 The American Physical Society

then since the theory's continuum limit has large β, the continuum limit of SU(2) would not confine static quarks. In Sec. IV, therefore, we consider a model of this phase transition which tends to suggest the transition is a special property of discrete gauge groups as expected. The model yields the prediction

$$\beta_c = \frac{\ln(1+\sqrt{2})}{1-\cos(2\pi/n)}, \tag{1.1}$$

with $n=6$ for $\bar{T}$, 8 for $\bar{O}$, and 10 for $\bar{I}$. This model also predicts a strong-coupling phase transition for Z_N gauge theories with $n=N$. Our predictions for β_c are all in reasonable agreement with the numerical results in Sec. III and in Ref. 5 for Z_N.

II. LATTICE GAUGE THEORIES FOR $\bar{T}$, $\bar{O}$, AND $\bar{I}$

$\bar{T}$ is the subgroup of SU(2) homomorphic to T, the subgroup of SO(3) mapping a tetrahedron centered on the origin onto itself. For each element of T there are two in $\bar{T}$ differing by a minus sign. Since the orientation of a tetrahedron centered on the origin can be given by specifying which of its four corners occupies a standard position and which of the three edges meeting at that corner occupies another standard position, it follows that T has 12 elements and $\bar{T}$ has 24. Similarly, a cube has 8 corners and 3 edges at each corner; thus the subgroup O of SO(3) preserving a cube has 24 elements and its image $\bar{O}$ in SU(2) has 48. An icosahedron has 12 corners and 5 edges meeting at each. Thus, I in SO(3) preserving an icosahedron has 60 elements and $\bar{I}$ in SU(2) has 120.

A lattice gauge theory for a discrete subgroup G of SU(2) on a periodic four-dimensional hypercubic lattice Λ can be constructed in the usual way.[1] To each ordered pair of nearest-neighbor lattice sites $[x,y]$ assign an element $U(x,y)$ of G with $U(x,y)=U(y,x)^\dagger$. For each oriented closed loop of nearest-neighbor sites $l=[x_1, x_2, \ldots, x_n]$, assign a loop variable

$$W(l)=\tfrac{1}{2}\,\mathrm{Tr}\,U(x_n,x_{n-1})\cdots U(x_2,x_1)U(x_1,x_n). \tag{2.1}$$

Loops differing by a cyclic permutation are identified, but for convenience later on we distinguish those differing by a reversal of order. Loops through four sites will be called plaquettes and written p. Following Creutz's convention, the action S for the theory is taken to be

$$S=\sum S(p), \qquad S(p)=1-W(p), \tag{2.2}$$

where the sum includes each plaquette on Λ in an arbitrarily chosen orientation. For any function $\mathfrak{F}$ of the link variables $U(x,y)$, the vacuum expectation value is

$$\langle \mathfrak{F} \rangle = Q^{-1}\sum \mathfrak{F}\exp(-\beta S), \tag{2.3}$$

where the sum is over all ways of choosing each link variable $U(x,y)$ as an element of G, Q is the partition function defined by the condition $\langle 1\rangle = 1$, and β is $4g_0^{-2}$ with g_0 the bare gauge coupling constant.

III. MONTE CARLO

It is impractical to evaluate the sum in (2.3) directly even for small groups and small lattices. For example, a periodic lattice of size $2\times2\times2\times2$ has $4\times2^4=64$ links. Thus if $G=\bar{T}$, the sum in (2.3) has $24^{64}\simeq 2\times10^{88}$ terms. An alternative (and equally impractical) way to find $\langle\mathfrak{F}\rangle$ is to average $\mathfrak{F}$ over an infinite ensemble E of field configurations with each possible configuration occurring a number of times proportional to $Q^{-1}\exp(-\beta S)$. The Monte Carlo evaluation of $\langle\mathfrak{F}\rangle$ consists of averaging $\mathfrak{F}$ over a finite and tractable approximation to E.

The finite ensembles we use are generated by an algorithm suggested by Wilson.[3] To describe the algorithm it is convenient first to define the set $C(G)$ for each of our discrete groups G to be the elements of G, other than the identity, which are closest to the identity. Thus $C(\bar{T})$ has eight members consisting of the SU(2) images of rotations by $\pm2\pi/3$ about an axis through one vertex of the tetrahedron and the center of the opposite triangle. $C(\bar{O})$ has six elements consisting of rotations by $\pm\pi/2$ about an axis through the centers of opposite squares. $C(\bar{I})$ has 12 elements consisting of rotations by $\pm2\pi/5$ about an axis through a pair of opposite vertices.

The finite ensembles we use to evaluate averages start with a single conveniently chosen set of link variables $U(x,y)$. Starting configurations are described later in more detail. Then each link variable on the lattice is successively examined and either changed or left unchanged according to the following procedure: At each link $[x,y]$ one of the elements V of $C(G)$ is chosen at random. The variable $U(x,y)$ is then multiplied on the left by V giving a new $U(x,y)'$. The difference between the action S with $U(x,y)$ in place and S' with $U(x,y)'$ in place is then found. If $S-S'$ is 0 or positive, $U(x,y)$ is replaced by $U(x,y)'$. If $S-S'$ is negative, we evaluate $q=\exp[\beta(S-S')]$ and choose a random number r between 0 and 1. If $r\leq q$, $U(x,y)$ is again replaced by $U(x,y)'$. If $r>q$, $U(x,y)$ is kept.

In this way a sequence of field configurations is

generated, one for each completed sweep of the lattice. Configurations with large values of $\exp(-\beta S)$ are clearly favored over those with $\exp(-\beta S)$ small. It can be shown[3] that as the ensemble of lattices is progressively increased by the method described the ensemble average of any $\mathfrak{F}$ converges to $\langle \mathfrak{F} \rangle$ no matter what starting configuration has been chosen. The finite ensemble we used to evaluate a particular $\langle \mathfrak{F} \rangle$ was typically considered large enough once doubling or quadrupling the ensemble's size left $\langle \mathfrak{F} \rangle$ sufficiently unchanged. The scatter of values of $\langle \mathfrak{F} \rangle$ calculated on subensembles of a quarter of the final set of configurations was then taken as a measure of the uncertainty in $\langle \mathfrak{F} \rangle$. More sophisticated methods can be used to ensure the accuracy of the final result and estimate its errors,[3] but we did not feel that great numerical accuracy was of sufficient importance in our calculations to justify the extra machine time which would have been required.

There are, of course, an infinite number of other algorithms which could be used to generate a sequence of configurations with the property that the running average of $\mathfrak{F}$ approaches $\langle \mathfrak{F} \rangle$ as the number of configurations grows. One possibility would be to choose a trial $U(x,y)'$ on each link randomly over all of G rather than by multiplying the old $U(x,y)$ by a random element of $C(G)$. The disadvantage of this procedure is that for moderate and large β once the iteration has run for a while most of the trial $U(x,y)$ generated will turn out to cause a large increase in S and will therefore be rejected. The algorithm we use rejects fewer trial $U(x,y)'$. As a result the algorithm requires less machine time to generate an ensemble which yields a satisfactory approximation to $\langle \mathfrak{F} \rangle$. Another possible algorithm is Cruetz's heat-bath method[2] in which the $U(x,y)$ on lattice links are successively replaced by new $U(x,y)'$ chosen randomly according to the probability distribution $Q^{-1}\exp(-\beta S)$. This procedure requires more machine time on each link but fewer sweeps of the entire lattice to give an adequate approximation to $\langle \mathfrak{F} \rangle$. On the whole, it is not clear whether this procedure or the one we adopted is faster. The programming for the method we have followed is simpler, however, and more easily adjusted when G is changed.

The first runs we carried out were calculations of the mean action per plaquette $\langle S(p) \rangle$ as a function of β. Our purpose was primarily to locate the phase transitions expected as a result of the arguments to be given in Sec. IV. We began with $\beta = 0$ and $U(x,y)$ for each link chosen randomly over G. The value of $\langle W(p) \rangle$ averaged over all plaquettes for an ensemble of such states is 0 and therefore $\langle S(p) \rangle = \langle 1 - W(p) \rangle$ is 1. We then increased β progressively in steps of 0.1 and for each β swept the lattice 40 times, taking as the starting configuration the final lattice configuration of the preceding β. For the group $\overline{I}$ we carried out 50 lattice sweeps at each β. For each β the average $\langle S(p) \rangle$ was calculated over all plaquettes in all 40 or 50 configurations. This was continued to a maximum of $\beta = 3$ for $\overline{T}$, $\beta = 4$ for $\overline{O}$, and $\beta = 7$ for $\overline{I}$. The results are shown in Figs. 1–3. For β at which two points are plotted, the upper point is the value obtained in this sequence of iterations. When the maximum β was reached, we set all the link variables to a single element of $C(G)$ since only fields in the gauge orbit of constant configurations contribute in the limit $\beta \to \infty$. Then we swept the lattice 40 times (50 for $\overline{I}$) and successively reduced β in steps of 0.1 back down to 0. These results are also shown in Figs. 1–3 and are the lower set of points for those β at which two values have been plotted. An estimate of the uncertainty at each β is provided by the difference between the upper and lower sets of points.

The solid lines at small β in each graph are the two leading terms, $1 - \beta/4$, of the small β expansion for $\langle S(p) \rangle$. These terms are the same for $\overline{T}$, $\overline{O}$, $\overline{I}$, and SU(2) itself. The solid line at large β is the leading term in the large-β (small-g_0) expansion for SU(2), $3/(4\beta)$. The results of Creutz for $\langle S(p) \rangle$ for SU(2) are shown in Fig. 4 along with the two asymptotic curves and our values for $\overline{I}$. For clarity we show only points obtained by successively decreasing β. It is clear that our results for all three discrete groups follow SU(2) quite well in the small-β region. In addition, the curve for $\overline{I}$ stays with SU(2) all the way up into the region in which the large-β expansion is valid. We found this result somewhat surprising.

Each of the graphs of mean plaquette action shows a hysteresis loop—a region in which the value of $\langle S(p) \rangle$ obtained by progressively raising β disagrees significantly with the value obtained by lowering β. These loops strongly suggest that a phase transition occurs somewhere in the interval of β over which the loop extends. To find the order of the transition and the critical β, iterations were performed on mixed lattice starting configurations. At various β within the range of the loop, initial lattices were constructed by splicing together half of the final configuration found on the upper branch of the hysteresis loop with half of the final configuration on the lower branch of the loop. Then this starting lattice was swept 600 times. For β at the low end of the hysteresis loop the upper values of $\langle S(p) \rangle$ eventually won the iteration contest, and for β at the upper end of the loop, the lower values of $\langle S(p) \rangle$ eventually won.

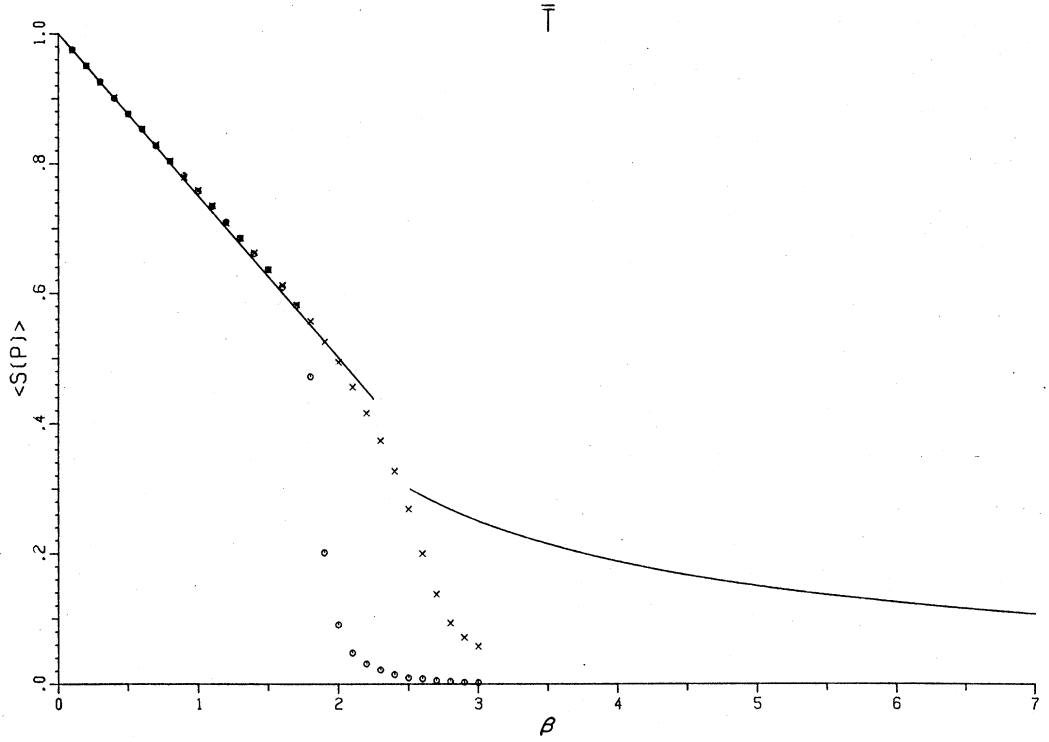

FIG. 1. $\langle S(p)\rangle$ as a function of β for $\bar{T}$ obtained by successively raising β from 0 to 3 (×) then lowering β from 3 back to 0 (○).

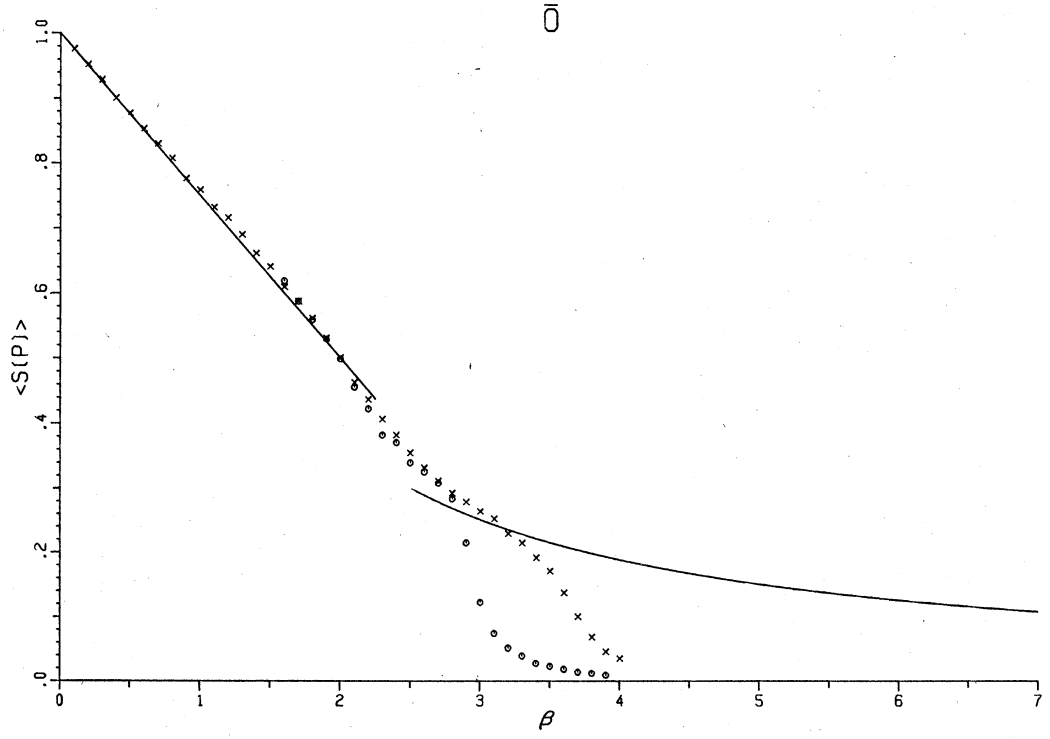

FIG. 2. $\langle S(p)\rangle$ as a function of β for $\bar{O}$ obtained by successively raising β from 0 to 4 (×) then lowering β from 4 back to 0 (○).

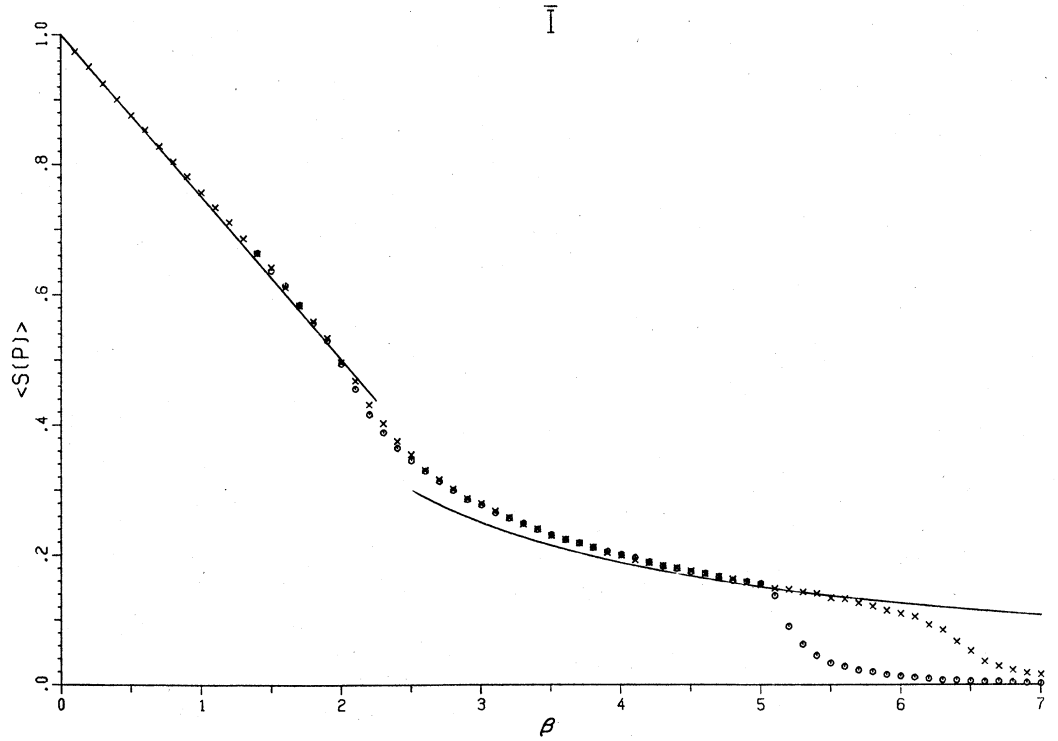

FIG. 3. $\langle S(p)\rangle$ as a function of β for $\bar{I}$ obtained by successively raising β from 0 to 7 (×) then lowering β from 7 back to 0 (○).

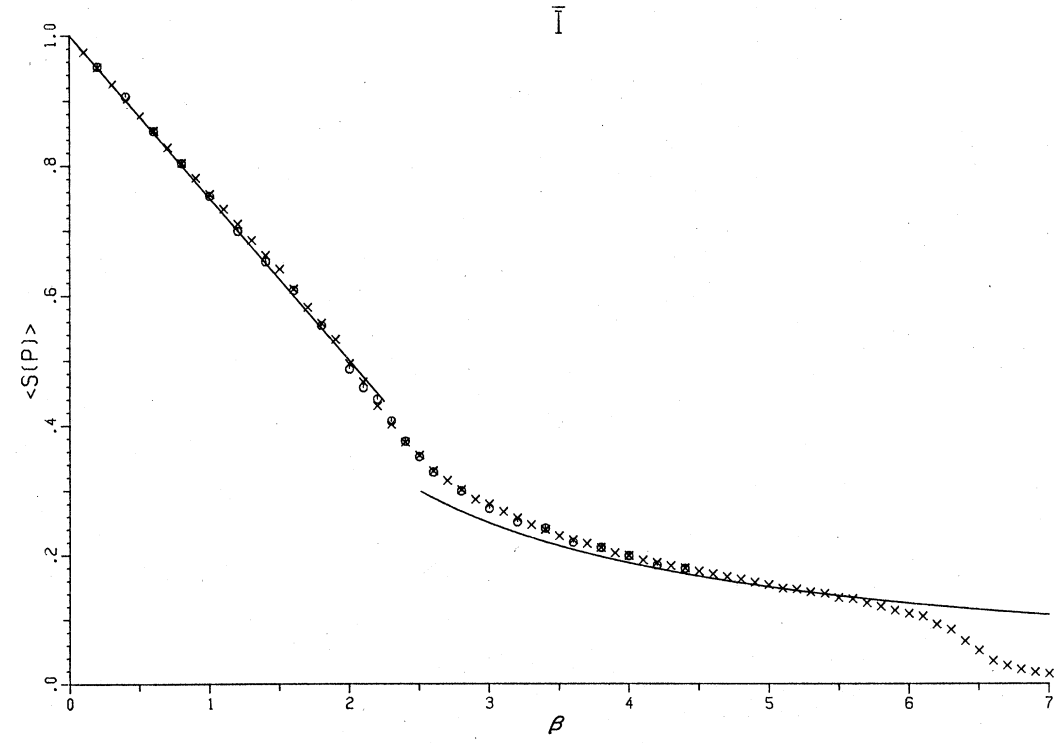

FIG. 4. $\langle S(p)\rangle$ as a function of β for $\bar{I}$ (×) and Su(2) (○). Values for SU(2) are from Ref. 2.

At a critical β within the loop range, the final equilibrium values of $\langle S(p)\rangle$ changed abruptly from the upper branch of the loop to the lower, suggesting a first-order phase transition. Figure 5 shows the result of such sweeps for $\overline{T}$ with $\beta=2.10$, 2.15, and 2.20. Figure 6 shows the results of $\overline{O}$ with β ranging from 3.20 to 3.30 in steps of 0.02. Figure 7 shows the results for $\overline{I}$ with β ranging from 5.96 to 6.06 in steps of 0.02. These results yield critical β_c of 2.15 ± 0.05 for $\overline{T}$, 3.25 ± 0.01 for $\overline{O}$, and 6.01 ± 0.01 for $\overline{I}$.

Since the mean action per plaquette for $\overline{I}$ duplicates that of SU(2) so well, a more detailed comparison between $\overline{I}$ and SU(2) seemed worthwhile. We therefore also evaluated the averages $\langle W(l)\rangle$ for square loops of sides 2, 3, 4, and 5. For a side of 1, $\langle W(l)\rangle$ is the quantity $\langle W(p)\rangle$ already discussed. Our results for $\overline{I}$ on an $8\times8\times8\times8$ lattice and β up to 3 are shown in Fig. 8 in comparison to Creutz's results for SU(2). Again the agreement between $\overline{I}$ and SU(2) is excellent. Our estimates of the errors for $\overline{I}$ in all cases are smaller than the plus signs used to indicate points. The errors are almost all less than 3% with the exception of the errors in the smallest values found for loops of sides 3, 4, and 5, in which cases the errors are as much as 20%. The curves in Fig. 8 are the leading term in the small-β expansion for $\langle W(l)\rangle$. Each of our values of $\langle W(l)\rangle$ was obtained from an ensemble of 100 lattices generated from a starting configuration given by the final configuration obtained in our evaluation of $\langle S(p)\rangle$. On each of the 100 lattices for each β, $\langle W(l)\rangle$ was found by averaging over all square loops of the appropriate size on the entire lattice. The dependence of the averages on lattice size for $\beta=3$ is shown in Fig. 9. The dependence on lattice size is less at lower β.

Now it is expected that for loops with large values of area A the averages $\langle W(l)\rangle$ will approach the area law

$$\ln\langle W(l)\rangle \to KA + o(A)$$

with some string tension K determined by β. If this holds, static quarks are confined. It is by no means clear that loops only up to size 5 are sufficiently large to extract K reliably beyond the region of small β. The numerical uncertainties in Monte Carlo values of $\langle W(l)\rangle$ add an additional complication. Nonetheless, if we simply ignore these problems and follow Creutz's fitting procedure, we can obtain a set of values of $K(\beta)$ for $\overline{I}$ to compare with $K(\beta)$ for SU(2). The results are shown in Fig. 10 along with Creutz's results.

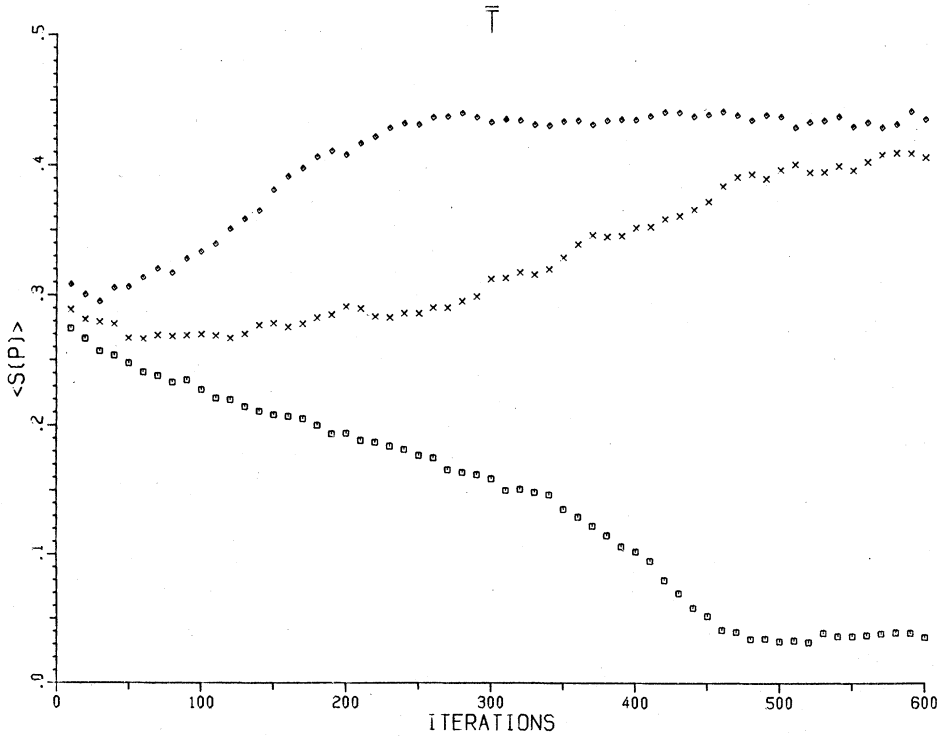

FIG. 5. $\langle S(p)\rangle$ for $\overline{T}$ for 600 iterations beginning with split lattice configurations. The values of β are 2.10 for the highest set of points, 2.15 for the middle set of points, and 2.20 for the lowest set of points.

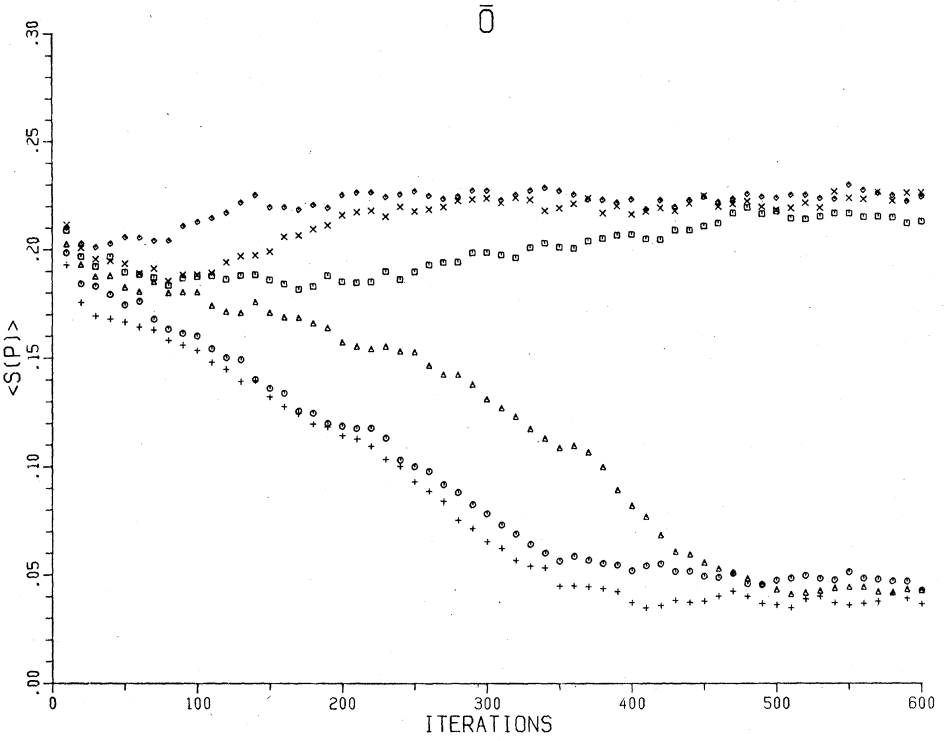

FIG. 6. $\langle S(p)\rangle$ for $\bar{O}$ for 600 iterations beginning with split lattice configuration. The values of β for the upper three sets of points are 3.20, 3.22, and 3.24, and for the lower three sets β is 3.26, 3.28, and 3.30.

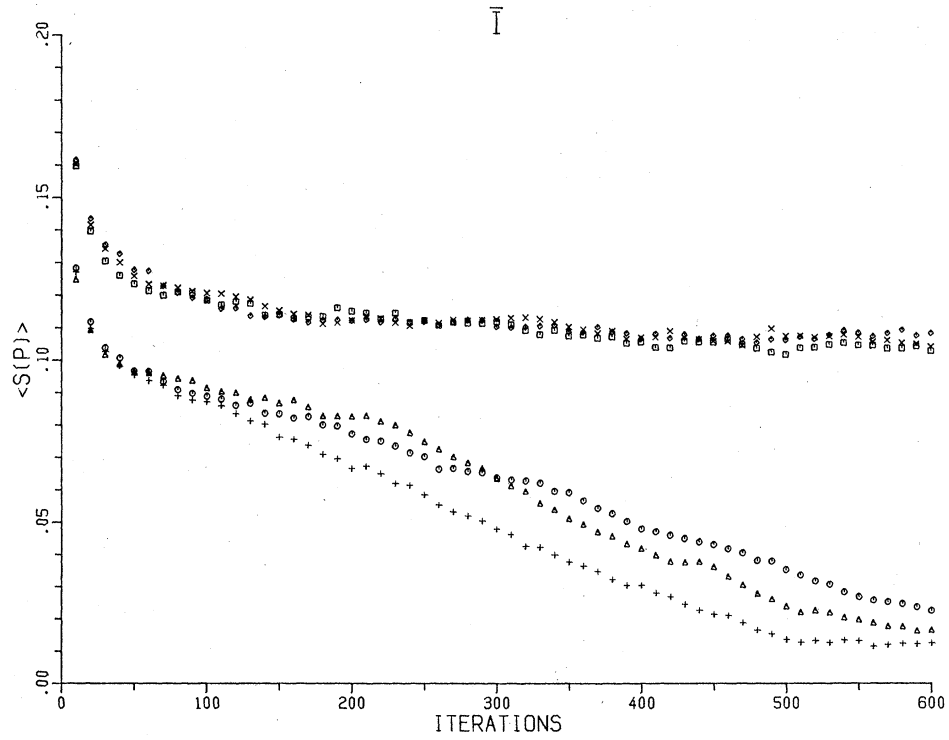

FIG. 7. $\langle S(p)\rangle$ for $\bar{I}$ for 600 iterations beginning with split lattice configurations. The values of β for the upper band of points are 5.96, 5.98, and 6.00. For the lower three sets of points β is 6.02, 6.04, and 6.06.

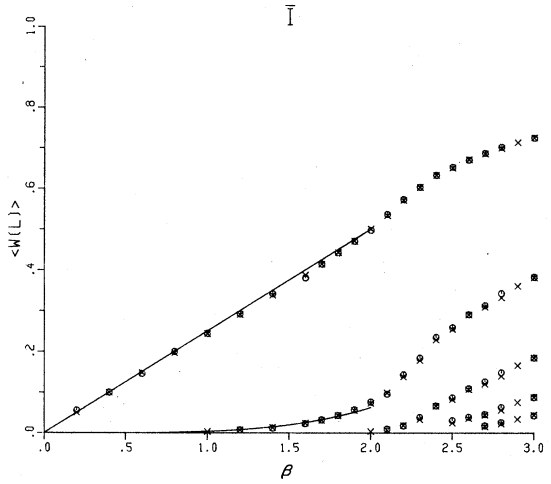

FIG. 8. $\langle W(l)\rangle$ for $\bar{I}$ (×) and SU(2) (0) for square loops of sides 1, 2, 3, 4, and 5 corresponding, respectively, to the highest through the lowest sets of points. All errors for $\bar{I}$ are smaller than the (×) indicating points.

The tension shown in Fig. 11 is measured in units of (lattice spacing)$^{-2}$. The plus signs are $K(\beta)$ for $\bar{I}$ and the crosses are $K(\beta)$ for SU(2). The solid line for small β is the leading term in the small-β expansion for $K(\beta)$. The solid line at large β is $K_0\exp(-6\pi^2\beta/11)$, which is predicted[2] by a continuum renormalization-group argument for SU(2) as the main behavior of $K(\beta)$ for β large. The value of K_0 is not predicted by this argument and

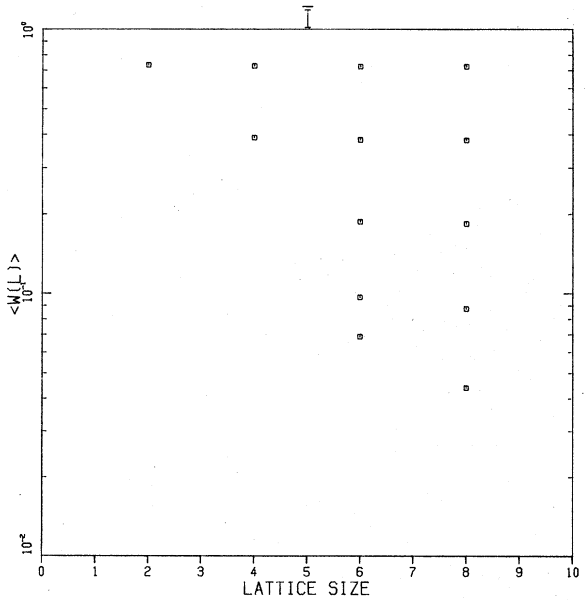

FIG. 9. $\langle W(l)\rangle$ for $\bar{I}$ for square loops as a function of lattice size with $\beta = 3.00$. Loops of sides 1, 2, 3, 4, and 5 correspond, respectively, to the highest through the lowest sets of points.

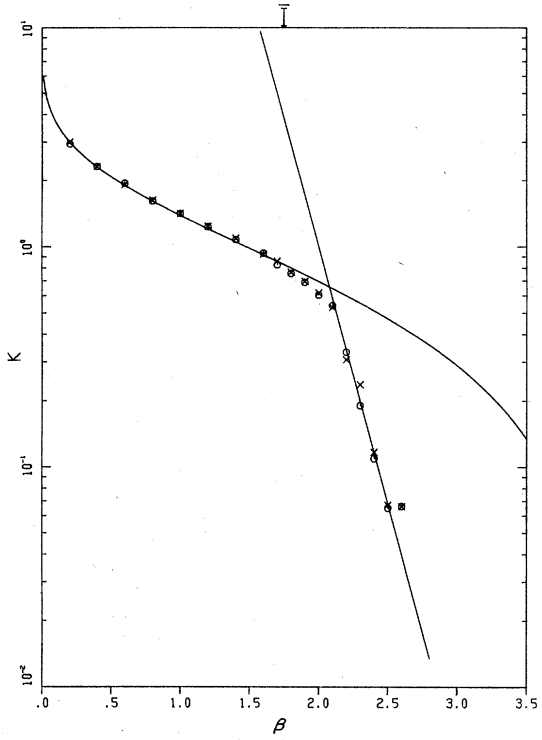

FIG. 10. String tension K in units of (lattice spacing)$^{-2}$ for $\bar{I}$ (×) and SU(2) (○). The SU(2) points are from Ref. 2.

was found instead by Creutz by a fit to his results for SU(2).

The points in this graph were extracted from values of $\langle W(l)\rangle$ as follows. For $\beta \leq 1.4$, loops of side 0 and 1 were fitted to

$$\ln\langle W(l)\rangle = KA + K'. \tag{3.1}$$

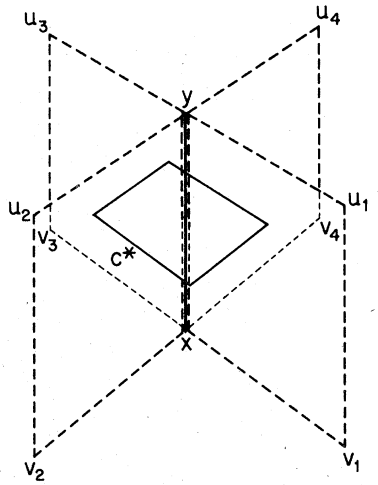

FIG. 11. The solid vertical line is a link $[x,y]$ with $U(x,v) = -1$, the dashed plaquettes are $[u_i,v_i,x,y]$, $i = 1,\ldots,4$, and the horizontal solid square loop is c^*.

For $1.6 \leq \beta \leq 2.1$ loops of side 0, 1, and 2 were fitted to

$$\ln W(l) = KA + K'\sqrt{A} + K''. \tag{3.2}$$

For $2.2 \leq \beta \leq 2.4$, loops of side 1, 2, and 3 were fitted to (3.2), and for $\beta = 2.5$ and 2.6, loops of side 1, 2, 3, and 4 were fitted to (3.2). The same procedure was used by Creutz for each of the SU(2) points. Once again, $\bar{I}$ closely follows SU(2) from the region in which small-β expansions work to the region in which the large-β approximation is reliable.

It is easily shown[2] that if $K(\beta)$ does agree with the large-β curve in Fig. 11, then restoring the lattice spacing a and fixing $K(\beta)$ in macroscopic units as a renormalization condition, we will obtain, as $a \to 0$, $g_0^2(a) = 4\beta^{-1}$, which goes as the continuum perturbation theory prediction

$$g_0(a)^2 = \frac{12\pi^2}{11 \ln(a_0/a)} + o\left(\ln^{-2}\left(\frac{a_0}{a}\right)\right)$$

for some constant a_0. In other words, $g_0(a)^2$ will exhibit asymptotic freedom. Thus in a sense Fig. 11 suggests not only that an SU(2) lattice theory exhibits asymptotic freedom but also that an $\bar{I}$ theory does so too, at least over a restricted range of g_0^2.

IV. PHASE TRANSITION

We will now present a model of the large-β phase transition found in gauge theories with discrete gauge groups. The model uses a reformulation of the path integral over gauge potentials on a D-dimensional lattice as a sum over closed $(D-2)$-dimensional surfaces on the dual lattice which may be thought of as the trajectories of $(D-3)$-dimensional lines of flux of the dual electromagnetic field. When β is large, the flux surfaces have finite size, but as β is lowered, the surfaces grow, and at a critical β the average size diverges, giving rise to a phase transition. We will begin with an exact discussion of the phase transition in a Z_2 gauge theory, and then show how this discussion can be adapted to yield an approximate picture of the transitions in theories with gauge group Z_N, $N > 2$, $\bar{T}$, $\bar{O}$, and $\bar{I}$. As mentioned in the Introduction, the picture we will develop tends to suggest that this class of transitions is a special property of discrete gauge groups and has no analog in gauge theories with continuous gauge groups.

Consider a Z_2 gauge theory on a D-dimensional hypercubic lattice Λ with free boundary conditions. The link variables $U(x,y)$ now range over the values 1 and -1 and loops $W(p)$ are defined by (2.1) without a trace or factor of $\frac{1}{2}$. The action S and vacuum expectation value $\langle \mathfrak{I} \rangle$ are given by (2.2) and (2.3), respectively, without change. In particular, the partition function Q in (2.3) is

$$Q = \sum \exp(-\beta S), \tag{4.1}$$

and the mean action per plaquette $S(p)$ is

$$\langle S(p) \rangle = -\frac{\partial}{\partial \beta} P^{-1} \ln Q, \tag{4.2}$$

where P is the number of plaquettes in Λ. Thus the phase transition found in Ref. 5 in Z_2 as a rapid change in $\langle S(p) \rangle$ as a function of β should appear as a singularity in the β dependence of the infinite-volume limit of $P^{-1}\ln Q$.

To obtain a picture of the phase transition in the Z_2 theory we will represent configurations of gauge fields on Λ as geometric figures on the lattice $\tilde{\Lambda}$ dual to Λ. The sites of $\tilde{\Lambda}$ are located at the centers of the D-cubes of Λ, the links of $\tilde{\Lambda}$ pass through the centers of the $(D-1)$-cubes of Λ, and so on. The m-dimensional element f^* of $\tilde{\Lambda}$ passing through the center of the $(D-m)$-dimensional element f of Λ will be called the dual of f.

Before stating in general how to convert fields on Λ to geometric objects on $\tilde{\Lambda}$, let us consider two simple examples which have small values of S and therefore make relatively large contributions to Q in (4.1). The smallest possible S is 0 occurring on the gauge orbit of the identity configuration with all $U(x,y) = 1$. On the dual lattice this will be represented by the absence of any object. The lowest $S > 0$ occurs on the gauge orbits of configurations with $U(x,y) = -1$ on one link $[x,y]$, and all other $U(x',y') = 1$. In this case we have $W(p) = -1$, $S(p) = 2$, on each of the $2D$-2 plaquettes $[u_i, v_i, x, y]$, $i = 1, \ldots, 2D$-2, passing through the link $[x,y]$. On all other plaquettes $W(p) = 1$, $S(p) = 0$. Thus $S = 2(2D$-2). The dual $[u_i, v_i, x, y]^*$ to each of the plaquettes $[u_i, v_i, x, y]$ is a $(D$-2)-dimensional element of $\tilde{\Lambda}$. The surface formed by taking these dual elements $[u_i, v_i, x, y]^*$ together is closed and can be regarded formally as the sum

$$c^* = \sum_i [u_i, v_i, x, y]^*. \tag{4.3}$$

c^* can be interpreted as the trajectory swept out by the $(D$-3)-dimensional lines of flux of the dual electromagnetic field arising from the gauge potential $U(x,y)$. An example of $[x,y]$, $[u_i, v_i, x, y]$, and c^* for $D = 3$ is shown in Fig. 11.

In both of the cases considered we have shown that fields $U(x,y)$ on Λ can be represented by closed $(D$-2)-dimensional figures c^* on $\tilde{\Lambda}$ with the action S of the original field given by $2A(c^*)$, where $A(c^*)$ is the number of elements in c^*.

This correspondence can be generalized to arbitrary fields $U(x,y)$ on Λ as follows: We first assign the field $U(x,y)$ to a linear combination $\underline{U}$ of links $[x,y]$ with coefficients in the additive group Z_2 on the integers 0, 1,

$$\underline{U}=\sum \tfrac{1}{2}[1-U(x,y)][x,y]. \tag{4.4}$$

Each link in (4.4) appears in this sum exactly once oriented, say, with coordinates fulfilling $y_i \geq x_i$, $i=1,\ldots,D$. Such linear combinations of links are called 1-chains over Z_2. Next we define the coboundary operator ∇ on 1-chains by the condition that it act linearly and take each oriented link $[x,y]$ into the sum of all D-2 coherently oriented plaquettes $[u_i,v_i,x,y]$ which contain $[x,y]$ in their boundary

$$\nabla[x,y]=\sum[u_i,v_i,x,y]. \tag{4.5}$$

Linear combinations of two-dimensional elements with coefficients in Z_2 are called 2-chains over Z_2; m-chains over Z_2 for any $m=0,\ldots,D$ can be defined similarly. If we now form $\nabla\underline{U}$, the dual

$$c^*=(\nabla\underline{U})^* \tag{4.6}$$

is the generalization to arbitrary fields of the c^* in (4.3). It follows from standard results of combinatorial topology[7] that all c^* given by (4.6) are closed in $\tilde{\Lambda}$. Moreover, since Λ and $\tilde{\Lambda}$ are topologically trivial, every closed (D-2)-chain in $\tilde{\Lambda}$ can be realized as $(\nabla\underline{U})^*$ for a unique gauge orbit of fields $U(x,y)$ on Λ. It is not difficult to convince oneself that S for a general field $U(x,y)$ is simply $2A(c^*)$ as before, where $A(c^*)$ is the number of elements summed in c^*.

It is perhaps useful to mention here some facts about chains. For a general discussion see Ref. 7. The addition and multiplication appearing in an m-chain over Z_N obey

$$\begin{aligned} &1\times f=f,\\ &0\times f=0,\\ &\alpha\times f+\beta\times f=(\alpha+\beta)f, \end{aligned} \tag{4.7}$$

where f is any m-dimensional element. For coefficients in Z_2 we have in particular

$$f+f=(1+1)f=0. \tag{4.8}$$

For any coefficient group Z_N we let any f with orientation reversed be equal to $-f$. Thus

$$\begin{aligned} &[x,y]=-[y,x],\\ &[u,v,x,y]=-[y,x,v,u]. \end{aligned} \tag{4.9}$$

The condition of closure for a chain can be stated conveniently using the boundary operator Δ taking m-chains into $(m-1)$-chains. The operator Δ is defined by the condition that it acts linearly and takes any element f into its boundary oriented coherently with the orientation of f. Since a point has no boundary we have

$$\begin{aligned} &\Delta[x]=0,\\ &\Delta[x,y]=[y]-[x],\\ &\Delta[u,v,x,y]=[u,v]+[v,x]+[x,y]+[y,u]. \end{aligned} \tag{4.10}$$

A closed chain c then fulfills $\Delta c=0$. It can be easily verified that any boundary Δc is closed and thus $\Delta\Delta c=0$.

We can now reformulate (2.3) and (4.1) as sums over closed (D-2)-chains* on $\tilde{\Lambda}$. For example, Q becomes

$$Q=M\sum\exp[-2\beta A(c^*)], \tag{4.11}$$

where the sum is over all closed (D-2)-chains c^* on $\tilde{\Lambda}$ and M is the total number of gauge transformations of any field $U(x,y)$ on Λ.

The total contribution to (4.11) of all chains c^* of size A is $N(A)\exp(-2\beta A)$, where $N(A)$ is the number of chains of size A. Thus when β is very large, only chains with small A will contribute significantly, while if β is made progressively smaller, larger and larger chains will contribute. If we hit a β small enough so that $N(A)$ for large A rises faster than $\exp(-2\beta A)$ falls, then chains of arbitrarily large size A will suddenly contribute, limited only by the boundaries of $\tilde{\Lambda}$. In the infinite-volume limit the system will undergo a phase transition. This is the mechanism of the phase transition found in Ref. 5.

By considering the effect on a Wilson-loop amplitude of a linear superposition of widely separated fluctuations of form (4.3), it can be shown that for large β the Wilson loop obeys a perimeter law. Since strong-coupling expansions prove an area law for small β the phase transition suggested by (4.11) should be the transition between these two regions of β.

A rough estimate of the critical β_c can be obtained for $D=4$ by noticing that if we have a two-dimensional surface with boundary and add one more plaquette to the boundary there will be five different ways of orienting it. Equations (4.7)–(4.9) imply the plaquettes in c^* never overlap each other. Thus we expect roughly $N(A)=5^A$ different ways of forming a surface of area A. The critical β_c becomes

$$\beta_c=\frac{\ln 5}{2}=0.804\cdots.$$

The correct result given by exact duality arguments[7] combined with the Monte Carlo results of

Ref. 5 is

$$\beta_c=\frac{\ln(1+\sqrt{2})}{2}=0.440\cdots . \tag{4.12}$$

Let us now adapt the preceding discussion to a theory with gauge group Z_N, $N>2$. In this case each $U(x,y)$ ranges over the values $\exp(2\pi im/N)$, $m=0,\ldots,N-1$, $S(p)$ in (2.2) must be replaced by $1-\mathrm{Re}W(p)$, and Eqs. (2.3), (4.1), and (4.2) hold without change. When β is large the main contribution to (4.1) will come from configurations with $S(p)$ for each plaquette given either by 0 or its minimum value greater than 0, which is $1-\cos(2\pi/N)$. The configurations which give rise to such $S(p)$ are all in the gauge orbit of fields with each $U(x,y)$ either 1, $\exp(2\pi i/N)$ or $\exp(-2\pi i/N)$. For such fields dual representations can be constructed similar to the representation for Z_2.

For each $[x,y]$ with $U(x,y)=1$, let $\alpha(x,y)=0$. If $U(x,y)=\exp(2\pi i/N)$, let $\alpha(x,y)=1$, and if $U(x,y)=\exp(-2\pi i/N)$, let $\alpha(x,y)=-1$. Then $\underline{U}$ is defined to be a 1-chain over Z_N, rather than over Z_2, given by

$$\underline{U}=\sum\alpha(x,y)[x,y]. \tag{4.13}$$

We define c^* again by

$$c^*=(\nabla\underline{U})^*. \tag{4.14}$$

It becomes a closed $(D-2)$-chain over Z_N. As a result of our simplified choice of $U(x,y)$, the coefficients which may appear in c^* are restricted to 0, 1, and -1. If we had begun with an arbitrary Z_N field, $U(x,y)$, the natural extension of our definition of $\alpha(x,y)$ would have led to a $\underline{U}$ with arbitrary coefficients in Z_N, realized as the integers mod N. The number of coefficients $A(c^*)$ differing from 0 in (4.14) is equal to the number of plaquettes p on Λ with $S(p)=1-\cos(2\pi/N)$ for the original field $U(x,y)$. Thus the action S for $U(x,y)$ becomes $[1-\cos(2\pi/N)]A(c^*)$.

The contribution to Q in (4.1) arising from fields $U(x,y)$ with all $S(p)$ either 0 or $1-\cos(2\pi/N)$ can now be represented on the dual lattice as

$$Q=M\sum\exp\{-\beta[1-\cos(2\pi/N)]A(c^*)\}, \tag{4.15}$$

where the sum is over all closed oriented $(D-2)$-chains c^* with coefficients 0, 1, or -1.

Every $c^*_{Z_N}$ which appears in (4.15) can be made into a $c^*_{Z_2}$ which appears in (4.11) simply by replacing all coefficients of -1 by 1. Thus the number of distinct $c^*_{Z_N}$ for each $c^*_{Z_2}$ is just the number of ways of orienting each $c^*_{Z_2}$ by replacing some of its coefficients of 1 by -1 consistent with the condition that $\Delta c^*_{Z_N}$ must vanish.

Expression (4.15) should exhibit a phase transition of the same sort that is shown by (4.11). Since (4.15) arises from a sum over a restricted set of field configurations $U(x,y)$, however, it is not immediately obvious what relation a transition in (4.15) should have to a transition in the exact Z_N theory. We believe the transition in (4.15) is closely related to the exact transition. If $N(S)$ is the number of configurations of $U(x,y)$ with fixed action S, then $\ln N(S)$ should be well approximated for large S by the set of configurations included in (4.15) with all $S(p)$ either 0 or $1-\cos(2\pi/N)$. For example, for reasonably large N, if we permit m plaquettes with $S(p)=1-\cos(4\pi/N)$, the lowest value above $1-\cos(2\pi/N)$, the increase in action will be nearly the same as letting $3m$ plaquettes which originally had $S(p)=0$ to be given $S(p)=1-\cos(2\pi/N)$. There are many more ways of adding $3m$ plaquettes with $S(p)=1-\cos(2\pi/N)$ than there are of raising m plaquettes from $1-\cos(2\pi/N)$ to $1-\cos(4\pi/N)$. If the configurations counted in (4.15) do yield a reliable approximation to $\ln N(S)$, then following our discussion of the phase transition in Z_2, Eq. (4.15) should also give a reliable value of the critical β_c.

On the other hand, by considering a few typical examples of large surfaces it is not hard to show that for large $A(c^*)$ the number of oriented $c^*_{Z_N}$ entering (4.15) for each $c^*_{Z_2}$ in (4.11) is typically at most of the order $2^{A/2}$ or $2^{A/4}$, while the total number of surfaces itself, as we showed, is roughly 5^A. Thus we expect $\ln N(A)$ for $c^*_{Z_N}$ in (4.15) to have about the same large-A asymptotic behavior as $\ln N(A)$ for $c^*_{Z_2}$ in (4.11). Therefore, the critical points of (4.11) and (4.15) should be related by

$$[1-\cos(2\pi/N)]\,\beta_c(Z_N)=2\beta_c(Z_2),$$

and by (4.12) we obtain

$$\beta_c(Z_N)=\frac{\ln(1+\sqrt{2})}{1-\cos(2\pi/N)}\approx\frac{0.881}{1-\cos(2\pi/N)}. \tag{4.16}$$

This formula agrees quite well with the results found in Ref. 5 and in effect explains the fit

$$\beta_c(Z_N)=\frac{0.78}{1-\cos(2\pi/N)}$$

found empirically in Ref. 5. A comparison of (4.16) with $\beta_c(Z_N)$ reported in Ref. 5 for $N=3$, 4, 5, 6, and 8 is shown in Fig. 12.

We are at last ready to adapt our model of the large-β phase transition to $\bar{T}$, $\bar{O}$, and $\bar{I}$. Let G be the particular groups we are considering. Motivated by the discussion of the transition in Z_N we will assume that a reasonable approximation to Q in (4.1) near the phase transition is obtained by restricting the allowed set of fields to those with each $U(x,y)$ either 1 or an element of $C(G)$

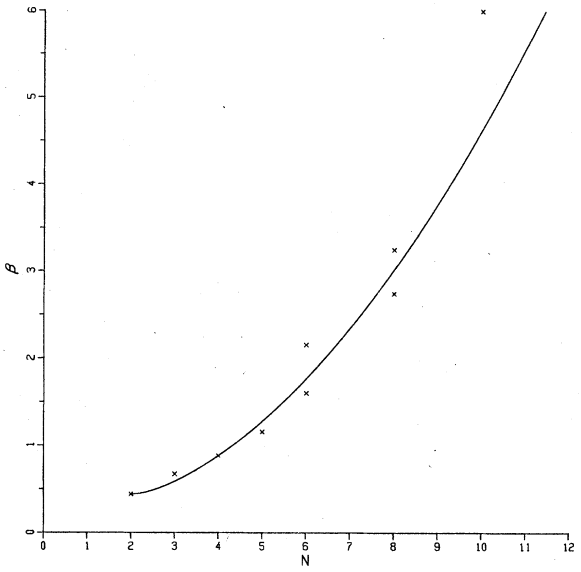

FIG. 12. Critical β_c compared with the predictions of Eqs. (4.16) and (4.18). The groups Z_N, with $N=2$, 3, 4, 5, 6, and 8 are the lower points at $N=2$, 3, 4, 5, 6, and 8, respectively. The upper points at $N=6$, 8, and 10 are $\overline{T}$, $\overline{O}$, and $\overline{I}$, respectively. Values for Z_N are from Ref. 5.

and each $S(p)$ either 0 or its minimum allowed value above 0. The set $C(G)$ is defined in Sec. III as the collection of elements of G closest to the identity but excluding the identity. We will also impose the restriction that if $U(x,y)$ is assigned to a link on the boundary of a plaquette p, the other $U(x,y)$ on that p must be either 1, $U(x,y)$ or $U(x,y)^{-1}$. We have not examined in detail the extent to which this condition is likely to restrict $N(S)$ for the set we are considering. We will return to this question shortly.

With these conditions, the allowed fields $U(x,y)$ contributing to the partition function Q for G split up into disjoint regions each of which has $U(x,y)$ given only by 1, g or g^{-1} for a particular $g\in C(G)$. Since G has a finite number of elements, for some N, $g^N=1$. Thus, the subgroup generated by g is isomorphic to Z_N. The minimum action $S(p)$ defined by (2.2) then turns out to be $1-\cos(2\pi/N)$, and the same N applies to all $g\in C(G)$. It follows by an adaptation of the discussion we gave for Z_N, that the partition function Q for G becomes

$$Q=M\sum\exp\{-\beta[1-\cos(2\pi/N)]A(c^*)\}m^{B(c^*)}, \tag{4.17}$$

where c^* is a closed $(D-2)$-chain on $\tilde{\Lambda}$, m is half the number of elements in $C(G)$, and $B(c^*)$ is the number of disjoint regions into which c^* can be separated. For any particular disjoint region of c^* there are m different ways to assign that region a $g\in C(G)$ consistent with our restrictions, hence the factor of $m^{B(c^*)}$.

This expression should exhibit a phase transition similar to the transition in (4.12). For large $A(c^*)$ we would expect the disjoint regions of typical c^* to become large so that the factor $m^{B(c^*)}$ makes a relatively unimportant contribution to $\ln N(A)$. We therefore expect the critical β_c here to be the same as for Z_N:

$$\beta_c(G)=\frac{\ln(1+\sqrt{2})}{1-\cos2\pi/N}\,. \tag{4.18}$$

For $\overline{T}$, $\overline{O}$, and $\overline{I}$ it turns out that N is 6, 8, and 10, respectively. A comparison of this prediction with our results discussed in Sec. III is shown in Fig. 12.

We now return briefly to the effect of our restriction on $U(x,y)$ which share a common plaquette. Clearly this approximation reduces $N(S)$ and therefore lowers our prediction for β_c below the correct value. At least as the lattice dimension D becomes large, however, the main contribution to $\ln N(S)$ will come from the number of ways large c^* in (4.17) can be embedded in $\tilde{\Lambda}$ and our condition on $U(x,y)$ will be a small correction to $\ln N(S)$. Looking at Fig. 12 it is clear that dimension 4 is already large enough for our approximation to be fairly accurate for $\overline{T}$, $\overline{O}$, and $\overline{I}$. It is least reliable for $\overline{I}$ presumably because $\overline{I}$ has the largest number of elements and therefore in this case our approximation does the most damage to the true $N(S)$.

The error in (4.18) will tend to increase for larger groups. A version of (4.18) for large-dimension discrete subgroups of SU(3), for example, could be completely unreliable.

Finally, it is perhaps worth mentioning the obvious fact that for continuous gauge groups the allowed values of each $S(p)$ extend continuously down to 0. Thus, in effect, the correct value of N in (4.18) is infinite, β_c becomes infinite and transitions of the sort we have discussed do not occur.

ACKNOWLEDGMENTS

This work was supported in part by the U. S. Department of Energy. We would like to thank Mark Sweeney and Michael Creutz for valuable discussions. We are also grateful to Michael Creutz for providing data for Figs. 4, 8, and 10.

[1]K. Wilson, Phys. Rev. D 10, 2445 (1974).
[2]M. Creutz, Phys. Rev. Lett. 43, 553 (1979); 43, 890 (E) (1979); Phys. Rev. D 21, 2308 (1980).
[3]K. Wilson, Cornell report, 1979 (unpublished).
[4]While this work was in process we received a report in which Monte Carlo techniques are used to determine the phase structure of lattice gauge theories for $\bar{T}$ and $\bar{O}$. The results agree with ours. See C. Rebbi, Phys. Rev. D 21, 3350 (1980).
[5]M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. Lett. 42, 1390 (1979); Phys. Rev. D 20, 1915 (1979).
[6]For earlier discussions of this phase transition in Z_N gauge theories see S. Elitzur, R. Pearson, and J. Shigemitsu, Phys. Rev. D 19, 3698 (1979); D. Horn, M. Weinstein, and S. Yankielowicz, *ibid.* 19, 3715 (1979); A. Ukawa, P. Windey, and A. Guth, *ibid.* 21, 1013 (1979); M. Einhorn, R. Savit, and E. Rabinovici, University of Michigan Report No. UMHE 79-25, 1979 (unpublished).
[7]P. Alexandroff, *Combinatorial Topology* (Graylock, Rochester, New York, 1956).
[8]R. Balian, J. M. Drouffe, and C. Itzykson, Phys. Rev. D 11, 2098 (1975); 11, 2104 (1975).