PHYSICAL REVIEW D VOLUME 21, NUMBER 2 15 JANUARY 1980

Possible third-order phase transition in the large-N lattice gauge theory

David J. Gross
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08540

Edward Witten
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138

(Received 10 July 1979)

The large-N limit of the two-dimensional U(N) (Wilson) lattice gauge theory is explicitly evaluated for all fixed $\lambda = g^2N$ by steepest-descent methods. The λ dependence is discussed and a third-order phase transition, at $\lambda = 2$, is discovered. The possible existence of such a weak- to strong-coupling third-order phase transition in the large-N four-dimensional lattice gauge theory is suggested, and its meaning and implications are discussed.

I. INTRODUCTION

One of the useful approaches to the study of the large-scale structure of non-Abelian gauge theories, and in particular quantum chromodynamics (QCD), is via the lattice formulation introduced by Wilson.[1] This approach is especially useful if the lattice coupling is sufficiently large so that one can employ the strong-coupling expansion. This expansion[1] provides, in an easily calculable fashion, a simple picture of the large-scale structure of QCD including the phenomenon of confinement. Although this picture is reasonable if, as we expect, infrared slavery drives the effective coupling in QCD to large values at some appropriate scale, in order to establish its connection with the continuum theory one must control the behavior of the lattice theory for arbitrarily small lattice couplings. The issue of whether QCD confines or not is equivalent, in the lattice formulation of the theory, to whether there exists a "phase transition" as one decreases the coupling (which plays the role of temperature in the analogous statistical-mechanics problem) from large to small values. The *absence* of a second-order phase transition would establish that QCD is both asymptotically free at short distances (weak coupling) and confining at large distances (strong coupling).

Another approach to QCD has been to exploit the only free "parameter" of a pure gauge theory, namely the order N of the gauge group. As shown by 't Hooft[2] there is a remarkable simplification of the perturbative expansion of QCD in the "large-N limit," i.e., $N \to \infty$ for fixed $\lambda \equiv g^2N$. The resulting topological structure of the surviving Feynman diagrams in the large-N limit—namely that of planar surfaces in index space—led 't Hooft to suggest that in such a limit one would recover a string model of hadrons. Recently it has been argued that such a string model might emerge in the lattice formulation of QCD either when one performs the large-N limit for sufficiently large values of λ,[3] or when the dimension of space-time is large enough.[4]

In this paper we shall solve for the behavior of U(N) lattice gauge theories in two space-time dimensions as $N \to \infty$ for fixed $\lambda \equiv g^2N$. The choice of U(N) instead of SU(N) is irrelevant in the $N \to \infty$ limit. The choice of two dimensions is dictated by our inability to solve the four-dimensional theory. Two-dimensional gauge theories are, of course, somewhat trivial. There are no transverse degrees of freedom in two dimensions, and the evaluation of most observables in the pure lattice gauge theory can be reduced to finite-dimensional integrals. The two-dimensional theory clearly confines for both weak and strong coupling (the Coulomb potential is linear in two dimensions), and there is no possibility of a second-order phase transition. On the other hand, the very simplicity of the theory allows one to calculate many interesting observables (the vacuum energy, the expectation value of a Wilson loop operator) as explicit functions of λ for $N = \infty$. What is perhaps surprising is that even in two dimensions the dependence on λ is highly nontrivial—in fact, we find a third-order phase transition for a finite value of λ.

It was shown by Brezin, Itzykson, Parisi, and Zuber[5] that the functional integrals in the large-N limit can be calculated by steepest-descent methods. In this limit, a particular configuration totally dominates the functional integral. In the lattice gauge theory we find that the elementary plaquette variables, which are unitary matrices W, have a determined distribution of eigenvalues $e^{i\alpha}$ at the $N = \infty$ saddle point. For weak coupling, $\lambda = g^2N$, one would expect this distribution to be peaked about $\alpha = 0$, corresponding to $W \approx 1$. We find that the eigenvalues are not only peaked about $\alpha = 0$, but restricted to lie within a finite domain:

$|\alpha| \le 2\sin^{-1}(\frac{1}{2}\lambda)$, for all $\lambda < 2$. As λ increases, this domain increases in size, and for $\lambda = 2$, $e^{i\alpha}$ can range over the whole unit circle. At this point there occurs a phase transition, i.e., the observables of the theory are described by different functions of λ for $\lambda \ge 2$ and for $\lambda \le 2$, which are not analytic continuations of each other. For large λ the distribution of eigenvalues becomes increasingly uniform, corresponding, as $\lambda \to \infty$, to a totally random [with respect to the Haar measure on U(N)] unitary matrix.

The phase transition at $\lambda = 2$ turns out to be of third order—the β function does not vanish (in fact, is always negative) but has a kink at $\lambda = 2$, and the derivative of the specific heat is discontinuous at $\lambda = 2$. We speculate on the existence of such a weak to strong third-order phase transition in the $N = \infty$ four-dimensional theory, and argue that *if* a noninteracting string picture emerges by interchanging the $N \to \infty$ limit with the strong-coupling expansion, then a phase transition *must* occur for $\lambda \ge 1$. The existence of such a phase transition has implications, as will be discussed below, for the problem of interpolating, even for finite N, between weak and strong coupling.

In Sec. II we review the structure of two-dimensional U(N) lattice gauge theories. In Sec. III we apply the method of steepest descent to evaluate explicitly the large-N limit of the theory, and calculate as a function of λ many physically interesting observables. Finally, in Sec. IV, we discuss some of the aspects of the third-order phase transition and present some speculations about four-dimensional QCD for large N.

II. SU(N) GAUGE THEORY IN TWO DIMENSIONS

Pure gauge theories in two space-time dimensions are essentially trivial. This is due to the lack of "transverse" dimensions and the resulting absence of physical gluons. The trivial nature of two-dimensional gauge theories manifests itself in the fact that one can reduce the calculation of the relevant physical quantities, i.e., the vacuum-to-vacuum amplitude, the vacuum expectation value of the Wilson loop operator, the β function, etc., to the evaluation of simple integrals. Indeed, having solved the gauge theory in a world consisting of but one plaquette one has essentially solved the full two-dimensional gauge theory.

Let us consider the lattice formulation of the two-dimensional U(N) gauge theory, as formulated originally by Wilson.[1] The dynamical variables are unitary matrices $U_{\vec{n},\vec{i}}$, associated with links on the lattice, where $a\vec{n}$ is a lattice site ($\vec{n} = n_0\vec{i}_0 + n_1\vec{i}_1$), $\vec{i}$ is one of the lattice vectors $\vec{i}_0$ or $\vec{i}_1$, and a is the lattice spacing. $U_{\vec{n},\vec{i}}$ is a unitary matrix in the fundamental, N-dimensional, representation of U(N) that parallel transports a matter field [in the N-dimensional representation of U(N)] from site $\vec{n}$ to $\vec{n}+\vec{i}$, and

$$(U_{\vec{n},\vec{i}})^\dagger = U_{\vec{n}+\vec{i},-\vec{i}} = (U_{\vec{n},\vec{i}})^{-1}. \tag{1}$$

The Wilson action is defined to be

$$S(U) \equiv \sum_P \frac{1}{g^2}\,\mathrm{Tr}\left(\prod_P U + \mathrm{H.c.}\right), \tag{2}$$

where the sum runs over all plaquettes (squares) on the lattice and

$$\prod_P U = U_{\vec{n},\vec{i}_0} U_{\vec{n}+\vec{i}_0,\vec{i}_1} U_{\vec{n}+\vec{i}_0+\vec{i}_1,-\vec{i}_0} U_{\vec{n}+\vec{i}_1,-\vec{i}_1}.$$

Ground-state expectation values of physical observables, i.e., functions of the U's, are given by

$$\langle O(U)\rangle = \frac{1}{Z}\int [DU]\exp[S(U)]O(U), \tag{3}$$

where Z is the vacuum-to-vacuum amplitude

$$Z = \int [DU]\exp[S(U)]$$

and the measure is $DU \equiv \Pi_{\vec{n},\vec{i}}\, dU_{\vec{n},\vec{i}}$, where $dU_{\vec{n},\vec{i}}$ is the Haar measure on the group U(N) which satisfies

$$DU = D(UV) = D(VU),$$

where V is an arbitrary unitary matrix, and which we normalize so that $\int dU_{\vec{n},\vec{i}} = 1$.

Some quantities of physical interest are the "free energy" F, which is proportional to the vacuum energy density E_0,

$$E_0 = \frac{F(g^2,N)}{g^2N} \equiv -\frac{1}{V}\ln Z, \tag{4}$$

where V = volume of the two-dimensional world, and g^2N plays the role of the temperature ($g^2N \equiv kT$) and the expectation value of the Wilson loop operator

$$W_L(g^2N) = \frac{1}{N}\left\langle \mathrm{Tr}\left[\prod_L U\right]\right\rangle, \tag{5}$$

where $\Pi_L\, U$ is an ordered product of U's on the links belonging to a closed loop L. In particular, for a rectangular loop of time extent Ta and spatial extent Ra, $W_L(g^2N)$ is related to the interaction energy $\epsilon(R)$ of static sources separated by distance Ra:

$$\epsilon(R) = -\frac{1}{Ta}\lim_{T\to\infty} \ln W_{L(T,R)}(g^2N). \tag{6}$$

The calculation of F or W_L can be reduced to a single dU integration by exploiting the gauge invariance of the theory, i.e., the invariance under

$$U_{\vec{n},\vec{i}} \to V_{\vec{n}} U_{\vec{n},\vec{i}} V^\dagger_{\vec{n}+\vec{i}} \tag{7}$$

for arbitrary unitary matrices $V_{\vec{n}}$. We then have the option of making a gauge choice. A convenient gauge is the one in which $U_{\vec{n},\vec{i}_0}=1$ for all $\vec{n}$, i.e., the $A_0=0$ gauge. In this gauge

$$S(U)=\frac{1}{g^2}\sum_{\vec{n}} \mathrm{Tr}(U_{\vec{n},\vec{i}_1}U^\dagger_{\vec{n}+\vec{i}_0,\vec{i}_1}+\text{H.c.}).$$

Z can then be easily evaluated by the change of variables

$$U_{\vec{n}+\vec{i}_0,\vec{i}_1} \equiv W_{\vec{n}} U_{\vec{n},\vec{i}_1}$$

so that

$$Z=\int \prod_{\vec{n}} [dW_{\vec{n}}] \exp\left[\sum_n \frac{1}{g^2}\mathrm{Tr}(W_n+W_n^\dagger)\right] = (z)^{V/a^2},$$

where V is the "volume" of our world (with free boundary conditions), V/a^2 is the number of plaquettes, and

$$z(g^2,N)=\int [dW]\exp\left[\frac{1}{g^2}\mathrm{Tr}(W+W^\dagger)\right], \qquad F/g^2N=-\frac{1}{a^2}\ln z. \tag{8}$$

Similarly, the Wilson loop operator can be written

$$W_L(g^2,N)=\frac{1}{N}\left\langle \mathrm{Tr}\prod_{k=0}^{R-1} U_{T\vec{i}_0+k\vec{i}_1,\vec{i}_1}\right\rangle = \frac{1}{N}\left\langle \mathrm{Tr}\prod_{k=0}^{R-1}\prod_{j=T-1}^{0} W_{j\vec{i}_0+k\vec{i}_1}\right\rangle, \tag{9}$$

where we have made the further gauge choice $(U_{0\vec{i}_0+n_1\vec{i}_0}=I)$. Since the action only depends on $\mathrm{Tr}(W_n+W_n^\dagger)$, it is invariant under $W_n \to VW_nV^\dagger$ for each W_n separately. Therefore, consider the integration over a particular $W_{\vec{n}}$ appearing in Eq. (9),

$$\frac{1}{N}\int dW_n \exp[(1/g^2)\,\mathrm{Tr}(W_n+W_n^\dagger)]\,\mathrm{Tr}(AW_nB)=\frac{1}{N}\int dV\int dW_n\exp[(1/g^2)\,\mathrm{Tr}(W_n+W_n^\dagger)]\,\mathrm{Tr}(AVW_nV^\dagger B),$$

where A and B represent the remaining products of W_n's. Using the fact that

$$\int dV\, V_{ij}V^\dagger_{kl}=\frac{1}{N}\delta_{il}\delta_{jk}, \tag{10}$$

it follows that $\mathrm{Tr}(AW_nB)$ can be replaced by $(1/N)\,\mathrm{Tr}(W_n)\,\mathrm{Tr}(BA)$. Iteration of this argument then yields

$$W_L(g^2,N)=[w(g^2,N)]^{RT}, \tag{11}$$

where w is the Wilson loop operator for a single plaquette

$$w(g^2,N)=\frac{1}{z}\int dW\frac{1}{N}\mathrm{Tr}W\exp[(1/g^2)\,\mathrm{Tr}(W+W^\dagger)]. \tag{12}$$

Correspondingly,

$$\epsilon(R)=-\frac{R}{a}\ln w(g^2,N). \tag{13}$$

Thus the two-dimensional gauge theory is reduced to a single integral, Eq. (8), which characterizes the one-plaquette world. The resulting physics contains no surprises. Since $w(g^2,N)=w^*(g^2,N)\le 1$, it follows that the string tension $\sigma\equiv\epsilon(R)/Ra=-(1/a^2)\ln w(g^2,N)$ is always positive and one is in the confining phase for all values of the coupling.

III. THE LARGE-N LIMIT

For any given N, the free energy F and the Wilson loop operator w can be explicitly calculated. In fact, Bars and Green[3] have derived an explicit expression for $z(g^2,N)$,

$$z(g^2,N)=\det \underline{M}, \qquad (\underline{M})_{i,j}=I_{i-j}(2/g^2),\quad i,j=1,\ldots,N. \tag{14}$$

This follows by noting that the integrand in Eq. (8) only depends on the eigenvalues α_i, $i=1,\ldots,N$, of W, and that W can be written as $W=TDT^\dagger$ where $D_{ij}=\delta_{ij}e^{i\alpha_i}$ and T is unitary. Furthermore, $dW=\text{const}\times dT\prod_{i=1}^N d\alpha_i\Delta^2(\alpha_i)$ where

$$\Delta^2(\alpha_i)=\prod_{i<j}\sin^2\left|\frac{\alpha_i-\alpha_j}{2}\right|=4^{-N}|\det\Delta|^2, \qquad (\Delta)_{j,k}=\exp[i(j\alpha_k)]. \tag{15}$$

Thus the evaluation of z reduces to the integral

$$z(g^2,N)=\text{const}\int_0^{2\pi}\prod d\alpha_i\Delta^2(\alpha_i)\times\exp\left(\frac{2}{g^2}\sum_{i=1}^N\cos\alpha_i\right), \tag{16}$$

where the constant is chosen so that $z(\infty,N)=1$. Expanding the determinant, one can evaluate all the integrals in terms of modified Bessel functions of the first kind and reexpress z as the above Hadamard determinant.

We are interested in the large-N limit, i.e., $N\to\infty$ for fixed $\lambda=g^2N$. The above expression is rather unwieldy for large N. We shall therefore adapt the method employed by Brezin, Itzykson, Parisi, and Zuber in their analysis of the large-N anharmonic oscillator.[5] In the large-N limit the steepest-descent method can be employed, and to leading order in $1/N$ we have

$$-E_0(\lambda)\equiv\lim_{N\to\infty}\frac{\ln z(g^2,N)}{N^2}=\lim_{N\to\infty}\frac{1}{N^2}\left(\frac{2}{g^2}\sum_{i=1}^{N}\cos\alpha_i+\sum_{i\neq j}\ln\left|\sin\frac{\alpha_i-\alpha_j}{2}\right|\right)+\text{const}, \tag{17}$$

where the eigenvalues α_i are given by the stationarity condition

$$\frac{2}{\lambda}\sin\alpha_i=\sum_{j\neq i}\cot\left|\frac{\alpha_i-\alpha_j}{2}\right|. \tag{18}$$

Also, in the large-N limit these equations can be replaced by their continuum version by introducing a nondecreasing function $\alpha(x)$, $0\leqslant x\leqslant 1$ such that

$$\alpha_i=\alpha(i/N),\quad i=1,\ldots,N, \tag{19}$$

in which case

$$-E_0(\lambda)\equiv\lim_{N\to\infty}\frac{\ln z(g^2,N)}{N^2}=\frac{2}{\lambda}\int_0^1 dx\cos\alpha(x)+P\int_0^1 dx\int_0^1 dy\ln\left|\sin\frac{\alpha(x)-\alpha(y)}{2}\right|+\text{const} \tag{20}$$

[the constant is adjusted so that $E_0(\infty)=0$], and

$$\frac{2}{\lambda}\sin\alpha(x)=P\int_0^1 dy\cot\frac{\alpha(x)-\alpha(y)}{2}, \tag{21}$$

where P refers to the principal part of the integral.

This equation can be solved by introducing, following Brezin *et al.*,[5] the density of eigenvalues

$$\rho(\alpha)=dx/d\alpha\geqslant 0,\qquad \int_{-\alpha_c}^{\alpha_c}d\alpha\,\rho(\alpha)=\int_0^1 dx=1. \tag{22}$$

Here we allow for the possibility that the eigenvalues lie in the region $|\alpha|\leqslant\alpha_c$, $\alpha_c\leqslant\pi$. Equation (21) then becomes

$$\frac{2}{\lambda}\sin\alpha=P\int_{-\alpha_c}^{+\alpha_c}d\beta\,\rho(\beta)\cot\left(\frac{\alpha-\beta}{2}\right). \tag{23}$$

For large λ we expect that the eigenvalues of W will be spread uniformly over the whole circle and $\alpha_c=\pi$. In that case, Eq. (21) can easily be solved by using

$$\cot\frac{\alpha-\beta}{2}=2\sum_{n=1}^{\infty}(\sin n\alpha\cos n\beta-\cos n\alpha\sin n\beta)$$

with the result that

$$\rho(\alpha)=(1/2\pi)[1+(2/\lambda)\cos\alpha]. \tag{24}$$

This indeed is positive for all α provided that $\lambda\geqslant 2$, and yields the unique solution of Eq. (23) and Eq. (22) for $\lambda\geqslant 2$.

For $\lambda\leqslant 2$ one must allow α_c to be less than π. In that case, to solve Eq. (23) define a function $F(Z)$

$$F(Z)\equiv\int_{-\alpha_c}^{+\alpha_c}d\beta\,\rho(\beta)\cot\frac{Z-\beta}{2} \tag{25}$$

which possesses the following properties:

(1) $F(Z)=F(Z+2\pi)$.

(2) $F(Z)$ is analytic for complex Z outside the real intervals $(-\alpha_c+2\pi N,\ \alpha_c+2\pi N)$.

(3) F is real for Z real outside the intervals $(-\alpha_c+2N\pi,\ \alpha_c+2N\pi)$, and when these intervals are approached

$$F(\alpha\pm i\epsilon)=\frac{2}{\lambda}\sin\alpha\mp i2\pi\rho(\beta), \tag{26}$$

since $\cot\frac{1}{2}(Z-\beta)$ is analytic in the Z plane except for simple poles at $Z=\beta+2N\pi$.

(4) $F(Z)\to\pm 1$ as $|Z|\to\infty$ in any direction except along the real axis, and $\operatorname{Im}Z\gtrless 0$, as a consequence of Eq. (22).

There exists a unique function which satisfies all of these properties:

$$F(Z)=\frac{2}{\lambda}\sin\alpha-\frac{4}{\lambda}\cos\frac{\alpha}{2}\left(\sin^2\frac{\alpha}{2}-\sin^2\frac{\alpha_c}{2}\right)^{1/2} \tag{27}$$

where

$$\sin^2(\tfrac{1}{2}\alpha_c)=\tfrac{1}{2}\lambda. \tag{28}$$

The square root is defined in the (multiply) cut α plane and chosen to be positive for $\alpha_c<\alpha<2\pi-\alpha_c$. It then follows that $F(Z)$ is periodic, since when $\alpha\to\alpha+2\pi$, both $\cos\frac{1}{2}\alpha$ and $[\sin^2(\frac{1}{2}\alpha)-\sin^2(\frac{1}{2}\alpha_c)]^{1/2}$ change sign. The discontinuity of $F(Z)$ then determines the density of eigenvalues to be

$$\rho(\alpha)=\frac{2}{\pi\lambda}\cos\frac{\alpha}{2}\left(\frac{\lambda}{2}-\sin^2\frac{\alpha}{2}\right)^{1/2}. \tag{29}$$

We note that this solution only makes sense for $\lambda\leqslant 2$, and for $\lambda=2$ it equals $(1/\pi)\cos^2(\frac{1}{2}\alpha)=1/2\pi(1+\cos\alpha)$ which coincides with our previous solution derived for $\lambda>2$.

To summarize, we note that the density of eigenvalues in the $N\to\infty$ limit is characterized by two separate analytic functions, one appropriate for large coupling ($\lambda\geq 2$) and one for small coupling ($\lambda\leq 2$):

$$\rho(\alpha)=\frac{dx}{d\alpha}=\begin{cases}\frac{2}{\pi\lambda}\cos\frac{\alpha}{2}\left(\frac{\lambda}{2}-\sin^2\frac{\alpha}{2}\right)^{1/2}, & \lambda\leq 2\quad |\alpha|<2\sin^{-1}\left(\frac{\lambda}{2}\right)^{1/2}\\ \frac{1}{2\pi}\left(1+\frac{2}{\lambda}\cos\alpha\right), & \lambda\geq 2\quad |\alpha|\leq\pi\,.\end{cases}\tag{30}$$

Thus there exists a phase transition (for $N=\infty$) at $\lambda=2$ between weak and strong coupling. The origin of the phase transition is clear. For very large λ the functional integral is dominated by the term $\Delta^2(\alpha_i)$ which causes the eigenvalues to repel, the Wilson action can be neglected to first approximation, and the density of eigenvalues is uniform, $\rho=1/2\pi$. On the other hand, for very small λ the Wilson action dominates, and the saddle point corresponds to $\alpha\approx O(\sqrt{\lambda})$. In fact, as $\lambda\to O$, the distribution of eigenvalues is given by Wigner's semicircle law,

$$\rho(\alpha)\simeq\frac{1}{\pi}\left(1-\frac{\alpha^2}{2\lambda}\right)^{1/2},\quad |\alpha|\leq\sqrt{2\lambda},\quad \lambda\approx 0\,.\tag{31}$$

The phase transition occurs precisely at the point at which the eigenvalues fill the whole unit circle.

We are now in a position to calculate $E_0(\lambda)$ for all λ:

$$-E_0(\lambda)=\frac{2}{\lambda}\int_{-\alpha_c}^{+\alpha_c}d\alpha\,\rho(\alpha)\cos\alpha+P\int_{-\alpha_c}^{\alpha_c}d\alpha\,d\beta\,\rho(\alpha)\rho(\beta)\ln\left|\sin\frac{\alpha-\beta}{2}\right|-\frac{1}{2\pi}\int_{-\pi}^{+\pi}d\alpha\ln\left|\sin\frac{\alpha}{2}\right|.\tag{32}$$

The integration is easily performed, yielding

$$-E_0(\lambda)=-\frac{Fa^2}{\lambda N^2}=\begin{cases}\frac{1}{\lambda^2}, & \lambda\geq 2\\ \frac{2}{\lambda}+\frac{1}{2}\ln\frac{\lambda}{2}-\frac{3}{4}, & \lambda\leq 2\,.\end{cases}\tag{33}$$

Here we see explicitly that the free energy is given by two different functions of λ, both analytic except at $\lambda=0$.

The expectation value of the Wilson loop operator $w(g^2,N)$ is easily constructed, since $w=-(\lambda^2/2N^2)\partial\ln z/\partial\lambda$ we have

$$w(\lambda)=\lim_{N\to\infty}w(g^2,N)=\begin{cases}1/\lambda, & \lambda\geq 2\\ 1-\lambda/4, & \lambda\leq 2\,,\end{cases}\tag{34}$$

and the string tension $\sigma(g^2,N)$ is given by

$$\sigma(\lambda)=\lim_{N\to\infty}\sigma(g^2,N)=\begin{cases}\frac{1}{a^2}\ln\lambda, & \lambda\geq 2\\ \frac{1}{a^2}\ln\frac{4}{4-\lambda}, & \lambda\leq 2\,.\end{cases}\tag{35}$$

It is also instructive to construct the β function for the $N=\infty$ theory. We imagine varying the lattice spacing and the value of the coupling λ so as to keep the string tension [or $\epsilon(R)$] fixed. This determines the effective coupling $\lambda(a)$ so that $\sigma[a,\lambda(a)]=\sigma$, namely

$$\lambda(a)=\begin{cases}e^{a^2\sigma}, & \lambda\geq 2\\ 4(1-e^{-a^2\sigma}), & \lambda\leq 2\,.\end{cases}\tag{36}$$

The β function, which yields the variation with length of the effective coupling,

$$-\frac{d\lambda(a)}{d\ln a}=-\beta(\lambda)=\begin{cases}2\lambda\ln\lambda, & \lambda\geq 2\\ 2(4-\lambda)\ln\frac{4}{4-\lambda}, & \lambda\leq 2\end{cases}\tag{37}$$

is plotted in Fig. 1.

The fact that $\beta(2)\neq 0$ means that the phase transition is not of second order. A second-order phase transition requires that the string tension vanish at the critical coupling, and that the weak-coupling phase is nonconfining. A naive extrapolation of the strong-coupling result [Eq. (37)] to $\lambda=1$ would predict such a phase transition; however, before this point, at $\lambda=2$, a phase transition of higher order occurs.

The order of the phase transition is easily seen to be three, namely the free energy; its first and second derivatives are all continuous but the third derivative of F is discontinuous at $\lambda=2$. In

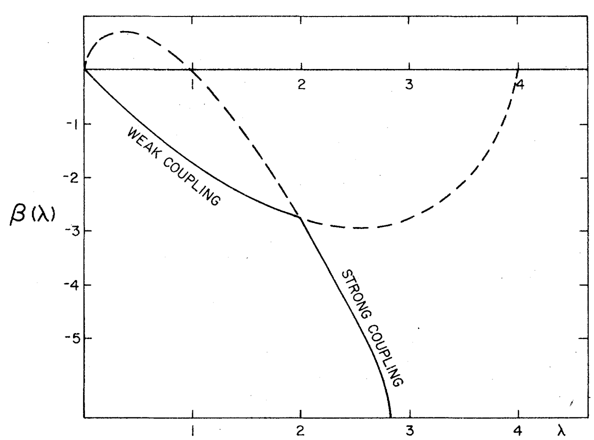

FIG. 1. The β function as a function of λ. The dashed lines are the (invalid) extrapolation of the weak- and strong-coupling results beyond the phase transition at $\lambda=2$.

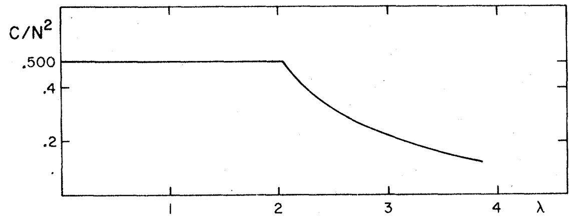

FIG. 2. The specific heat per degree of freedom, C/N^2, as a function of λ (temperature).

fact, if we regard the theory as a statistical-mechanical system with temperature $T=\lambda$ the internal energy $E(\lambda)$ per unit volume is given by

$$-E=-T^2\,\partial F/\partial T=2N^2 w(\lambda) \tag{38}$$

and is continuous at $\lambda=2$. The specific heat

$$C=\frac{\partial E}{\partial T}=2N^2\times\begin{cases}\frac{1}{\lambda^2}, & \lambda\geq 2\\ \frac{1}{4}, & \lambda\leq 2\end{cases} \tag{39}$$

is also continuous; however, the first derivative of the specific heat clearly is discontinuous at $\lambda=2$ (see Fig. 2).

IV. REMARKS AND CONCLUSIONS

In this section we shall discuss some of the interesting aspects of the large-N two-dimensional gauge theory and attempt to draw some conclusions that might be relevant to the behavior of four-dimensional lattice gauge theories.

First we remark that the large-N limit of the theory could have been derived, for large $g^2N=\lambda$ (actually $\lambda\geq 2$), by interchange of the $N\to\infty$ limit and the strong-coupling expansion. Consider, for example, the evaluation of $z(g^2,N)$ [Eq. (8)] and expand the integral in powers of $1/\lambda$ (for fixed N). Thus

$$z=\sum_{n=0}^{\infty}\int[dW]\left(\frac{1}{n!}\right)^2\left(\frac{N}{\lambda}\right)^{2n}(\mathrm{Tr}W)^n(\mathrm{Tr}W^\dagger)^n\,. \tag{40}$$

Now it is easy to prove, by expanding $(\mathrm{Tr}W)^n$ in characters of U(N) that

$$\int[dW](\mathrm{Tr}W)^n(\mathrm{Tr}W^\dagger)^n=n!\,,\quad n\leq N\,. \tag{41}$$

Therefore, if we interchange the large-N limit with the strong-coupling expansion we would derive

$$\lim_{N\to\infty}z(\lambda,N)=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{N^2}{\lambda^2}\right)^n=\exp\left[N^2\left(\frac{1}{\lambda^2}\right)\right] \tag{42}$$

and thus $-E_0(\lambda)=1/\lambda^2$, in accord with our *exact* result, Eq. (33) for $\lambda\geq 2$.

The interchange of perturbative expansions with the limit $N\to\infty$ is commonly employed to define the large-N limit. Thus in continuum QCD the equivalence of the large-N limit and the sum of planar graphs is established by interchanging the $N\to\infty$ limit with the perturbative expansion in powers of λ. There is no known reason to suspect this interchange. However, in the case of the strong-coupling expansion the interchange is highly suspicious. The terms in Eq. (40), for $n>N$, that one is dropping are of order $(N/\lambda)^{2k}(1/N^m)$ $(k>N)$, and thus while suppressed, compared to the leading terms, are not at all negligible.

Our result shows that while it might be the case that the $N\to\infty$ limit and the strong-coupling expansion are interchangeable for sufficiently large coupling (here $\lambda\geq 2$), they will not be so for small coupling. In fact, it is clear that Eq. (42) could not be valid for all λ, since it implies that

$$1\geq w(\lambda)=\left\langle\frac{1}{N}\mathrm{Tr}W\right\rangle=\frac{1}{\lambda} \tag{43}$$

and therefore must fail for $\lambda\leq\lambda_c$, $\lambda_c\geq 1$ (in fact, $\lambda_c=2$). Furthermore, the existence of our phase transition follows once one derives Eq. (42) for large coupling, since the resulting $w(\lambda)$, Eq. (43), is analytic for all $\lambda>0$ and yet cannot be the correct $w(\lambda)$ for $\lambda<1$.

In the real world (four-dimensional QCD), life is much more complicated. However, if one could interchange the large-N limit with the strong-coupling expansion one could derive (for large λ) a lattice version of the string model. As shown by Bars and Green,[3] if one uses the "approximation"

$$\int du\,\exp[(N/\lambda)\,\mathrm{Tr}(UA+A^\dagger U^\dagger)]\simeq\exp[(N/\lambda^2)\,\mathrm{Tr}AA^\dagger]\,, \tag{44}$$

which would follow from exchanging the limit $N\to\infty$ with the expansion of Eq. (44) in powers of $1/\lambda$, then one can integrate the four-dimensional U(N) gauge theory link by link. One then derives the noninteracting lattice string model where the expectation value of a Wilson loop operator is given by

$$\left\langle\frac{1}{N}\mathrm{Tr}\,W_L\right\rangle=\sum_{S_L}\left(\frac{1}{\lambda}\right)^{A(S_L)}, \tag{45}$$

where the sum runs over all planar surfaces S_L bounded by the loop L, and $A(S_L)$ is the area of the surface.

As indicated above we have no reason to trust Eq. (44), except for $\lambda\simeq\infty$, and indeed there are corrections for a generic matrix $A\neq I$ for finite λ. Thus there is no solid reason to expect the string model [Eq. (45)] to emerge in four dimensions as $N\to\infty$. *If*, however, one did prove that Eq. (45) was correct and convergent for $\lambda\geq\lambda_0$ (it is reasonable to expect that $\lambda_0\approx 6$),[4] then it must be the case that there is a phase transition at $\lambda=\lambda_c\geq \mathrm{Max}(\lambda_0,1)$. This follows by considering a one-plaquette Wilson loop given by

$$1\geq\left\langle\frac{1}{N}\mathrm{Tr}\,W_1\right\rangle=\sum_{S_1}\left(\frac{1}{\lambda}\right)^{A(S_1)}\geq\frac{1}{\lambda}, \quad (46)$$

which is a sum of positive terms, is greater than $1/\lambda$, and becomes greater than one for some $\lambda>\lambda_0$.

Therefore, we conclude that *if* the noninteracting string model is valid on the lattice, as $N\to\infty$ for large λ, a phase transition of the type discussed in this paper must occur. It is therefore unlikely that the continuum theory in the large-N limit is described by a noninteracting string model. This is hardly surprising. 't Hooft's analysis of the large-N behavior of the continuum theory[2] only ensures planarity in index space and not in real space-time. Furthermore, it is hard to see how the soft behavior at large momentum of the string model could possibly be consistent with the pointlike interactions that hold in an asymptotically free gauge theory.

Second we note that the phase transition discussed in this paper is quite different from the second-order phase transitions which one normally searches for in lattice gauge theories. The latter are characterized by a discontinuity, or divergence, in the specific heat, an infinite correlation length at the critical coupling = temperature, and a qualitative difference in the behavior of the Wilson loop for large loops in the two phases. For example, the naive extrapolation of the strong-coupling result would predict such a phase transition at $\lambda=1$, where $\beta(\lambda)$ and the "string tension" vanish, resulting in the lack of confinement for $\lambda<1$. Our phase transition is of a different nature. It arises in the "thermodynamic limit" $N\to\infty$, which yields an infinite number of degrees of freedom even in a finite volume. Its origin resides in the fact that for small λ the functional integral is strongly peaked about plaquette matrices close to the identity whereas for large λ the integral receives contributions from arbitrary, random, unitary matrices. For $N=\infty$ a particular unitary matrix, up to similarity transformations, dominates the integration, and the above tendency is so enhanced that for $\lambda<2$ the matrices that contribute are restricted to a finite portion of the group manifold peaked about $W=I$. This region increases with increasing λ and for $\lambda=2$ fills the whole manifold. At this point there occurs the phase transition to the strong-coupling phase, where the distribution becomes increasingly uniform as $\lambda\to\infty$. It is apparent that the phase transition will be third order; namely, the internal energy or the expectation value of the Wilson loop is clearly continuous at $\lambda=2$ and there is no reason at this point for the string tension to vanish. It is possible to find other "order parameters" which illustrate in a more dramatic fashion the nature of the phase transition at $\lambda=2$. Consider, for example, the expectation value of a power of a single plaquette matrix, $w_k=(1/N)\langle\mathrm{Tr}(W^k)\rangle$. This variable is a measure of the randomness of the distribution of W's. For a uniform distribution all w_k would vanish. In the large-N limit, w_k is easily calculated:

$$w_k=\lim_{\substack{N\to\infty\\ k\geq 2}}\int_{-\alpha_c}^{+\alpha_c}d\alpha\,\cos k\alpha\,\rho(\alpha)$$

$$=\begin{cases}0, & \lambda\geq 2\\ (1-\frac{1}{2}\lambda)\left[\dfrac{P_k'(1-\lambda)}{k(k+1)}+\dfrac{P_{k-1}'(1-\lambda)}{k(k-1)}\right], & \lambda\leq 2\,.\end{cases} \quad (47)$$

Thus for $\lambda\geq 2$ the distribution of W's is as random as could be—all $w_{k\geq 2}=0$ and $w_1=1/\lambda$, whereas for $\lambda\leq 2$, $w_k(\lambda)\neq 0$. Note that all w_i are once differentiable, but their second derivatives are discontinuous, at $\lambda=2$.

We see no reason why such a third-order, weak-to strong-coupling phase transition would not occur in the large-N limit of the four-dimensional gauge theory. The occurrence of such a phase transition would not mean that the large-N theory does not confine. It would, however, imply that the weak- and strong-coupling lattice theories are not described by the same analytic functions, and that one cannot deduce the properties of the continuum theory from the $(N=\infty)$ strong-coupling theory.

Finally, we note that the limit $N\to\infty$ is crucial to the existence of the phase transition. For finite N, the functions F, w, etc., are all analytic functions of λ for all $0<\lambda\leq\infty$. Thus for any finite N there will be no phase transition. Clearly, as $N\to\infty$ an infinite number of zeros of z, which lie in the complex λ plane for finite N, accumulate to form a natural boundary which presents the analytic continuation from $\lambda>2$ to $\lambda<2$. For any finite N none of these zeros (logarithmic singularities of the free energy) will lie on the real axis and they will not be dense.

If there exists an analogous weak-to-strong phase transition in the four-dimensional theory, we again would expect it to occur only in the $N\to\infty$ limit. For finite N, therefore, one could continue from strong to weak coupling (assuming no other phase transitions). However, one would expect to see a sign of $N=\infty$ phase transition for large enough $N(3?)$ whose manifestation would be a sharp transition at $\lambda \simeq \lambda_c$ from weak-coupling to strong-coupling behavior. We note that precisely such a sharp transition is consistent with the results of Wilson [numerical integration of an SU(2) lattice gauge theory],[6] of Kogut, Pearson, and Shigemitsu[7] [Padé of the SU(3) lattice gauge theory strong-coupling expansion], and of Callan, Dashen, and Gross[8] (semiclassical treatment of the transition from weak to strong coupling). Increased understanding of the possible $N=\infty$ phase transition could be helpful in probing QCD in the region of transition from weak to strong coupling.

ACKNOWLEDGMENTS

We would like to thank I. Bars, R. Dashen, M. Peskin, A. Ukawa, and P. Windey for helpful conversations. The research of D. G. was supported in part by NSF Grant No. PHY78-01221. The research of E. W. was supported in part by NSF Grant No. PHY77-22864 and by the Harvard Society of Fellows.

[1] K. Wilson, Phys. Rev. D 10, 2445 (1974); J. Kogut and L. Susskind, *ibid*. 11, 395 (1975).
[2] G. 't Hooft, Nucl. Phys. B72, 461 (1974).
[3] I. Bars and F. Green, I. A. S. report, 1979 (unpublished).
[4] D. Foerster, Phys. Lett. 77B, 211 (1978); A. Polyakov, *ibid*. 82B, 247 (1979).
[5] E. Brezin, C. Itzykson, G. Parisi, and J. B. Zuber, Commun. Math. Phys. 59, 35 (1978). For some related work see M. Casartelli, G. Marchesini, and E. Onofri, University of Parma report, 1979 (unpublished).
[6] K. Wilson, private communication.
[7] J. Kogut, R. Pearson, and J. Shigemitsu, Phys. Rev. Lett. 43, 484 (1979).
[8] C. Callan, R. Dashen, and D. Gross, Phys. Rev. D 20, 3279 (1979).