Nuclear Physics B196 (1982) 189–204

ON THE PHASE STRUCTURE OF VECTOR-LIKE GAUGE THEORIES WITH MASSLESS FERMIONS

T. BANKS and A. ZAKS[1]

Department of Physics and Astronomy, Tel Aviv University, Tel Aviv, 69978 Israel

Received 16 June 1981
(Corrected version received 22 October 1981)

We present a systematic expansion for studying an infrared-stable fixed point of gauge theories with massless fermions. These results are combined with information from strong coupling and large-N_F expansions to sketch the phase diagrams for lattice gauge theories with massless fermions. On the basis of these diagrams we conclude that a large class of continuum models with fermions in real representations of the gauge group do not have spontaneous chiral symmetry breaking. We also argue that the transition between chirally symmetric and asymmetric phases is generally first order.

1. Introduction

It has long been known [1] that the asymptotic freedom of QCD disappears if the number of fermion species or the dimension of their color representation is too large. The critical number of Dirac fermions in the fundamental representation of SU(3) for which the first coefficient of the β function vanishes is $N^* = 16.5$. To our knowledge no one has attempted to understand the structure of models with $N_F > N^*$. Since the asymptotically free fixed point has disappeared one must introduce a regulator in order for such an attempt to make sense.

Consider, then, the hamiltonian lattice gauge theory with Kogut–Susskind fermions which corresponds to the formal continuum SU(N) gauge theory with $N_F > N^*$ fermions in the fundamental representation*. For $g \ll 1$ the β function of this theory is positive while for $g \gg 1$ it is negative [2]. Thus, (assuming that β is continuous) there is a critical value $g = g_c$ for which $\beta = 0$. According to standard lore this implies that the model undergoes a second-order phase transition at g_c. Since g_c is ultraviolet stable we should (assuming that there are no *other* transitions) be able to define exactly two continuum theories by approaching g_c from above or below. The theory defined via the upper phase should be confining (whatever this means when there are matter fields in the fundamental representation) and exhibit chiral

[1] Supported in part by the US–Israel Binational Science Foundation.

* Here and henceforth N_F refers to the number of fermion fields in the formal continuum theory which is twice the number of single component lattice fields. It is possible to write down lattice theories that have odd numbers of continuum fermion fields by breaking cubic invariance a la Nielsen and Ninomiya [3]. We feel sure that such theories will follow the pattern to be outlined here although we have not studied them in detail.

symmetry breaking if the fermions are massless. Its short-distance behavior will be scale invariant with non-canonical dimensions.

The short-distance behavior of the lower phase is similar. Its infrared behavior depends on the fermion masses. Massive fermions drop out of the renormalization group equations when the energy falls below their masses. Thus the infrared behavior will be like that of pure gauge theory. On the other hand, if the fermions are massless, the lower phase is infrared free with Green functions which scale as powers of logarithms. In this case the theory does not seem to admit of a particle interpretation (cf. the massless abelian Thirring model). Finally we note that chiral symmetry is not broken in the lower phase.

We have lately embarked on a program designed to verify and confirm the simple and satisfying picture outlined above. The idea was to vary the number of fermions in order to bring the critical point g_c into the range of validity of either the strong coupling or weak coupling expansion. In the course of this investigation we remembered two curious facts which may change the picture. Firstly, as N_F is increased from zero, the second coefficient of the β function [4] changes sign before the first [at $N_F = 8.05$ for an SU(3) gauge theory]. This indicates the presence of a non-trivial zero in the perturbative β function. What is more interesting however, is that as $N_F \uparrow N^*$ this zero occurs at smaller and smaller values of g^2, which implies that it can be reliably studied in perturbation theory. We envision a systematic expansion of the critical behavior at this fixed point in powers of $(N_F - N^*)$. The existence of this fixed point may drastically change the simple picture proposed above.

Secondly, we realized that there was an ambiguity in the definition of the β function at strong coupling and that in gauge theories with fermions different definitions of the β function behave in a qualitatively different manner (this is a physical effect not related to the artificial renormalization scheme dependence of perturbative QCD).

These two curious facts have caused us quite a bit of confusion. While we now believe that we have found a new picture of the phase structure which is consistent with all the data available to us, our shock at finding the previous picture false has made us wary of phase diagrams based on folkfore and hand-waving arguments. We have decided therefore to divide this paper into two parts. In sect. 2 we discuss continuum Yang–Mills theory for $N_F \leq N^*$. We present the $(N_F - N^*)$ expansion and discuss the (infrared) critical behavior at the perturbative fixed point, $g^* = O(N_F - N^*)$.

We search for (and do not find) other small zeros of β. Finally we discuss the crucial question of whether the expansion can be trusted at physical (integer) values of N_F. This section is fairly solid and straightforward. Sect. 3 is a highly speculative attempt to construct a phase diagram for SU(N) gauge theories with fermions. We use the results of sect. 2, strong coupling information, and a weak coupling–large-N_F expansion which is described in the appendix. Unfortunately, we also make frequent use of terms like "maximal simplicity", "reasonable" and "intuitively appealing", so our phase diagram does not stand on a very firm basis. It is a conjecture consistent with the evidence available to us.

However, if it is correct, it implies a host of interesting things including a wide extension of the class of theories with massless composite fermions [5]. It also predicts a first-order phase transition in theories with fermions in the fundamental representation.

In sect. 2 we concentrate mostly on an SU(3) gauge theory though other theories will give qualitatively similar results. In sect. 3 the gauge group is fixed but unspecified because most of the arguments are quite general. We do not discuss the large-N_c limit.

2. The IR-stable fixed point and the $N_F - N^*$ expansion

The 3-loop β function of an SU(3) gauge theory with N_F Dirac fermions in the presentation R is

$$\beta(g) = -\left(\beta_0 \frac{g^3}{16\pi^2} + \beta_1 \frac{g^5}{(16\pi^2)^2} + \beta_2 \frac{g^7}{(16\pi^2)^3}\right), \tag{1}$$

$$\beta_0 = 11 - \tfrac{4}{3} T(\mathrm{R}) N_F, \tag{2}$$

$$\beta_1 = 102 - (20 + 4C_2(\mathrm{R}))T(\mathrm{R})N_F, \tag{3}$$

$$\beta_2 = \left(\tfrac{2857}{2} - \tfrac{5033}{18} N_F + \tfrac{325}{54} N_F^2\right), \qquad (\mathrm{R} = \text{fundamental}). \tag{4}$$

Here $C_2(\mathrm{R})$ is the second Casimir operator and

$$T(\mathrm{R}) = \tfrac{1}{8} C_2(\mathrm{R}) \dim(\mathrm{R}). \tag{5}$$

β_0, β_1 are invariant under changes of renormalization scheme which are analytic at $g = 0$. β_2 is computed [7] in the minimal subtraction scheme.

Note that β_1 changes sign at $N_F' = 102/[(20+4C_2)T]$, which is always smaller than $N^* = 33/4T$, the point at which β_0 vanishes. Thus for $N^* > N_F > N_F'$ the two-loop β function has a non-trivial zero and if $N_F = N^*(1 - \frac{1}{11}\epsilon)$, then the zero is at

$$\frac{g^{*2}}{16\pi^2} = \frac{\frac{1}{3}\epsilon}{21 + 11C_2 - (5 + C_2)\varepsilon}. \tag{6}$$

For ε very small and positive this fixed point should be in the range of validity of the perturbation expansion so that we can believe our calculation. Of course, $\varepsilon \ll 1$ does not correspond to a physical (integral) value of N_F, but we may imagine analytically continuing our theory to physical values by expanding in powers of ε. This is analogous to the $4-d$ [8] and $(d-2)$ [9] expansions in the theory of critical phenomena. In fact that $N^* - N_F$ expansion is *a priori* on a better theoretical footing than either of these expansions. N_F appears analytically in the functional integral representation of the gauge theory, so we can give a non-perturbative meaning to the theory with non-integral N_F.

Standard lore has it that the zero of the β function that we have found controls the infrared behavior of the asymptotically free gauge theory with massless fermions at those values of N_F for which the zero exists. Since the zero is simple, Green functions will scale at large distances or low momenta with non-canonical powers. We will use statistical mechanics terminology and call the anomalous powers critical indices.

The most important of these indices is the correlation index

$$\nu = \left(-\frac{d\beta}{dg}\bigg|_{g^*} \right)^{-1} . \tag{7}$$

To leading order in ε it is given by

$$\nu = \frac{1}{3}\left(\frac{4\pi}{g^*}\right)^2 = -\varepsilon^{-1}(21 + 11C_2(\mathrm{R}) + \mathrm{O}(\varepsilon)) . \tag{8}$$

If we could apply hyperscaling [10] we could now compute the specific heat index $\alpha = 2 - 4\nu$. We find

$$\alpha = 2 + 4\varepsilon^{-1}(21 + 11C_2(\mathrm{R}) + \mathrm{O}(\varepsilon)) . \tag{9}$$

Note that $\alpha \geqslant 2$, a fact which will be important in our discussion in sect. 3. Actually, as we will emphasize there, the use of hyperscaling in eq. (9) is not valid.

Finally, we record the anomalous dimension of the chiral order parameter $\bar{\psi}\psi$:

$$\gamma_{\bar{\psi}\psi} = -\frac{g^2}{2\pi^2} = \frac{-2\varepsilon}{63 + 33C_2(\mathrm{R})}(1 + \mathrm{O}(\varepsilon)) , \tag{10}$$

and Peskin's [11] index for the Wilson loop which is

$$\eta = -2C_2(\mathrm{R})\left(\frac{g^*}{4\pi}\right)^2 = -\frac{2C_2(\mathrm{R})}{63 + 33C_2(\mathrm{R})}(\varepsilon + \mathrm{O}(\varepsilon^2)) . \tag{11}$$

It is important to estimate the corrections to these results in order to understand their relevance for physical values of N_F. Every fermion loop introduces a factor that goes like $N_F(=N^*)$ dim (R) (the factor of N^* can also be seen directly in the path integral representation) so we might expect the real expansion parameter to be

$$\frac{\varepsilon N^* \dim(\mathrm{R})}{63 + 33C_2(\mathrm{R})} . \tag{12}$$

For the fundamental representatgion of SU(3) this is

$$\tfrac{99}{214}\varepsilon \ (=0.46\varepsilon) \tag{13}$$

which is $\frac{11}{71}$ for $\varepsilon = \frac{1}{3}$ ($N_F = 16$) the first physical value. This seems like a reasonably small value.

More reliable error estimates can be obtained by explicit 2-loop computations of the quantities described above. This can be done for the case of the fundamental representation because we have the third coefficient of the β function. Note that the scheme dependence of this coefficient is irrelevant for our purposes because we are calculating anomalous dimensions at a fixed point. We must only take care to evaluate the perturbation expansion of the anomalous dimensions in the MS scheme (or in $\overline{\mathrm{MS}}$ which differs from MS only via a change in the scale parameter).

The $N^* - N_{\mathrm{F}}$ expansion of the fixed point of the β function in the MS scheme is

$$\left(\frac{g^*}{4\pi}\right)^2 = a\varepsilon + b\varepsilon^2 + c\varepsilon^3 + \mathrm{O}(\varepsilon^4)\,, \tag{15}$$

$$a = \frac{-1}{\beta_1(N^*)} = \frac{1}{63 + 33C_2(\mathrm{R})} = 0.01\,,$$

$$b = \frac{1}{\beta_1(N^*)^2}\left(15 + 3C_2(\mathrm{R}) - 2\frac{\beta_2(N^*)}{\beta_1(N^*)}\right) = -1.2\times 10^{-3}\,.$$

If we know the perturbation expansion of some physical quantity P through two-loop order,

$$P = A\frac{g^2}{16\pi^2} + B\frac{g^4}{(16\pi^2)} + C\frac{g^6}{(16\pi^2)^3} + \mathrm{O}(g^2)\,, \tag{16}$$

then its $N^* - N_{\mathrm{F}}$ expansion is

$$P = aA\varepsilon + (a^2B + bA)\varepsilon^2 + (cA + 2abB + a^3C)\varepsilon^3 + \mathrm{O}(\varepsilon^4)\,. \tag{17}$$

For ν^{-1} we have

$$A = 3\beta_0(N_{\mathrm{F}}^*)\,, \qquad B = 5\beta_1(N_{\mathrm{F}}^*)\,, \qquad C = 7\beta_2(N_{\mathrm{F}}^*)\,, \tag{18}$$

so

$$\nu^{-1} = (-3a + 5\beta_1 a^2)\varepsilon^2 + (3b + 10ab\beta_1 + 7\beta_2 a^3)\varepsilon^3\,. \tag{19}$$

The first term is larger than the second by at least a factor of ten as long as

$$\varepsilon \leqslant \left|\frac{0.3a + 0.5\beta_1 a^2}{3b + (75 + 15C_2)a^2 + 10ab\beta_1 + 7a^2\beta_2}\right| = 0.41\,. \tag{20}$$

We do not know of a two-loop calculation of $\gamma_{\bar{\psi}\psi}$ but we can use the two-loop calculations of the anomalous dimensions of the twist two operators which appear in deep inelastic scattering [12]. The values A_S and B_S (for the spin S operator) appear in table 1 along with the coefficient of the first two terms in the $N^* - N_{\mathrm{F}}$ expansion.

To summarize this section, we have found a non-trivial zero of the β function of Yang–Mills theories with certain fermion representations. The fixed point can be investigated in an expansion in the number of flavors around a non-integral

TABLE 1

S	A_S	B_S	$\bar{A}_S$	$\bar{B}_S$
2	7.11	27.36	0.071	−0.0058
3	8	−9.65	0.08	−0.0106
4	13.96	−32.87	0.139	−0.02
5	16.17	−49.92	0.162	−0.0244
12	25.07	−106.97	0.25	−0.04

S = spin, A,B coefficient of loop expansion, $\bar{A}$, $\bar{B}$ coefficient of $N_F - N^*$ expansion.

value N^*. For the SU(3) gauge theory with fermions in the fundamental representation the expansion seems to be well behaved down to physical values of N_F (at least $N_F = 16$). This would appear to imply that the $N_F = 16$ theory does not exhibit chiral symmetry breaking since the two-point function of the chiral order parameter $\bar{\psi}\psi$ satisfies the cluster property. Previously it was supposed that all the models with $N_F < \frac{33}{2}$ had the same qualitative infrared structure (chiral symmetry breaking, linear Regge trajectories, etc.), but the $N^* - N_F$ expansion predicts that for some range of N_F between N^* and $N_F' = 8.05$ these models are chirally invariant, have no mass gap and do not have a particle interpretation.

3. The phase diagram of gauge theories with $N_F < N^*$

We now want to confront the results of the $(N^* - N_F)$ expansion with non-perturbative information about lattice gauge theories. To do this we must clarify the meaning of the β function in strong coupling lattice theories. In general [13], the renormalization group transformation for a cut off field theory operates on the infinite-dimensional space of short-range hamiltonians compatible with a given symmetry. A single-parameter renormalization group equation is valid only in the vicinity of a very special type of fixed point. The use of single-parameter β functions in strong coupling lattice gauge theory is based on the following logic:

We want to approach the $g^2 = 0$ fixed point, which is known to be asymptotically free and have a single relevant parameter (in a model with chiral symmetry). If we define a β function in terms of some physical quantity (e.g. the string tension) then the absence of zeros of β guarantees that this quantity extrapolates smoothly to small g. It is then reasonable to assume that it has a continuum limit, i.e. that its β function coincides at small g with the universal β of the continuum.

A priori there is no reason to expect that β functions defined in terms of different physical quantities have the same qualitative behavior at strong coupling. If they do not, different aspects of the phase structure of the theory may be revealed in different β functions.

This is precisely what happens in lattice gauge theories with massless Kogut–Susskind fermions [15]. The theory is divided into sectors according to the eigenvalues of the external color density:

$$\sum_{i=\pm 1}^{3} E_i^\alpha(x) = \rho^\alpha(x) . \tag{21}$$

For strong coupling the lowest excitation energy in the ρ sector $M_\rho(g) = E_\rho - E(\text{vac})$ behaves like $1/g^2$ or g^2 depending on whether ρ can be screened by fermions on a local, site by site, basis without exciting color electric flux. Defining $\beta_\rho(g)$ by

$$\left.\begin{aligned} &\beta_\rho(g(a)) = -a\frac{\mathrm{d}g}{\mathrm{d}a} \\ &\frac{\mathrm{d}}{\mathrm{d}a}(m_\rho(g(a))/a) = 0 \end{aligned}\right\} \Rightarrow \beta_\rho(g) = -(\mathrm{d}\ln m_\rho/\mathrm{d}g)^{-1} , \tag{22}$$

we find $\beta_\rho > 0$ (<0) if ρ is (not) screenable. In particular if the fermions are in a representation invariant under the center of the group then we can define a string tension β function β_{T} (<0 for large g). If chiral symmetry breaking takes place (as it certainly does for $N_{\mathrm{F}} > 2$) then we can define another β function β_m (>0 for large g) in terms of the chiral order parameter, which we take to be the baryon mass.

We will now try to sketch out a phase diagram in the g^2-N_{F} plane for the system on the basis of our knowledge of the β functions for large and small coupling and of the order parameters at large g. The reader is warned that what follows is rank speculation. We merely try to construct the simplest picture that is consistent with all the data at our disposal.

We begin with very large N_{F}. Here β_{T} is negative for large g^2 and positive for small* g^2 indicating a zero at some intermediate value $g_{c_1}^2(N_{\mathrm{F}})$. The zero is ultraviolet stable so the string tension vanishes at this point and confinement disappears. For small g^2 we have a "Coulomb" phase as shown in the appendix. As far as the tension is concerned we seem to have a standard second-order transition. Now consider β_m. (We know that chiral symmetry is broken for large g^2 and N_{F} so it is meaningful to talk of β_m.) Since β_m is positive for both large and small g^2, it is simplest to assume that it has no zeros. However, it is not reasonable to allow the spontaneously induced fermion mass to persist below some finite $g_{c_2}^2(N_{\mathrm{F}})$ for the renormalization group formula implies that it diverges at $g^2 = 0$. Such behavior would imply that perturbation theory (for say the $\bar{\psi}\psi$ two-point function on the lattice) was not even asymptotic at $g^2 = 0$. This can probably be rigorously proven

* At small g^2 we know that the variation of all gauge-invariant Green functions with the lattice spacing can be compensated by a universal change in g accompanied by wave function renormalizations and local subtractions. This implies that any dynamically induced mass must scale according to the universal β function [eq. (1)] up to terms of $\mathrm{O}(g^7)$.

to be false. Combined with the strict positivity of β_m the vanishing of m below g_{c_2} implies that $g^2_{c_2}(N_F)$ is a line of first-order phase transitions. Since the fermions in this theory do not feel confining forces we have no prejudices about the relative values of g_{c_1} and g_{c_2} [16]. The two lines of transitions may even coincide or cross at some particular values of N_F.

There is a satisfying intuitive explanation for the apparent first-order nature of the chiral transition in gauge theories. First-order transitions are generally caused by instabilities: an infinitesimal ordering causes the system to want to become more ordered. Chiral symmetry breaking is caused by the condensation of virtual fermion-antifermion bound states. If the bare coupling is small and N_F is large, then the effective coupling never gets large enough to bind these states because it is screened by fermion vacuum polarization. As the bare coupling is raised, a point comes when the constituents of the condensate are barely bound, giving mass to the fermions. This reduces the vacuum polarization, thus increasing the effective coupling, which increases the condensate and thus the fermion mass. But this reduces the vacuum polarization–a classic instability. We find this explanation quite convincing.

Let us now move down to small values of N_F. Here the weak coupling β function is negative and shows no signs of a zero, while β_T is also negative for large g^2. It is consistent to assume that the string tension is a smooth function which vanishes only at zero.

If we now raise N_F for fixed small g^2, we pass from this confining phase into the "Coulomb" gluon phase we have discussed above. We believe that this signals another line of phase transitions. Although N_F is not the usual sort of thermodynamic variable (physical theories are defined only at discrete values) we have not been able to think of any system in which one can pass *analytically* between two distinct phases by varying the number of fields or the dimension of space.

An additional argument for such a line of transitions comes from the infrared stable zero that we have investigated in sect. 2. For $g^2 \ll 1$ β_T should coincide with the weak coupling β function and so for $N_F \leqslant N^*$ it has a weak coupling zero. Thus it will vary as a function of N_F according to fig. 1. Note that the nontrivial ultraviolet stable fixed point represents the continuation of the 2nd order transition which we have argued to exist for large N_F.

The tension-related phase transitions of this model are summarized in fig. 2. Note that whereas we have argued that the phase transition is second order on the right-hand portion of the curve, it is first order on the left-hand portion. The string tension blows up at the transition and then drops to zero. The changeover between first- and second-order behavior occurs at the point (g_0, N_F^0) (fig. 2) at which the two zeros of β_T coalesce.

Now let us examine the chiral properties of the model for small N_F. In ref. [5] it was shown that at large g^2, the model has chiral symmetry breaking for $N_F > 2$ but not for $N_F = 2$. It was conjectured that chiral symmetry remains unbroken at $N_F = 2$

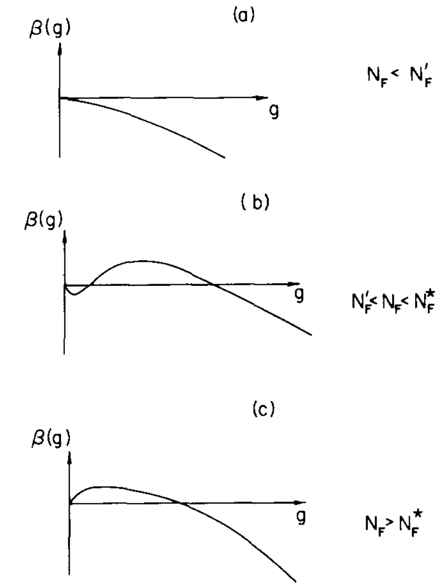

Fig. 1. 2 zero β function.

for all g^2. This indicates the need for a line of phase transitions separating $N_F = 2$ and 4 for large g^2. The simplest way to complete the phase diagram is to connect this line to the line of chiral transitions we have found for large N_F (fig. 3). It is tempting to identify the crossing point of the chiral and confinement transition lines with (g^0, N_F^0) but we have found no convincing argument that this is so. In fact, we are not even sure that the lines should cross.

Fig. 3 is consistent with all the data we have on the model, and as far as we can tell, with general physical principles. If it is correct it has some interesting consequences. Perhaps the most important is the implication that the asymptotically free continuum theory is confining and has composite massless fermions for all $N < N^*$. This goes much further than the results of ref. [5], and implies that the phenomenon noted there is quite general.

Another result of fig. 3 is that there is no continuum theory (for the class of fermion representations to which it applies) with both chiral symmetry breaking

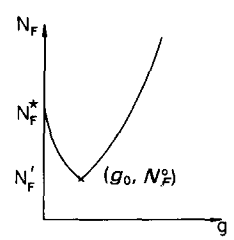

Fig. 2. Tension related transitions in adjoint model.

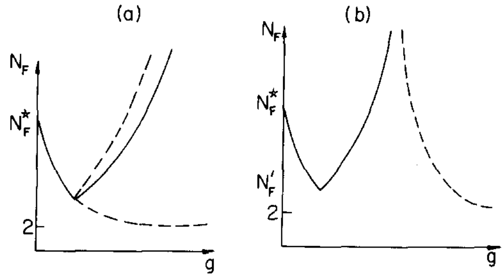

Fig. 3. Tension and chiral transition in adjoint model. (Dashed line: line of chiral transitions.)

and nontrivial ultraviolet structure. The fermion mass does not vanish at the non-trivial ultraviolet stable zero, g_T^* of β_T, unless the lines of transitions do not cross, in which case chiral symmetry breaking does not occur at the fixed point. In the case that the lines do cross, the continuum theory defined at g_T^* would be rather strange. The fermions must decouple because their mass goes to infinity with the cutoff – so the theory is a pure gauge theory with power-law violations of scaling at short distances. Perhaps this is an argument that fig. 3b is preferable to fig. 3a.

Turning now to the case of fermions in the fundamental representation, we cannot define a string tension because any external color density can be screened within a finite number of lattice spacings. However, we still expect a phase transition associated with the gluonic sector of the theory. For very large N_F and small g^2 we expect that the theory contains massless gluons (see appendix) while for large g^2 no such particles exist. We would also like to believe that for small enough N_F, there are no massless gluons even as $g^2 \to 0$ (we are sure that we aren't alone in this piece of wishful thinking). Thus we should expect the massless gluon phase to be separated from the rest of the phase diagram by a line of the same general shape as the solid line in fig. 3.

For large N_F the discussion of the chiral order parameter is almost the same as that for the case of adjoint representation fermions. The only difference is that since the fermions in the strong coupling phase feel confining forces, we believe that the chiral order parameter will be non-zero as long as the transition to a massless gluon phase has not occurred [16]. Therefore, the dashed line (chiral transition) in fig. 4 is definitely not to the right of the solid line (transition to massless gluons).

For small N_F the chiral properties of fundamental and adjoint fermions are completely different. Chiral symmetry breaking occurs at large g^2 for every integer value of N_F. The chiral β function β_m is positive for large g^2 and negative for small g^2 (if we assume chiral symmetry breaking persists for small g^2). Thus it has an infrared-stable zero at some intermediate g^2 at which the order parameter diverges. We believe that this indicates a first-order transition* at which the order parameter

* A. Aharony has emphasized to one of us (TB) that divergent masses at a fixed point are an indication of first-order transitions.

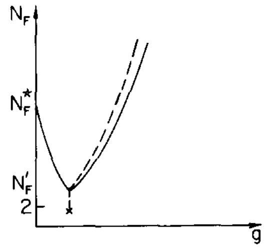

Fig. 4. Gauge and chiral transition in fundamental model.

jumps without vanishing. Alternatively we can assume that chiral symmetry breaking vanishes at some intermediate g^2 leading to a first-order transition at which the chiral ordering drops suddenly. It cannot be second order because β_m has no ultraviolet-stable zeros.

To distinguish between these alternatives we analytically continue the ground state energy to non-integer values of N_F. This can be done because N_F appears analytically in the functional integral representation of the ground-state energy after integrating out the fermions. $N_F = 0$ is the pure gauge theory and we believe that its ground-state energy is an analytic function of g^2. Therefore, we cannot allow the line of singularities that we have claimed to exist (in both scenarios) for physical values of N_F to extend down to arbitrarily small values. Rather, it must end, as indicated in fig. 4. Since the weak and strong coupling regions are analytically connected (albeit through unphysical values of N_F) we conclude that chirality is broken in the weak coupling region.

This is a welcome result, for we expect the weakly coupled SU(3) gauge theory with 2 fundamental fermions to be a good approximation to the real (chirally asymmetric) world. Note, however, that our results have unhappy consequences for strong coupling expansions – they imply a first-order phase transition for every physical value of N_F. Existing strong coupling calculations for $N_F = 2$ [14] seem to provide an analytic extrapolation from large to small g^2. However, in these calculations an extra operator was added to the hamiltonian which spontaneously breaks chirality at zeroth order in g^2. This operator was added to avoid the inconvenience of degenerate perturbation theory and to prevent some unpleasant level crossings in particle energies. It was argued to be "irrelevant" at the weak coupling fixed point. If our arguments are correct it may have been "crucial" to the avoidance of a phase transition at intermediate coupling.

If the weak coupling, small-N_F phase has chiral symmetry breaking, then the chiral order parameter blows up along the line of fixed points of β that we have found in the $N^* - N_F$ expansion. This presumably indicates the first-order transition into the chirally symmetric (see appendix) large-N_F, small-g^2 phase. It is reasonable to assume that the massless gluons also appear as we cross this line although we have no really tight argument that this is so. If it is, then the full phase diagram looks like fig. 4.

This is the most simple phase diagram consistent with what we know. However, there is a hand-waving argument that the massless gluon phase should not coexist with chiral symmetry breaking. The point is that this phase is assumed to be unstable against infrared fluctuations for arbitrarily small g^2 in the pure gauge theory. If the fermions acquire a dynamical mass then gauge field fluctuations with wavelengths longer than the fermion Compton wavelength should not be screened and the massless gluon phase is unstable.

In the theory with fundamental fermions this immediately implies that the dotted and solid lines in the upper part of fig. 4 coincide. It also provides the rationale for assuming that the massless gluon and chiral phase transitions coincide as N_F is decreased for small g^2. In the theory with adjoint fermions the above argument reinforces our contention that fig. 3b is preferable to fig. 3a. The possibility that the dotted and solid lines coincide in the upper part of fig. 3b is still left open.

There are several discrepancies between the present analysis and that of the last section. For example, there we argued that the infrared behavior of the weakly coupled gauge theory between $N_F = N^*$ and $N_F = N'$ was given by the $N^* - N_F$ expansion. As a consequence the theory in this region had no chiral symmetry breaking, no confinement no mass gap, and no particle interpretation, in disagreement with figs. 3, 4.

We believe that the present analysis is correct. In our scenario for the fundamental representation chiral symmetry breaking occurs. The fermions acquire mass and do not contribute to vacuum polarization at sufficiently low energies. Thus it is inconsistent to use the *massless* fermion β function to study the infrared behavior of the theory and determine whether or not chirality is broken. Its use already prejudges the issue.

To put this in a more mathematical and general way, a fixed point of the renormalization group equations only determines the asymptotic behavior of Green functions which are not singular at the fixed point. (This should be familiar to practitioners of perturbative QCD. The $m_q = 0$, $g^2 = 0$ fixed point only determines the asymptotic behavior of quantities which are free of mass singularities.) If the string tension or chiral order parameter are generated they must [eq. (22)] be singular at the fixed point. These singularities propagate to all Green functions and the renormalization group equations do not determine the infrared asymptotic behavior.

We believe that the fixed point (6) and its associated anomalous dimensions *do* determine the infrared behavior of the chirally symmetric nonconfining phase between $N_F = N'$ and N^*. Above N^* the infrared behavior of this phase is like that of massless free field theory modified by logs. We don't see any reason on earth to be interested in these anomalous dimensions, but they can be calculated in the $N^* - N_F$ expansion.

Finally, note that statistical mechanical folklore has it that the signal for first-order phase transitions in renormalization group analyses is that the critical index $\alpha = -1$.

We have claimed that the line of non-trivial zeros is a line of first-order transitions, but our value of α is always >2. In fact it is easy to see from eqs. (7)–(9) that this bound is valid for any infrared-stable point. Of course the point is that the folklore refers to the value of α at an ultraviolet-stable fixed point. We have no right to use the hyperscaling relation $\alpha = 2 - d\nu$ at an infrared-stable point where masses are not zero.

4. Conclusions

By combining a systematic expansion around the weak coupling, infrared-stable fixed point with information from strong coupling and large-N_F expansions, we have made a conjectural map of the phase diagrams of lattice gauge theories with fermions. We cannot sufficiently emphasize the flimsiness of the arguments on which these diagrams stand. We have often resorted to criteria of simplicity and "reasonableness" when no other arguments were available.

Two particular points that we have completely ignored are the possibility of chirally symmetric phases with massive parity doublets and possible singularities of the β functions. Parity doubling is ruled out by the anomaly constraints [17] in the continuum theory, but on the lattice there are no such restrictions. Poles and discontinuities in the β function are expected to be absent for general reasons [13]. However, our strong coupling β functions are not really a complete renormalization group and we do not know if these arguments apply to them. *A priori* we can see nothing wrong with a scenario in which the baryon mass rises as g^2 decreases from ∞, reaches a maximum and then falls to zero at $g^2 = 0$. This would lead to a pole in β_m.

The region of the phase diagram which would be most affected by such poles is the line of transitions with a small number of fermions in the fundamental representation. A pole in β_m would eliminate the *raison d'etre* for this line. We therefore consider it the most suspicious part of our scenario.

If the phase diagrams are correct, however, then they teach us many interesting things:

(i) The absence of chiral symmetry breaking that was found [5] in strong coupling for certain theories with real fermions and $N_F = 2$ is a feature of these theories for all $N_F < N^*$ at the weak coupling fixed point which defines the continuum theory. (Note that our arguments did not use the fact that the fermions were in the adjoint representation and are valid for all real representations with the same qualitative chiral properties at strong coupling. Note also that the 't Hooft anomaly constraints [17] are trivially satisfied by assuming the appropriate multiplicity of "screened" fermions in the bound state spectrum.)

(ii) There seems to be no way to define a continuum gauge theory with non-trivial ultraviolet behavior and chiral symmetry breaking. More generally, we find that the chiral phase transition is generically first order. We have given an intuitive explanation of this in terms of the Nambu–Jona-Lasinio mechanism and vacuum polarization.

(iii) One must be very careful in using an infrared-stable zero of the β function in a massless theory to predict the low-momentum behavior. These predictions can be upset by spontaneously generated masses. This is in marked contrast to ultra-violet-stable points where masses are unimportant.

(iv) Gauge theories with fundamental fermions have a phase transition at which the chiral order parameter is discontinuous (but does not go to zero). This implies problems for strong coupling analyses of these systems. However, these problems may be avoided by adding "irrelevant" operators to the strong coupling hamiltonian.

It is clear that much work remains to be done to test our conjectured phase diagrams. The most straightforward way to approach this problem is via systematic strong coupling and/or large -N_F expansions. We hope to be able to present the results of such calculations in the future.

Appendix

THE LARGE-N_F EXPANSION

In the text we presented renormalization group arguments that the large-N_F, small-g^2 region of the theory is chirally symmetric and contains massless fermions and gluons. Here we want to derive the same result by performing a large-N_F expansion with $g^2 N_F = \lambda$ fixed and $\ll 1$. The large-N_F expansion should be valid for all λ but we have not yet analyzed the intermediate coupling region.

We write the partition function of our lattice gauge theory as

$$Z = \int \prod_{x,\mu} [\mathrm{d}U_\mu(x)] \exp\left[N_F\left(\sum_{x,\mu} \frac{1}{\lambda} \operatorname{tr}(U_{\mu\nu}(x) + U^+_{\mu\nu}(x)\right) + \operatorname{tr}\ln D(U)\right], \quad (A.1)$$

where $D(U)$ is the operator

$$D(U)_{x,x'} = i \sum_\mu \alpha_\mu(x)[U_\mu(x)\,\delta_{x,x'+\mu} - U^+_\mu(x)\,\delta_{x,x'-\mu}], \quad (A.2)$$

$$\alpha_\mu(x) = (-1)^{x_\nu}, \qquad \nu \text{ follows } \mu \text{ in cyclic order}, \quad (A.3)$$

For large N_F the integral can be done by the method of steepest descent and for small λ the stationary points are

$$U_\mu(x) = V(x)V^+(x+\mu). \quad (A.4)$$

To study the fluctuations around these points we write

$$\det(D) = \int \mathrm{d}\bar\psi\, \mathrm{d}\psi\, \mathrm{e}^{\bar\psi D \psi}, \quad (A.5)$$

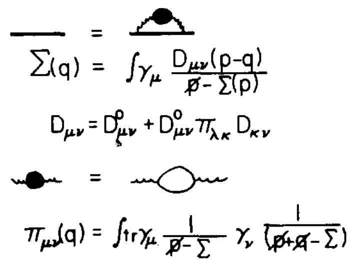

Fig. 5. Schwinger Dyson equations.

$$U_\mu(x) = V(x) W_\mu(x) V^\dagger(x+\mu)\,, \tag{A.6}$$

and expand around $W = 1$.

The V integration is seen to be trivial by defining

$$\chi = V^+\psi\,, \qquad \bar\chi = \bar\psi V\,, \qquad \bar\psi D(U)\psi = \bar\chi D(W)\chi\,. \tag{A.7}$$

W and χ are gauge-invariant fields, but we must fix the ambiguity in the definition of W_μ by some "gauge-fixing condition". Then we can write $W_\mu = e^{iA_\mu}$ and expand to leading order in A_μ. The spectrum then consists of free massless fermions and vector bosons whose inverse propagator is the free massless one plus the fermion vacuum polarization diagram of fig. 5. This propagator has no singularity at $p^2 \neq 0$. Since we have the exact solution of the large-N_F theory for λ in some finite range around zero, the claims about this region made in the text are verified.

Above some $\lambda_c = O(1)$ we may imagine that there is a solution of the steepest descent equations with action below that of (A.4), leading to a phase transition. Since (A.4) always remains stable (essentially because of the screening effect of the fermions: also there are no Landau ghosts on a lattice), this transition would have to be first order.

If the new minimum were not invariant under single unit translations, then it would lead to chiral symmetry breaking. Support for this scenario comes from the $\lambda \to \infty$ limit. In this limit we can solve the theory using the methods of ref. [18]. The U_μ integral splits into a product of single link integrals which can be performed exactly (though in a useful form only for very small or very large gauge groups) [19]. This gives an effective action for ψ which we write in terms of auxiliary fields. In the large-N_F limit the dynamics of the auxiliary fields is that of a generalized antiferromagnet and breaks chiral (single unit translation) symmetry.

There is, however, an alternative scenario which is also consistent with these results. The lines of phase transitions in figs. 3 and 4 might converge to $\lambda \to \infty$ as $N_F \to \infty$. Then we would have to study $1/N_F$ corrections in order to understand the transition. The appropriate tool would be the standard Schwinger-Dyson equation (see fig. 5) (on the lattice)

$$\Sigma = \int D(\Sigma_0 - \Sigma)^{-1} ,$$

with D given by the leading $1/N_F$ propagator including the vacuum polarization. Thus D is itself a functional of Σ. This equation is very interesting for it incorporates the mechanism that we claim makes the chiral transition first order.

We plan to return to the study of the large-N_F limit at all values of λ in a future publication.

References

[1] D. Gross and F. Wilczek, Phys. Rev. Lett. 30 (1973) 1346; Phys. Rev. D8 (1973) 3633
[2] J. Kogut, R. Pearson and J. Shigemitsu, Phys. Rev. Lett. 43 (1979) 484
[3] H.B. Nielsen and M. Ninomiya, Nucl. Phys. B193 (1981) 173
[4] W.E. Caswell, Phys. Rev. Lett. 33 (1974) 244
[5] T. Banks and V. Kaplunovsky, TAUP preprint (1981)
[6] V. Kaplunovsky, in preparation
[7] O.V. Tarosov, A.A. Vladimirov and A. Yu Zharkov, Phys. Lett. 93B (1980) 429
[8] M. Fisher and K.G. Wilson, Phys. Rev. Lett. 28 (1972) 240
[9] E. Brezin, J. Zinn-Justin, Phys. Rev. Lett. 36 (1976) 691; Phys. Rev. B14 (1961) 4976
[10] E. Brezin, J. Zinn-Justin and J.C. Le Guillon, Phase transition and critical phenomena, vol. 5a, ed. C. Domb and M.S. Green (Academic Press)
[11] M. Peskin, Phys. Lett. 94B (1980) 161
[12] E.G. Floratos, D.A. Ross and C.T. Sachrajda, Nucl. Phys. B129 (1977) 66
[13] K. Wilson and J. Kogut, Phys. Reports 12 (1974) 75
[14] T. Banks et al. Phys. Rev. D13 (1977) 1111
[15] L. Susskind, Phys. Rev. D16 (1977) 3031
[16] T. Banks and A. Casher, Nucl. Phys. B169 (1980) 103
[17] G. 't Hooft, Cargèse Lectures, 1979
[18] H. Klugberg-Stern et al., Saclay preprint (1981)
[19] I. Bars and F. Green, Phys. Rev. D20 (1979) 3311
[20] D. Gross and E. Brezin, Phys. Lett. 97B (1980) 120