PHYSICAL REVIEW D VOLUME 21, NUMBER 8 15 APRIL 1980

$1/N$ expansions and the phase diagram of discrete lattice gauge theories with matter fields

John B. Kogut
Physics Department, University of Illinois at Urbana-Champaign, Urbana, Illinois, 61801
(Received 16 January 1980)

Ising lattice gauge theory with matter fields is generalized to a family of gauge-invariant N-state models. These models are soluble in the $N \to \infty$ limit. In both two and three spatial dimensions the $N \to \infty$ limits of these models possess three phases separated by lines of first-order transitions. One phase admits free charge, the second confines, and the third experiences the Higgs phenomenon. $1/N$ corrections are calculable and terminate the boundary between the confining and Higgs phases. A closed region of free charge persists for all $N \geq 1$.

I. INTRODUCTION

There has been considerable interest recently in the phase diagram of Ising lattice gauge theory with matter fields.[1] Although this model has several special and simple features, it presents some problems common to lattice quantum chromodynamics. For example, in both theories the Wilson correlation function is short-ranged due to screening and an order parameter distinguishing confinement and free-charge phases is not known.[2] However, some information concerning the phase diagram of Ising gauge theory with matter fields has been obtained using expansion methods[1,2] and Monte Carlo simulations.[3,4]

Ising gauge theories with matter fields are characterized by (at least) two temperaturelike variables—one controlling the fluctuations of the gauge fields and the other controlling the matter fields. Conventional expansion methods are useful here only when one of the parameters is near an extreme, either zero or infinity, and the other is free to vary. The computer simulations have the advantage that they can sometimes yield reliable data when *both* parameters are on the order of unity and one is observing the interior of the system's phase diagram. These calculations have produced some expected as well as unexpected results which will be discussed in detail in the latter sections of this article. Such ponderous methods, however, challenge us to find simple, physical arguments and calculations to explain or rederive the resulting phase diagrams. This article presents such an approach.

The strategy of the approach is the following. Instead of concentrating just on Ising lattice gauge theory with matter fields, we generalize this theory to an entire family, N-state Potts lattice gauge theories with matter fields. This represents a natural generalization from the Ising-type model in which link and site variables can assume two possible values to ones in which those variables can assume N possible values. These models have the important property that they are *soluble* in the $N \to \infty$ limit and $1/N$ corrections are easily calculated and understood. We will present calculations of the free energy and latent heats and will be able to understand the phase diagram of the $N=2$ model in three and four dimensions in simple terms. The only approximation will consist of the $1/N$ expansion itself—the coefficients of this expansion are complicated functions of the temperaturelike variables in the problem and they are computed *exactly*. Therefore, the method allows us to probe everywhere within the *interior* of the phase diagram for each model of fixed N. An additional bonus of the method is that its results are good even for relatively small values of N. This will allow us to make some interesting semiquantitative remarks about Ising lattice gauge theory with matter fields. These results include the following:

(1) If the gauge field temperature is low and the matter field temperature is high, the theory resides in a phase in which free charge can exist.

(2) This phase is completely surrounded by a line of transitions separating it from a phase which can be labeled "confinement-Higgs."

(3) There is a line of first-order transitions extending from the boundary of the phase of free charge a finite distance into the region of confinement. In $D+1=3$ dimensions this line rests on the self-dual line of the model.

We shall see that these results and the physics behind them can be read off a calculation of very modest proportions. This is, perhaps, the nicest feature of this approach to the problem.

II. N-STATE POTTS LATTICE GAUGE THEORY WITH MATTER FIELDS

Before we consider the models and calculations of immediate interest, it is useful to review some relevant facts about Ising lattice gauge theory with

matter fields.[1,2] That model contains Ising variables which reside on sites $\vec{r}$ of a $(D+1)$-dimensional cubic lattice,[5]

$$\sigma(\vec{r}) = \pm 1 . \tag{2.1}$$

These variables simulate matter fields which belong to the fundamental representation of the gauge group. In addition there are Ising variables which reside on links $(\vec{r},\mu)$

$$U_\mu(\vec{r}) = \pm 1 , \tag{2.2}$$

which simulate gauge fields. The Euclidean action S of this model consists of two terms: a hopping term for the matter fields and a four-spin interaction for the gauge fields

$$S = \beta \sum_{\vec{r},\mu} \sigma(\vec{r}) U_\mu(\vec{r}) \sigma(\vec{r}+\hat{e}_\mu) + \kappa \sum_{\vec{r},\mu\nu} U_\mu(\vec{r}) U_\nu(\vec{r}+\hat{e}_\mu) U_\mu(\vec{r}+\hat{e}_\nu) U_\nu(\vec{r}) , \tag{2.3}$$

where $\hat{e}_\mu$ is a unit lattice vector in the direction μ and the four U variables in the second term belong to the four edges of a plaquette. (Since the reader is presumed to be familiar with lattice gauge theory, our notation will usually be abbreviated.) An important feature of S is its invariance under local Z_2 gauge transformations

$$\begin{aligned} &\sigma(\vec{r}) \to \sigma(\vec{r}) s(\vec{r}) , \\ &U_\mu(\vec{r}) \to s(\vec{r}) U_\mu(\vec{r}) s(\vec{r}+\hat{e}_\mu) , \end{aligned} \tag{2.4}$$

where $s(\vec{r})$ is an Ising variable. Direct substitution shows that Eq. (2.3) is invariant under the transformation Eq. (2.4).

A number of facts are known about the action Eq. (2.3). Consider its phase diagram in the (β,κ) plane and concentrate on its boundaries.

(1) $\beta = 0$. The matter fields decouple and the theory reduces to Z_2 Ising gauge theory. In $D+1 = 3$ dimensions this theory is dual[1] to the Ising model which has an isolated critical point. In $D+1 = 4$ dimensions, the theory is self-dual,[1] and Monte Carlo[6] and strong-coupling studies[7] have shown that the theory undergoes a first-order phase transition at its self-dual point κ^*. For $\kappa < \kappa^*$ the theory confines and for $\kappa > \kappa^*$ it does not.

(2) $\kappa = \infty$. The gauge fields are frozen out and the theory reduces to the Ising model which has an isolated critical point β_c in both $D+1 = 3$ and 4. For $\beta > \beta_c$ the theory is ordered, $\langle \sigma(\vec{r}) \rangle = 0$, while for $\beta < \beta_c$ it is not.

(3) Other boundaries. On both the $\kappa = 0$ and $\beta = \infty$ boundaries the theory becomes trivial.

Less is known about the interior of the phase diagram, but there are two important results concerning the immediate vicinity of the boundaries.

(1) In $D+1 = 3$ dimensions, the critical points on the $\beta = 0$ and $\kappa = \infty$ lines are the end points of critical lines. For example, near the gauge-field fixed point κ_c, the coupling to matter fields produces a finite renormalization in κ,[1,2]

$$\kappa_{\text{eff}} = \kappa + \tfrac{1}{8}\beta^4 + \cdots . \tag{2.5}$$

In $D+1 = 4$ dimensions, the Ising critical point persists at least a small distance into the phase diagram.

(2) There is a strip along the $\kappa = 0$ boundary and the $\beta = \infty$ boundary where the free energy is analytic.[2] So the confining phase of the gauge theory $(\kappa < \kappa_c)$ and the magnetized phase of the Ising model $(\beta > \beta_c)$ are not qualitatively different.

Piecing together these facts leads to the phase diagram shown in Fig. 1. The dashed line represents a common prejudice that the region of free charge (no confinement along the $\beta = 0$ axis for $\kappa > \kappa_c$) is separated by a phase boundary from the confinement region. It is quite puzzling, however, that no one has found a simple symmetry criterion to distinguish these two phases.

Detailed information concerning the interior of the phase diagrams has been obtained recently using Monte Carlo simulation methods.[4] The phase diagram for $D+1 = 3$ dimensions is shown schematically in Fig. 2. Here the lines of second-order

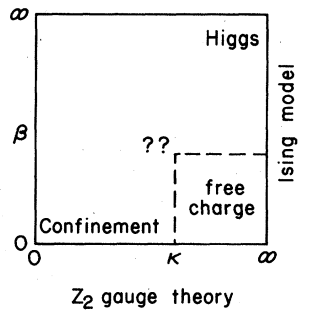

FIG. 1. Theoretical expectations for the phase diagram of Ising lattice gauge theory with matter fields in $D+1 = 3$ dimensions.

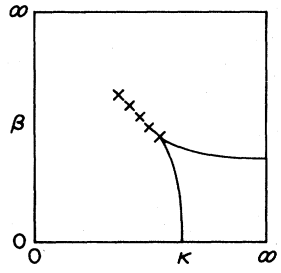

FIG. 2. Monte Carlo phase diagram in $D+1 = 3$ dimensions. The $\times$'s indicate a line of first-order transitions and the solid lines inside the phase diagram denote second-order transitions.

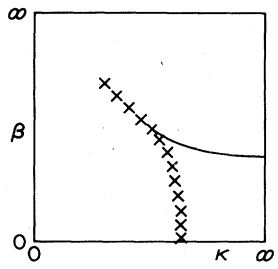

FIG. 3. Same as Fig. 2 but in $D+1=4$ dimensions.

transitions enclose the region of free charge. In addition a line of first-order transitions emerges from their meeting point and extends a finite distance into the phase diagram before terminating. As will be discussed below, this model is mapped unto itself under a duality transformation and the line of first-order transitions lies on the self-dual line. In four dimensions[3] there is again a line of first-order transitions, but now it extends to the pure gauge-field transition as shown in Fig. 3.

One aim of this article is to shed light on the origin of these lines of first-order transitions. We will concentrate on the three-dimensional models and will use a Hamiltonian formulation of the problem. We seek a method of analysis which is equally reliable in the interior of the phase diagram as on its boundaries. To begin we generalize the model Eq. (2.1) to an N-state Potts system. Write Eq. (2.3) in a more suggestive form,

$$S=\beta\sum_{\vec{r},\mu}\delta_{n_\mu,0}+\kappa\sum_{\vec{r},\mu\nu}\delta_{m_{\mu\nu},0}\,, \tag{2.6}$$

where

$$\begin{aligned} n_\mu(\vec{r})&=n(\vec{r})+m_\mu(\vec{r})+n(\vec{r}+\hat{e}_\mu)\,,\\ m_{\mu\nu}(\vec{r})&=m_\mu(\vec{r})+m_\nu(\vec{r}+\hat{e}_\mu)\\ &\quad+m_{-\nu}(\vec{r}+\hat{e}_\mu+\hat{e}_\nu)+m_{-\mu}(\vec{r}+\hat{e}_\nu)\,, \end{aligned} \tag{2.7}$$

and $n(\vec{r})$ is an Ising site variable ($n(\vec{r})=0,1$), $m_\mu(\vec{r})$ is an Ising link variable ($m_\mu(\vec{r})=0,1$), and the Kronecker symbols are taken mod (2). The N-state model we are interested in uses Eqs. (2.6) and (2.7), but generalizes the variables $n(\vec{r})$ and $m_\mu(\vec{r})$ to take N possible values,

$$\begin{aligned} n(\vec{r})&=0,1,2,\ldots,N-1\,,\\ m_\mu(\vec{r})&=0,1,2,\ldots,N-1\,. \end{aligned} \tag{2.8}$$

These models are a natural generalization of Ising lattice gauge theory in the sense that for every N the matter field lies in the fundamental representation of the gauge group Z_N [the local invariance of the N-state model is the obvious generalization of Eq. (2.4)] and, as our analysis will show, the phase diagrams of the N-state models change smoothly as N varies from one to infinity. The generalization is useful because the $N\to\infty$ limit is soluble and $1/N$ corrections are easily computed and reveal features of the Ising case throughout the entire (β,κ) plane.

It is a straightforward exercise in transfer-matrix techniques to obtain the Hamiltonian for the N-state models. Recall that for the $N=2$ Ising case,[8]

$$\begin{aligned} H=&-\lambda\sum_l\sigma_1-\frac{1}{\lambda}\sum_p\sigma_3\sigma_3\sigma_3\sigma_3\\ &-\frac{1}{x}\sum_i\tau_1-x\sum_l\tau_3\sigma_3\tau_3\,, \end{aligned} \tag{2.9}$$

where $\sigma_{1,3}$ are Pauli matrices defined on links l, $\tau_{1,3}$ are Pauli matrices defined on sites i, and p indicates a plaquette. In $D=2$ spatial dimensions H maps onto itself under a duality transformation,[1,2,8]

$$\begin{aligned} &\sigma_1\to\tilde{\tau}_3\tilde{\sigma}_3\tilde{\tau}_3\,,\\ &\sigma_3\sigma_3\sigma_3\sigma_3\to\tilde{\tau}_1\,,\\ &\tau_1\to\tilde{\sigma}_3\tilde{\sigma}_3\tilde{\sigma}_3\tilde{\sigma}_3\,,\\ &\tau_3\sigma_3\tau_3\to\tilde{\sigma}_1\,, \end{aligned} \tag{2.10}$$

where the tilde variables are again Pauli matrices. Therefore

$$H(\lambda,x)\to H(x,\lambda)\,. \tag{2.11}$$

The simplicity of Eq. (2.11) is the motivation for the parametrization of the model chosen in Eq. (2.9). Note that the model is self-dual on the line $x=\lambda$. In later analysis it will prove convenient to use the variables

$$\begin{aligned} \xi&=\tfrac{1}{2}(x+\lambda)\,,\\ \eta&=\tfrac{1}{2}(x-\lambda)\,. \end{aligned} \tag{2.12}$$

Then the duality transformation is the mapping $\xi\to\xi$ and $\eta\to-\eta$. Equation (2.11) implies that the ground-state energy E of the model satisfies

$$E(\xi,\eta)=E(\xi,-\eta)\,. \tag{2.13}$$

The phase diagram of the Hamiltonian version of the model then appears as in Fig. 4 where the regions and labels have been borrowed from Fig. 2. Note that the self-dual line lies at 45° on this figure and the duality transformation simply folds the figure along this crease.

The generalization of Eq. (2.9) to the N-state Potts model is straightforward. In fact, it parallels the discussions of Z_N gauge theories with matter fields which have been discussed extensively in the literature.[9] One begins with the Euclidean formulation Eq. (2.6), singles out a τ direction, constructs a transfer matrix with an operator representation, and lets the lattice spacing in the τ

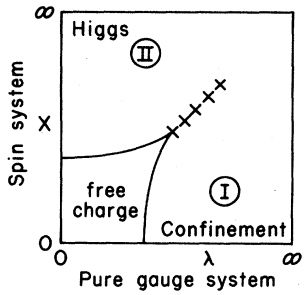

FIG. 4. Hamiltonian phase diagram for Ising lattice gauge theory with matter fields in $D+1=3$ dimensions corresonding to Fig. 2.

direction scale to zero. The resulting Hamiltonian for the general N-state Potts gauge theory with matter fields is

$$H=-\frac{\lambda}{2N}\sum_{l}\sum_{m=1}^{N-1}(R_l^m+\text{H.c.}) -\frac{1}{2N\lambda}\sum_{p}\sum_{m=1}^{N-1}[(CCCC)_{\mu\nu}^m+\text{H.c.}] -\frac{1}{2Nx}\sum_{i}\sum_{m=1}^{N-1}(V_i^m+\text{H.c.}) -\frac{x}{2N}\sum_{i}\sum_{m=1}^{N-1}[(B_i^\dagger C B_j)^m+\text{H.c.}], \tag{2.14}$$

where R and C are link variables describing the gauge fields and V and B are site variables describing the matter fields. On a given link we have a Z_N algebra

$$RC=e^{-i2\pi/N}CR, \quad R^\dagger C=e^{i2\pi/N}CR^\dagger, \quad C^N=R^N=1, \tag{2.15a}$$

and on a given site we have the analogous relations for the matter field operators,

$$VB=e^{-i2\pi/N}BV, \quad V^\dagger B=e^{i2\pi/N}BV^\dagger, \quad V^N=B^N=1. \tag{2.15b}$$

A straightforward generalization of the construction of the duality transformation for the Ising case shows that *each* N-state model satisfies ($D=2$)

$$H(x,\lambda)=H(\lambda,x), \tag{2.16}$$

so the ground-state energy of the theory satisfies

$$E(x,\lambda)=E(\lambda,x) \quad (\text{all } N). \tag{2.17}$$

Therefore, the (x,λ) or the (ξ,η) parametrizations of the models are quite convenient. To describe the state of particular links and sites, one may choose to diagonalize either R or C and V or B. In the basis in which R is diagonal we have

$$R|L\rangle=e^{-i(2\pi/N)L}|L\rangle, \tag{2.18a}$$

where $L=0,1,\ldots,N-1$ and $|L\rangle$ refers to a given link. The Z_N algebra then implies that C is a ladder operator in this space,

$$C|L\rangle=|L+1\rangle. \tag{2.18b}$$

Analogous equations for the site operators V and B can be written down,

$$V|n\rangle=e^{-i(2\pi/N)n}|n\rangle, \quad B|n\rangle=|n+1\rangle. \tag{2.19}$$

The Hamiltonian Eq. (2.14) can be written in a particularly useful form in this basis. Recalling the identity

$$\frac{1}{N}\sum_{m=0}^{N-1}\cos\left(\frac{2\pi}{N}Lm\right)=\delta_{L,0} \tag{2.20}$$

we have

$$H=\lambda\sum_{l}\left(\frac{1}{N}-\delta_{L,0}\right)-\frac{1}{2N\lambda}\sum_{p}\sum_{m=1}^{N-1}[(CCCC)_{\mu\nu}^m+\text{H.c.}] +\frac{1}{x}\sum_{i}\left(\frac{1}{N}-\delta_{n,0}\right)-\frac{x}{2N}\sum_{i}\sum_{m=1}^{N-1}[(B_i^\dagger C B_j)^m+\text{H.c.}]. \tag{2.21}$$

It will also prove useful to sometimes use the basis in which C is diagonal,

$$C|m\rangle=e^{i(2\pi/N)m}|m\rangle, \quad m=0,1,\ldots,N-1 \tag{2.22a}$$

on each link. Then R is a ladder operator on that link,

$$R|m\rangle=|m+1\rangle, \quad |m+N\rangle=|m\rangle. \tag{2.22b}$$

If the matter fields are treated as in Eq. (2.19), the Hamiltonian Eq. (2.14) can be written in this mixed basis as

$$H=\frac{1}{\lambda}\sum_{p}\left(\frac{1}{N}-\delta_{\Sigma m,0}\right)-\frac{\lambda}{2N}\sum_{l}\sum_{p=1}^{N-1}(R_l^p+\text{H.c.}) +\frac{1}{x}\sum_{i}\left(\frac{1}{N}-\delta_{n,0}\right) -\frac{x}{2N}\sum_{i}\sum_{p=1}^{N-1}[(B_i^\dagger e^{-(2\pi/N)m}B_j)^p+\text{H.c.}], \tag{2.23}$$

where $\sum m$ means the sum of the link variables m around a plaquette. Equations (2.21) and (2.23) will be useful for developing $1/N$ expansions.

III. 1/N EXPANSIONS AND PHASE DIAGRAMS

Consider the Hamiltonian Eq. (2.21) in $D=2$ spatial dimensions. Suppose that N is very large and let $\lambda>1$ and $x<1$ so we are working below the self-dual line in region I of Fig. 4. Then the

Kronecker symbols in Eq. (2.21) are the dominant terms and the ground-state energy density is

$$E_{\mathrm{I}} = -(2\lambda + 1/x) + O(1/N) . \tag{3.1}$$

The factor of 2 in Eq. (3.1) records the fact that in $D=2$ spatial dimensions there are twice as many links as plaquettes on a square lattice. In Eq. (3.1) we have anticipated the crucial fact that corrections to the first term are suppressed by powers of N. Note also that the first term in Eq. (3.1) is also the strong-coupling result ($\lambda \gg 1$, $x \ll 1$) that applies for all N. The compatibility of these two expansion methods is important. The $1/N$ expansion is, however, particularly useful for this problem because the coefficients of the expansion can be calculated *exactly* as simple functions of λ and x allowing us to probe the interior of the phase diagram reliably.

Before calculating $1/N$ effects we should obtain the theory's phase diagram in the limiting $N \to \infty$ case. The duality transformation gives us the vacuum energy density in region II of the phase diagram shown in Fig. 4,

$$E_{\mathrm{II}} = -(2x + 1/\lambda) + O(1/N) . \tag{3.2}$$

And finally we must account for the fact that the ground state may be qualitatively different in the small-x, small-λ region of the phase diagram and compute the vacuum energy density which matches onto the small-x or small-λ boundaries of the diagram. The form Eq. (2.23) of the Hamiltonian is convenient for this purpose. Call this region of the phase diagram III as in Fig. 5 and read off Eq. (2.23) that

$$E_{\mathrm{III}} = -(1/\lambda + 1/x) + O(1/N) . \tag{3.3}$$

Equations (3.1)–(3.3) contain some interesting and surprising physics. First consider E_{I} as λ decreases and E_{III} as λ increases. These formulas are exact in the $N \to \infty$ limiting case so we can use them when λ and/or x are of order unity. For given x, E_{III} becomes larger than E_{I} as λ passes through $1/\sqrt{2}$. But the ground-state energy density

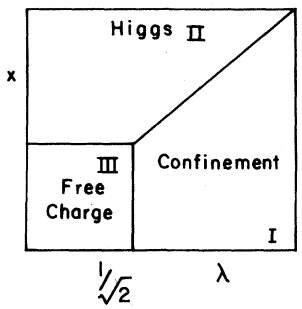

FIG. 5. $N \to \infty$ phase diagram in $D+1=3$ dimensions.

of the model is a continuous, single-valued function so we learn that the ground state changes its character abruptly for all $x < \lambda$ along the line $\lambda^2 = \frac{1}{2}$. This is a line of *first*-order phase transitions because the slope of the ground-state energy changes discontinuously here. And this is a phase boundary between a ground state III which admits free charge and one I which confines. The latent heat across the phase boundary can be defined as

$$\Delta = -\frac{\partial}{\partial\lambda}(E_{\mathrm{I}} - E_{\mathrm{III}}) \tag{3.4a}$$

and we compute

$$\Delta = 4 + O(1/N) \tag{3.4b}$$

across the line $\lambda^2 = \frac{1}{2}$ extending from $x=0$ to $x=\lambda$. The region for $x > \lambda$ is obtained from the duality transformations. We learn that the free-charge region III is *closed* and is surrounded by a line of first-order transitions in the $N \to \infty$ limit as shown in Fig. 5. We will see that the inclusion of $1/N$ corrections provides curvature to the line of transitions between regions I and III.

Next, consider the interface between the Higgs II and confinement I regions. The duality transformation has given E_{II} and it is clear that on the self-dual line $x=\lambda$, E_{I} matches E_{II} in magnitude but *not* slope. It is convenient to use the variables ξ and η as in Eq. (2.13), so that

$$E_{\mathrm{II}}(\xi,\eta) = E_{\mathrm{I}}(\xi,-\eta) , \quad \eta \geq 0 \tag{3.5}$$

and one can consider a latent heat across the self-dual line,

$$\Delta(\xi) = \frac{1}{2}\frac{\partial}{\partial\eta}(E_{\mathrm{I}} - E_{\mathrm{II}})\Big|_{\eta=0} . \tag{3.6}$$

To leading order in $1/N$,

$$\Delta(\xi) = 2 + 1/\xi^2 + O(1/N) , \tag{3.7}$$

so we learn that a line of first-order transitions separates the confining phase from the Higgs phase in the $N \to \infty$ limit. This is shown in Fig. 5. Thus in this limit there is an *intrinsic* difference between the Higgs and confinement mechanisms. Rigorous theorems assure us that for fixed and finite N such a distinction must disappear.[2] We will see below that the $1/N$ expansion predicts that the line of first-order transitions terminates and that regions I and II are in fact analytically connected.

Let us obtain the $1/N$ corrections to Eq. (3.1). First there are the terms we read off the Hamiltonian Eq. (2.21),

$$\frac{1}{N}\left(2\lambda + \frac{1}{x}\right) . \tag{3.8}$$

Next we must consider the nontrivial second and

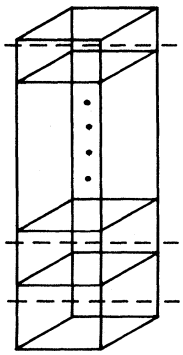

FIG. 6. Perturbation-theory graph contributing to the $1/N$ expansion. The horizontal dashed lines denote energy denominators and the horizantal squares denote the action of the operator $(CCCC)^m_{\mu\nu}$.

fourth terms in the Hamiltonian. Consider that term in Rayleigh-Schrödinger perturbation theory in which the third term acts an arbitrary number of times on a given plaquette as pictured in Fig. 6. The squares in this figure indicate the action of $(CCCC)^m_{\mu\nu}$, and each dashed horizontal line indicates an energy denominator. (Such pictorial representations of Rayleigh-Schrödinger perturbation theory have been discussed extensively elsewhere, so the reader is presumed to be familiar with these rules.) The sum of these graphs is a geometric series

$$\sum_{m=2}^{\infty}\left(-\frac{1}{N\lambda}\right)^m\frac{(N-1)(N-2)^{m-2}}{(-4\lambda)^{m-1}} = -\left(\frac{N-1}{N^2}\right)\left(\frac{\lambda^{-2}}{4\lambda-[(N-2)/N]\lambda^{-1}}\right). \quad (3.9)$$

The factors of N here, $(N-1)(N-2)^{m-2}$, count the possible number of intermediate states in the N-state model. For example, if $\sum_{m=1}^{N-1}(CCCC)^m_{\mu\nu}$ is applied to the $L=0$ zeroth-order vacuum it creates $(N-1)$ intermediate states each raised in energy by 4λ. But if the same operator is applied to the same plaquette a second time it can create a state different from the zeroth-order vacuum from only $N-2$ of its terms, hence, the $(N-2)^{m-2}$ in Eq. (3.9). Note that the leading N dependence in Eq. (3.9) is

$$\frac{1}{N}\left(-\frac{\lambda^{-2}}{4\lambda-\lambda^{-1}}\right)+O(1/N^2). \quad (3.10)$$

The final $1/N$ correction to the vacuum energy occurs when the fourth term in the Hamiltonian applies to a link an arbitrary number of times. This is shown in Fig. 7 and the same methods that gave Eq. (3.9) yield

$$-2\left(\frac{N-1}{N^2}\right)\left(\frac{x^2}{2x^{-1}+\lambda-[(N-2)/N]x}\right), \quad (3.11a)$$

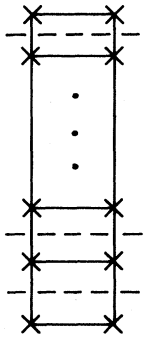

FIG. 7. Pictorial representation of the operator $(B_iCB_j)^m$ contributing to the $1/N$ expansion.

whose leading N dependence is

$$\frac{1}{N}\left(-\frac{2x^2}{2x^{-1}+\lambda-x}\right)+O(1/N^2); \quad (3.11b)$$

collecting everything we have

$$E_I=-\left(2\lambda+\frac{1}{x}\right)+\frac{1}{N}\left[2\lambda+\frac{1}{x}+\frac{1}{\lambda}\left(\frac{1}{1-4\lambda^2}\right)-\frac{2x^2}{2+x\lambda-x^2}\right]+O(1/N^2). \quad (3.12)$$

It is easy to obtain several more terms in this $1/N$ expansion. The detailed calculation of the $1/N^2$ terms involves the graphs Figs. 10–15 and is contained in the Appendix.

Using Eq. (3.12) and the definition Eq. (3.7) for the latent heat on the self-dual line, we can compile the expansion

$$\Delta(\xi)=\Delta^{(0)}(\xi)+\frac{1}{N}\Delta^{(1)}(\xi)+\frac{1}{N^2}\Delta^{(2)}(\xi)+\cdots. \quad (3.13)$$

The coefficients $\Delta^{(i)}(\xi)$, $i=0$, 1, and 2, are recorded for a range of ξ in Table I. They contain

TABLE I. $1/N$ expansion coefficients for the latent heat on the self-dual line.

ξ^2	$\Delta^{(0)}$	$\Delta^{(1)}$	$\Delta^{(2)}$
0.40	4.5	−32.248 88	−129.015 09
0.45	4.222 22	−21.052 5	23.113 01
0.50	4.0	−15.75	29.504 93
0.55	3.818 182	−12.841 39	23.445 00
0.60	3.666 67	−11.098 78	18.126 02
0.65	3.538 46	−9.997 5	14.506 00
0.75	3.333 33	−8.812 5	10.702 16
1.00	3.0	−8.222 22	9.042 96
1.25	2.8	−8.812 5	10.157 11
1.50	2.666 67	−9.87	11.186 23
1.75	2.571 11	−11.201 39	10.944 53
2.00	2.5	−12.734 69	8.276 13
2.50	2.4	−16.293 21	−7.495 23
3.00	2.333 33	−20.429 75	−46.183 40

some interesting physics. The coefficient $\Delta^{(0)}(\xi)$ was discussed in Eq. (3.7). Consider the $1/N$ corrections $\Delta^{(1)}(\xi)$ listed in Table I. $\Delta^{(1)}(\xi)$ receives contributions from Eqs. (3.8), (3.9), and (3.11). Note that it is *negative* definite and tends to diminish the zeroth-order latent heat. In fact, $\Delta^{(1)}(\xi)$ grows large as ξ^2 tends to $\frac{1}{4}$ and as ξ^2 tends to infinity and, thereby, *terminates* the line of first-order transitions at both ends of the phase diagram. These features of $\Delta^{(1)}(\xi)$ are easy to understand. The termination of the first-order line for ξ^2 small follows from Eqs. (3.9) and (3.10) and the related Fig. 6. These equations give the contribution of partially renormalized plaquette excitations to the ground-state wave function. These graphs shift the mass of the bare plaquette excitation from 4λ to $4\lambda-(1-2/N)\lambda^{-1}$. So, to leading order in N and on the self-dual line ($\lambda=x=\xi$), these excitations become tachyons at $\xi^2=\frac{1}{4}$. This is a signal that the real ground state changes its character near this point and *electric-flux condensation* occurs for $\xi^2\lesssim\frac{1}{4}$. This type of vacuum permits the existence of *free charge*.[10] Of course, these are the same conclusions we reached when examining the $N\to\infty$ limit. The shape of the free-charge region can be estimated from Eq. (3.12). In the region of the phase diagram below the self-dual line, the ground state becomes a condensate of electric flux for $\lambda\approx\frac{1}{2}$ independent of x. This phase boundary extends up to the self-dual line where there is an intersection at $\xi\approx 1/\sqrt{2}\approx 0.71$. The shape of the boundary of the free-charge region above the self-dual line is obtained by folding this result over the $x=\lambda$ line. The higher-order $1/N$ corrections give the expected curvature to the phase boundary but do not change it dramatically. Note that although we have not calculated the order of the transition across the boundary, the $1/N$ expansion suggests that its order is the same as that of the theory on the boundary because the mechanism—electric-flux condensation—is the same all along the phase boundary. Thus in $D=2$ spatial dimensions the free-charge region is surrounded by a boundary of second-order transitions. This agrees with the Monte Carlo studies.[4]

Now consider large ξ. In this limit $\Delta^{(1)}(\xi)$ grows as $-\xi^4$,

$$\Delta(\xi)\sim 2-\xi^4/N \quad (\xi^2\gg 1). \tag{3.14}$$

Ignoring higher-order effects in N we see that for fixed but large N, $\Delta(\xi)$ vanishes for

$$N\gtrsim\xi^4. \tag{3.15}$$

This qualitative conclusion is also supported by higher-order calculations which will be discussed below.

Before continuing we should discuss the proper use and interpretation of $1/N$ expansions for latent heats. We have very few rigorous results here to guide us, so we will rely on model calculations. Recall the N-state Potts spin system in $D+1=2$ dimensions. These systems depend on only one variable, temperature, and also map into themselves under a duality transformation. Call the self-dual point $T^*=1$. $1/N$ expansions can be done for their latent heats giving[11]

$$\Delta(T^*=1)=1-\frac{6}{N}+\frac{2}{N^2}+\frac{8}{N^3}+\cdots, \tag{3.16}$$

which has a zero at $N_c\cong 5.5$. It is known rigorously that Δ vanishes identically for $N\leq 4$ but is nonzero otherwise.[12] Thus, the $1/N$ expansion is a rather good guide to the phase structure of this class of models. Note that the third-order polynomial Eq. (3.16) actually becomes negative for $N_c\lesssim 5.5$. In fact, the exact Δ vanishes nonanalytically at $N=4$. The polynomial, of course, misses this fact and the negative values of Δ indicate that one is expanding about the wrong vacuum for $N<4$ (latent heats must be positive for thermodynamic stability). So the expansion is reliable down to that N where Δ vanishes and not beyond. A more sophisticated estimate of N_c might allow for the possibility that $\Delta(T^*=1)$ vanishes with a power as N approaches N_c from above,

$$\Delta(T^*=1)\sim(N-N_c)^{\gamma}, \quad N\geq N_c. \tag{3.17}$$

The logarithmic derivative of Δ would then have a simple pole

$$\Delta'/\Delta\sim\gamma/(N-N_c). \tag{3.18}$$

The $1/N$ expansion for the logarithmic derivative of Δ reads

$$\frac{1}{\Delta}\frac{d\Delta}{d(1/N)}\sim -6\left(1+\frac{16/3}{N}+\frac{26}{N^2}\right), \tag{3.19}$$

which can be fitted to a simple pole using a [1/1] Padé approximate,

$$\frac{1}{\Delta}\frac{d\Delta}{d(1/N)}\sim -6\left(\frac{1+\frac{11}{24}(1/N)}{1-\frac{39}{8}(1/N)}\right), \tag{3.20}$$

which has a pole at

$$N_c=4.875 \tag{3.21}$$

which is a slight improvement over the result read off the polynomial. Actually it is known that Δ vanishes with an essential singularity as $N\to 4$ from above, so even Eq. (3.21) can be improved. Of course, with so few terms in the $1/N$ expansion it is not very sensible to face the full sophistication of the problem. Anyway, we learn from this example that the $1/N$ expansion is a good guide to the phase structure of these models and is useful even for relatively small values of N when the

TABLE II. [1/1] Padé approximates for the expansion in Table I.

ξ^2	$\Delta(\xi)$
0.40	$4.5\left(\frac{1-11.167\,03/N}{1-4.000\,61/N}\right)$
0.45	$4.222\,22\left(\frac{1-3.888\,24/N}{1+1.097\,88/N}\right)$
0.50	$4\left(\frac{1-2.064\,17/N}{1+1.873\,33/N}\right)$
0.55	$3.818\,18\left(\frac{1-1.537\,41/N}{1+1.825\,81/N}\right)$
0.60	$3.666\,67\left(\frac{1-1.537\,41/N}{1+1.633\,16/N}\right)$
0.65	$3.538\,88\left(\frac{1-1.374\,42/N}{1+1.450\,96/N}\right)$
0.75	$3.333\,33\left(\frac{1-1.429\,32/N}{1+1.214\,43/N}\right)$
1.00	$3\left(\frac{1-1.640\,82/N}{1+1.099\,85/N}\right)$
1.25	$2.8\left(\frac{1-1.994\,74/N}{1+1.152\,58/N}\right)$
1.50	$2.666\,67\left(\frac{1-2.567\,90/N}{1+1.133\,36/N}\right)$
1.75	$2.571\,11\left(\frac{1-3.379\,56/N}{1+0.977\,07/N}\right)$
2.00	$2.5\left(\frac{1-4.444\,00/N}{1+0.649\,00/N}\right)$
2.50	$2.4\left(\frac{1-7.248\,86/N}{1-0.460\,02/N}\right)$
3.00	$2.333\,33\left(\frac{1-11.016\,20/N}{1-2.260\,60/N}\right)$

latent heat vanishes. Studies of N-state gauge models have proved equally successful.[13] With this experience we will use our results for the gauge system with matter fields for all N.

Consider Table I again. Imagine letting N decrease for various fixed values of ξ. We are interested in the possibility that a finite line of first-order transitions persists on the self-dual line even when N is 2 or 3, say. One can either look for zeros of the $1/N$ expansion treated as a polynomial in $1/N$, or one can replace the polynomial by a [1/1] Padé approximate and search for zeros of its numerator. The approximates are recorded in Table II. Fortunately both treatments of the $1/N$ expansion give the same semiquantitative results.

Setting $N=2$ in Table II, we see that the Ising gauge theory with matter fields is predicted to have a line of first-order transitions on its self-dual line from

$$\xi^2 \approx 0.5 \text{ to } \xi^2 \approx 1.25 . \tag{3.22}$$

From strong-coupling expansions of the Hamiltonian formulation of the Ising model we also know that there are critical points on the boundaries of the phase diagram at[14]

$$\begin{aligned} x_c &= 0.405, \quad \lambda = 0 \\ \lambda_c &= 0.405, \quad x = 0 . \end{aligned} \tag{3.23}$$

Therefore, using the results from the $1/N$ expansion and the other rigorous results known about this model one is led to draw the phase diagram in Fig. 8. Note the beautiful consistency between Eqs. (3.23) and (3.22)—the beginning of the line of first-order transitions begins in the correct region of the phase diagram to smoothly connect onto the second-order transitions on the boundaries. Equation (3.13) also agrees quite well with the estimate $\lambda \approx \frac{1}{2}$ obtained from the flux condensation idea.

It is interesting to pinpoint those graphs in the calculation of Δ which cause the termination of the line of first-order transitions as ξ increases. At $O(1/N^2)$ the graphs (A4) and (A8) of the Appendix contribute negative terms to $\Delta^{(2)}(\xi)$ which dominate when ξ^2 becomes large and cause the latent heat to vanish. These graphs are connected clusters which survive in the percolation $N \to 1$ limit of the family of models. Note from Table II that all members of the family of models are predicted to have a region of first-order transitions on the self-dual line except for the case $N \lesssim 1$. Thus it is tempting to speculate that in the $N \to 1$ limit the first-order transitions are missing, but as N grows a line of such transitions begin to grow into the phase diagram from the point where the lines of critical points meet on the self-dual line. For $N > 2$, the lines of critical points which reach the

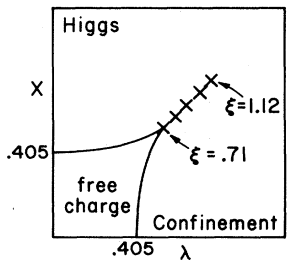

FIG. 8. Numerical $1/N$ results for the phase diagram of Ising lattice gauge theory with matter fields in $D+1=3$ dimensions.

boundaries most likely become first-order lines because the pure Potts spin systems have first-order transition points in two spatial dimensions.

IV. DISCUSSION

The analysis presented here emphasized the $1/N$ expansions for the family of N-state models in $D+1=3$ dimensions. The simplicity of the duality transformation made this analysis particularly easy. The generalization of our calculations to $D+1=4$ dimensions is straightforward. In the $N\to\infty$ limit, we calculate using Eqs. (2.21) and (2.23) the ground-state energy densities in the three regions.

$$E_{\mathrm{I}} = -(3\lambda + 1/x) + O(1/N),$$
$$E_{\mathrm{II}} = -(3/\lambda + 3x) + O(1/N), \tag{4.1}$$
$$E_{\mathrm{III}} = -(3/\lambda + 1/x) + O(1/N).$$

Comparing E_{I} and E_{III} we find a line of first-order transitions at $\lambda=1$ separating these phases. Comparing E_{II} and E_{III} we find a boundary at $x=1/\sqrt{3}$. And finally, comparing E_{I} and E_{II} we find a boundary between these two phases,

$$x = \frac{1}{6\lambda}[3(\lambda^2-1) + (9\lambda^4 - 6\lambda^2 + 9)^{1/2}], \quad \lambda \geq 1. \tag{4.2}$$

The resulting phase diagram is shown in Fig. 9. It is easy to calculate the $1/N$ corrections to the latent heats along these lines and find that the boundary between the confining and Higgs phases terminates due to percolation graphs.

Although the $1/N$ expansion exposed a pocket of free charge in the phase diagram, the calculations presented here were not powerful enough to determine the order of transition across the boundary if N is small. A more elaborate calculation which accurately determines the position of the phase boundary is necessary for that. It appeared from the few graphs we did calculate that the order of the transition does not change as the line leaves an edge of the phase diagram. $1/N$ expansions for the ground-state energy inside the free-charge pocket could be used with Eq. (3.12) to search for a latent heat on the free-charge phase boundary.

Since the confining and Higgs phases of these models are separated in the $N\to\infty$ limit, one should be able to find an order parameter to label them. However, when N is finite this order parameter should become trivial since then the phases are analytically connected.

The calculation of the latent heat across the self-dual line in the $D=2$ family of models suggested that only the $N\to 1$ limit of the models has no first-order transitions. Since the $N\to 1$ limit reduces the model to a mixed percolation model, this estimate may in fact be quite precise. (The percolation limit of lattice gauge theories will be discussed elsewhere.[15]) It is interesting to consider how the phase diagrams of the $D=2$ family of models changes as N increases. In the $N\to 1$ limit the pocket of free charge would be surrounded by a line of second-order transitions. As N increases above unity, a line of first-order transitions begins to develop along the self-dual line beginning at the boundary of the free-charge region. As N passes through 2 the boundary lines surrounding the free-charge region become first order. As N tends to infinity the line of first-order transitions on the self-dual line extends into the phase diagram without bound (its length grows as $|\xi| \sim N^{1/4}$). A similar pattern should occur in $D=3$ except that the phase boundary between the free-charge and confining regions of the phase diagram

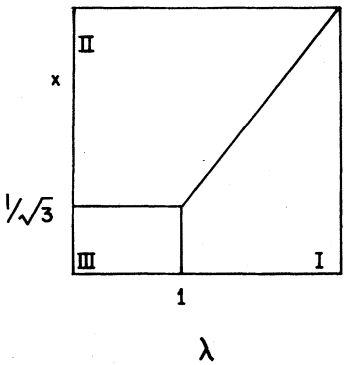

FIG. 9. $N\to\infty$ phase diagram in $D+1=4$ dimensions.

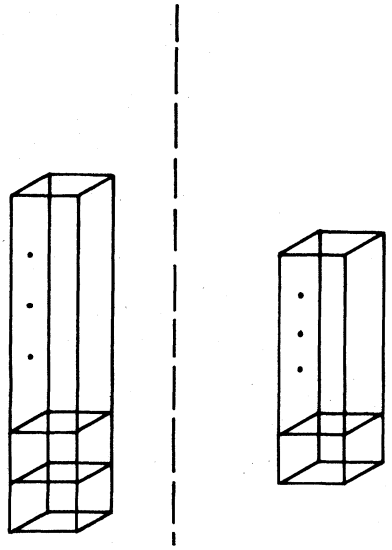

FIG. 10. Graphs contributing to $O(1/N^2)$ and consisting of two disconnected towers of $(CCCC)^m_{\mu\nu}$ operators.

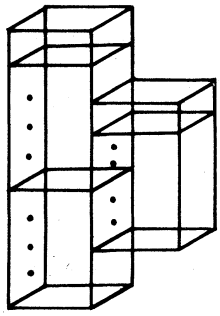

FIG. 11. Graphs as in Fig. 10 but each plaquette has a link in common.

should become first order as N increases above unity. We believe that this is so because $1/N$ expansions for pure gauge Potts lattice theories indicate that they experience first-order transitions for all N greater than one.[13,15]

We have used the $1/N$ expansion very boldly here. It would be useful if some mathematical properties of the expansion could be determined. Does it have a finite radius of convergence around $1/N=0$? Does the latent heat vanish with an essential singularity as N decreases to a critical value?

And finally it would be interesting to carry out the $1/N$ expansions for these models using the Euclidean formulation. Then detailed numerical comparisons could be made with the Monte Carlo data[4] which inspired the present investigation.

ACKNOWLEDGMENTS

The author thanks E. Fradkin, S. Shenker, J. Stack, and J. Shigemitsu for discussions. The solubility of the $N\to\infty$ limit of N-state Potts lattice gauge theories was derived independently by S. Shenker (unpublished). This work was supported in part by the National Science Foundation under Grant No. NSF-PHY79-00272.

APPENDIX

We will sketch the calculation of the $O(1/N^2)$ terms of the ground-state energy and record the

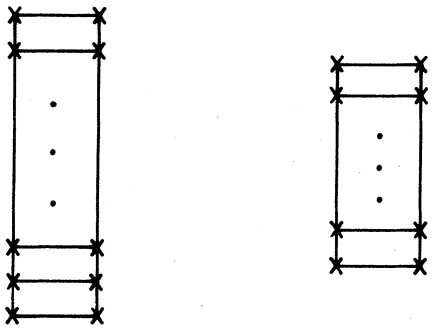

FIG. 12. Graphs contributing to $O(1/N^2)$ and consisting of two disconnected towers of $(B_iCB_j)^m$ operators.

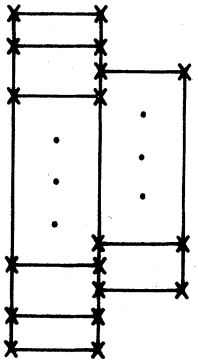

FIG. 13. Graphs as in Fig. 12 but the link operators have a site in common.

results here. First there are the $1/N^2$ pieces of the graphs shown in Figs. 6 and 7. From Eqs. (3.9) and (3.11a) we have

$$(\text{Fig. } 6)=\frac{1}{N^2}\left[\frac{1}{\lambda(4\lambda^2-1)}+\frac{2}{\lambda(4\lambda^2-1)^2}\right]+\cdots, \tag{A1}$$

$$(\text{Fig. } 7)=\frac{1}{N^2}\left[\frac{2x^2}{2x+\lambda-x}+\frac{4x^3}{(2x+\lambda-x)^2}\right]. \tag{A2}$$

Next there are graphs whose leading N dependence is just $1/N^2$. Since Figs. 6 and 7 contribute $O(1/N)$ + corrections, such graphs arise when the perturbation terms of H act on different plaquettes or links of the lattice. For example, the graphs of Fig. 10 are clearly $O(1/N^2)$+ higher-order corrections. We compute

$$(\text{Fig. } 10)=\frac{1}{N^2}\left[\frac{5}{\lambda(4\lambda^2-1)^3}\right]+\cdots, \tag{A3}$$

where the factor of 5 counts the number of plaquettes excluded by another plaquette (the dashed

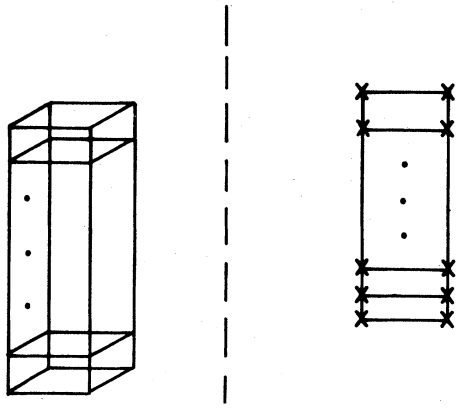

FIG. 14. Graphs contributing to $O(1/N^2)$ and consisting of two disconnected towers of $(CCCC)^m_{\mu\nu}$ operators and $(B_iCB_j)^m$ operators.

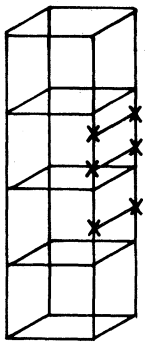

FIG. 15. Graphs as in Fig. 14 but the operators have a link in common.

vertical line indicates that the two plaquettes have no links in common). Next there are the graphs in which the two plaquettes have a single link in common

$$(\text{Fig. } 11)=\frac{1}{N^2}\left[\frac{-8}{\lambda(4\lambda^2-1)^2(7\lambda^2-2)}\right]. \tag{A4}$$

This result is obtained by summing a geometric series as in the derivation of Eq. (3.9). The remaining graphs, Figs. 12–15, give contributions

$$(\text{Fig. } 12)=\frac{1}{N^2}\left[\frac{14x^7}{(2+x\lambda-x^2)^3}\right], \tag{A5}$$

$$(\text{Fig. } 13)=\frac{1}{N^2}\left[\frac{-24x^7}{(2+x\lambda-x^2)^2(3+2x\lambda-2x^2)}\right], \tag{A6}$$

$$(\text{Fig. } 14)=\frac{1}{N^2}\frac{x^2}{\lambda^2}\left[\frac{4}{(4\lambda-\lambda^{-1})^2(2/x+\lambda-x)}+\frac{4}{(4\lambda-\lambda^{-1})(2/x+\lambda-x)^2}\right], \tag{A7}$$

$$(\text{Fig. } 15)=\frac{1}{N^2}\left[\frac{-4x^2}{\lambda^2(4\lambda+2/x-\lambda^{-1}-x)}\times\left(\frac{1}{4\lambda-\lambda^{-1}}+\frac{1}{2/x+\lambda-x}\right)^2\right]. \tag{A8}$$

Collecting these results gives $E_I(x,\lambda)$ to second order in $1/N$. Differentiating with respect to η then gives the expansion for $\Delta(\xi)$ Eq. (3.13). The coefficients are listed in Table I. Aside from the discussion in Sec. III of the text, the detailed formulas are not very illuminating.

[1] F. Wegner, J. Math. Phys. 12, 2259 (1971).
[2] E. Fradkin and S. Shenker, Phys. Rev. D 19, 3682 (1979).
[3] M. Creutz, Phys. Rev. D 21, 1006 (1980).
[4] J. Stack and G. Jongeward, University of Illinois report (unpublished).
[5] D will refer to the spatial dimensions when we construct a transfer matrix.
[6] M. Creutz, C. Rebbi, and L. Jacobs, Phys. Rev. Lett. 42, 1390 (1979).
[7] R. Balian, J. M. Drouffe, and C. Itzykson, Phys. Rev. D 11, 2098 (1975).
[8] D. Horn and S. Yankielowicz, Phys. Lett. 85B, 347 (1979). The speculative phase diagram suggested in this article has been ruled out by the calculation of Ref. 4.
[9] 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 (1980).
[10] An elementary discussion and review of related phenomena can be found in J. Kogut, Rev. Mod. Phys. 51, 659 (1979).
[11] R. Pearson, J. Shigemitsu, and J. Kogut, Institute for Theoretical Physics, Santa Barbara report (unpublished).
[12] R. Baxter, J. Phys. C 6, L 445 (1973).
[13] R. Pearson, J. Shigemitsu, D. K. Sinclair, and J. Kogut, University of Illinois report (unpublished).
[14] R. Elliot and P. Pfeuty, J. Phys. C 4, 2370 (1971).
[15] D. K. Sinclair, J. Shigemitsu, and J. Kogut (unpublished).