PHYSICAL REVIEW D VOLUME 44, NUMBER 4 15 AUGUST 1991

Coupled fermion gap and vertex equations for chiral-symmetry breakdown in QCD

Joannis Papavassiliou* and John M. Cornwall
Department of Physics, University of California, Los Angeles, California 90024-1547
(Received 21 February 1991)

The fermion gap equation for QCD is usually posed as a simple variant of the Baker-Johnson-Willey equation, with *ad hoc* cutoffs for infrared singularities in the running charge imposed as needed, and no confinement effects taken into account. While we too omit confinement effects, we replace the *ad hoc* cutoffs with a physically consistent picture of QCD, with a dynamically generated gluon mass m serving as an infrared regulator, and study not only the fermion gap equation but also nonlinear vertex equations which determine the running charge. Fermion loops are included in the vertex equations, which allows us to study the dependence of $\alpha_s(q^2)$ on the fermion constituent mass M. As one might expect from the one-loop running charge, $\alpha_s(0)$ is an increasing function of M. When we use a standard form of the fermion gap equation and leave out all confinement effects, the relations between $\alpha_s(0)$ and M which follow from the gap equation and the vertex equation are inconsistent with each other, for any value of the gluon mass m. Although our model is crude, this suggests that confinement plays an important role in chiral-symmetry breakdown in QCD. Furthermore, lack of confinement above the deconfining temperature may explain why the chiral-symmetry-restoration temperature is so close to the deconfining temperature.

I. INTRODUCTION

Countless papers have been written on chiral-symmetry breakdown (CSB) as embodied in fermion gap equations in gauge theories since the days of Baker, Johnson, and Willey (BJW) [1]. There have been applications to fixed-point (QED-like) theories [2], to QCD [3–10], to technicolor [11], and to gauged Nambu–Jona-Lasinio models [12]. For the most part, these works are based on simple modifications of the original BJW work, which was intended to apply to weakly coupled fixed-point QED with emphasis on control of the ultraviolet behavior. Thus the gap equation is studied in the Landau gauge, where ultraviolet vertex corrections are unimportant, and even for a confining theory such as QCD only single-gluon forces are kept, with no reference to the confining forces.

However, since QCD is not a fixed-point theory, certain modifications are needed to incorporate a running charge and asymptotic freedom. Here is where the trouble starts, because the perturbatively defined running charge is infrared singular. Different authors have used different sorts of cutoffs, usually prescribed in an *ad hoc* way; for example, Higashijima [4] uses the perturbative running charge down to a certain momentum, then stops the running, while Atkinson and Johnson [6] add what amounts to a gluon mass term in the running charge but keep the gluon propagator massless. Haeri [8], as well as Haeri and Haeri [10], consistently invoke both a running charge and a gluon propagator modified by gluon mass generation [13]; thus their infrared cutoff is not *ad hoc,* but physically motivated. We will do the same in this work. Various other *ad hoc* prescriptions occur in the literature [7,9].

Another kind of infrared singularity is important, both for QCD and for fixed-point theories; this is associated with vanishing fermion mass. Such a singularity occurs in the usual gap equation when it is linearized, so that both the gluon and the fermion propagators are massless. One obvious way [14] to get rid of this kind of fermion infrared singularity is to keep the fermion self-energy in the denominator of the fermion propagator, which we will term the nonlinear BJW equation. However, this is not the only kind of fermion-mass infrared singularity. Another kind appears in the running charge itself, and comes from fermion loops (i.e., the unquenched theory). Of course, if CSB does occur then the fermion mass M is not zero, and no singularity arises. There is an important difference between would-be gluonic and fermionic infrared mass singularities; the running charge $\alpha_s(q^2)$ at small momentum decreases as the gluon mass increases, and increases as the fermion mass increases.

Similarly, the fermion gap equation shows that as $\alpha_s(0)$ is increased, larger values of M result; for $\alpha_s(0)$ less than a critical value α_c, there is no CSB (i.e., $M=0$). There are thus two relations between $\alpha_s(0)$ and M, one from the gap equation and one from the running charge, and it is our purpose here to study, in a physically motivated picture of QCD, these two relations to see if they are consistent. In order to furnish a basis for comparison with earlier work, we will use a standard form of the gap equation, and omit all references to confinement.

Our new contribution is an analysis of some coupled nonlinear Schwinger-Dyson equations which determine the running charge $\alpha_s(q^2)$ and its dependence on M. We have already studied these equations for quarkless QCD [15], and will not describe the program leading to them in full detail. It is enough for now to say that the vertex equations are highly simplified versions of real-world Schwinger-Dyson equations which have the crucial prop-

erties of being gauge invariant and usable at all momenta from 0 to ∞. (The usual way of introducing the running charge via the renormalization group is, of course, useful only at large momentum.)

If, in these equations, the gluon is kept massless the solution necessarily has a spacelike singularity, as one would have expected from the one-loop running charge. Moreover, the β function which can be extracted from the vertex is a sum to all orders in g^2, with the coefficients all of one sign and showing factorial growth (as is known to happen from other considerations [16]). The cure for these singularities is that a gluon mass is spontaneously generated, as was shown some time ago [13] from a study of the gauge-invariant gluon propagator, constructed with the same techniques used later [15] to calculate gauge-invariant gluon vertices. Various considerations, including the original mass-generation study [13] as well as fitting [15] the vertex-equation solution to $\alpha_s(0)\simeq 0.5$, suggest that the gluon mass m is around 500 MeV. However, for our purposes here we will let m be a free parameter (it is not determined by the vertex equations alone), which we can tune to adjust $\alpha_s(0)$. Obviously decreasing m results in increasing $\alpha_s(0)$. In fact, there is a critical value m_c below which $\alpha_s(q^2)$ becomes singular.

The running charge receives contributions from fermions as well as gluons, so in this paper we modify the vertex equations to account for fermionic terms. Fermion loops, like gluon loops, are infrared singular at $M\simeq 0$ except that they contribute with opposite sign: Decreasing M decreases $\alpha_s(0)$.

The study of the vertex equation, then, leads us to one relation between M and $\alpha_s(0)$, at fixed gluon mass m. A second relation comes from the fermion gap equation, which has the same general trend: The smaller $\alpha_s(0)$ is, the smaller the fermion mass M, and finally, as mentioned above, for $\alpha_s(0)$ less than some critical value, CSB becomes impossible.

The question now is whether the two relations we have discussed are consistent. One may imagine the following possibilities.

(1) The relations are the same, and give no information beyond the gap equation itself. This would be an inexplicable miracle.

(2) They are consistent, and yield (for fixed m) determinate values for M and $\alpha_s(0)$ rather than only one relation between them.

(3) They are inconsistent. Insofar as we can trust our crude models, the third alternative is what we find.

The reason for inconsistency is simply that the gluon mass m required to achieve CSB in the fermion gap equation is so small that the vertex equation is singular, i.e., $m<m_c$. Thus there is no self-consistent way, within the context of the model, to make the running charge large enough to account for CSB. In fact, the situation is even worse: Not only is there no CSB, but with the fermionic mass $M=0$, QCD itself fails because of the fermionic infrared singularities in $\alpha_s(0)$.

So the model we use is wrong; where does it fail? We believe the answer is that confinement has not been accounted for; confining forces will certainly break chiral symmetry, as we discuss below. To be fair, there are many other deficiencies in both the gap equation and the vertex equation which could lead to the failure of the model. For example, the kind of BJW equation we (and others) use is best suited to dealing with ultraviolet momenta, not infrared; for small momenta in gauge theories, there is another approach [8,10,17] based on the gauge technique. The gauge technique is accurate in the infrared in its simplest form and can also be made accurate in the ultraviolet, as King [17] and Haeri [8] have shown. Moreover, a recent investigation [10] of the ultraviolet-improved gauge-technique gap equation, which uses massive gluons as we do, leads to one of our conclusions: There is no CSB for the gluon mass larger than about 0.1 GeV, far smaller than the physical value. (These authors do not study the vertex equation, as we do here.) The vertex equations we use are highly simplified, but they work so well [15] in both the ultraviolet and the infrared regimes that we have no real reason to suspect that they are the cause of the inconsistencies we have found. As for the fundamental aspect of the gluon mass, numerous lattice workers [18] have confirmed the existence of such a mass, with a value at least as great as the original estimate [13] of 500 MeV.

One should note that quite aside from any of our present considerations, the critical values of $\alpha_s(0)$ quoted by several workers from the fermion gap equation are substantially larger than one finds either phenomenologically or on the lattice. Higashijima [4] gives $\alpha_c=0.74$, and Atkinson and Johnson [6] give $\alpha_c=1.2$. But charmonium phenomenology [19] yields $\alpha_s(0)\simeq 0.5$, and recent lattice calculations [20] give $\alpha_s(0)=0.45$. A gluon mass of 500 MeV also yields $\alpha_s(0)\simeq 0.5$, when used [13] in our vertex equations. So it might not be surprising that single-gluon exchange is not enough to drive CSB. Nevertheless, it is worthwhile to examine, as we do, whether there can be consistency between CSB and the Schwinger-Dyson equations that determine the running charge, quite aside from what appear to be reasonable values for $\alpha_s(0)$ from lattice or phenomenological considerations.

This is not the place to discuss in detail how confinement might repair this inconsistency, but we will review very briefly some of the salient points in this regard as well as what appears to be the reason for why confinement is widely neglected nowadays in CSB.

A decade or so ago, several authors [21–23] pointed out that confinement was sufficient to ensure CSB. Unfortunately, others, in referring to these works, sometimes interpreted them as saying that confinement was necessary for CSB, which is quite a different thing.

These ideas were soon challenged by computer simulations in quenched SU(2). Kogut *et al.* [24] claimed that the essential parameter governing CSB was the product g^2C_F, where g^2 (at an appropriate scale) is the gauge coupling constant, and $C_F=I(I+1)$ is the Casimir eigenvalue for the fermions of isospin I which developed a condensate ($\langle\bar{\psi}\psi\rangle\neq 0$). It is, of course, natural to interpret this product as measuring the strength of short-range single-gluon forces and to claim that such forces are dominant in CSB. Kogut *et al.* [24] also found CSB both for

fermions which could be confined [isospin $N+\frac{1}{2}$ for SU(2), where N is an integer ≥ 0] and those which were not (integral isospin), adding weight to the case against an important role for confinement.

However, things are not so simple. Even for unconfined fermion representations there are stringlike forces which rise linearly for a while, then break with the formation of a gluon pair [25]. For $I=1$, the string-breaking energy to form such a pair is quite comparable to the energy at which the $I=\frac{1}{2}$ string breaks because of formation of a $q\bar{q}$ pair. Moreover, the strength of the linearly rising string forces grows with increasing I (although not at the same rate as for single-gluon exchange). The simplest model of the string strength envisages a particle of isospin I as being in effect composed of N quarks, where $N=2I$. Such a particle will be joined to a similar one by N strings, so the string force should be proportional to N or to $2I$. The increase of the string tension with I has been seen numerically [26] and explicitly modeled [27] in a picture of confinement [28] based on vortices of finite extent [29].

In view of these remarks, it is not clear to us that the results of Kogut *et al.* [24] decisively rule out confining forces as important for CSB. It is likely that both confining forces and short-range single-gluon forces contribute, and indeed the work we report here suggests that single-gluon forces are not enough. While it is not our purpose here to belabor the issue of confinement, it may be worthwhile to review how confinement leads to CSB in one model [23], which uses the time-honored device of imitating confinement with an effective propagator.

$$+g_{\mu\nu}\frac{\mathcal{M}^2}{k^4}+k_\mu k_\nu \text{ terms} , \tag{1.1}$$

where $\mathcal{M}\simeq 1.9$ GeV. We hasten to say that the effective propagator has nothing to do with the true gluon propagator, which is short ranged because of gluon mass generation [13]; single-gluon effects must be added [30] to the propagator (1.1). Let us leave the latter aside for the moment, and ask what the confining potential does by itself. If we naively calculate the one-loop quark self-energy, the result is ultraviolet convergent but linearly infrared divergent. So we regularize (1.1) to

$$g_{\mu\nu}\frac{\mathcal{M}^2}{(k^2-\mu^2)^3} \tag{1.2}$$

and then find [23], for the one-loop quark mass,

$$M^{(1\text{ loop})}=\frac{\mathcal{M}^2}{16\pi\mu}+O\left[\frac{\mathcal{M}^2}{M}\ln\mu\right] . \tag{1.3}$$

As $\mu\to 0$ the quark mass approaches infinity, which might naively be taken as a signal for confinement. However, the static potential derived from (1.2) has a similar infrared divergence; in a color-singlet $q\bar{q}$ state one finds

$$V(r)=-\frac{\mathcal{M}^2}{8\pi\mu}+\frac{\mathcal{M}^2}{8\pi}r+O(\mu) \tag{1.4}$$

so that the physically relevant quantity

$$2M+V(r)=\frac{\mathcal{M}^2 r}{8\pi}+O\left[\frac{\mathcal{M}^2}{M}\ln\mu\right] \tag{1.5}$$

is free of linear infrared divergences. There is no such cancellation for quark combinations which are not color singlets [31]. The remaining logarithmic divergence is replaced by a finite quantity, after the recognition that calculating the quark self-energy with a loop graph containing a free-quark propagator is unjustified when confining forces are present. Various ways of getting around this difficulty were discussed in Ref. [23], with the upshot that the effective quark mass M is approximately $\mathcal{M}/2\pi\simeq 300$ MeV. The only importance of this formula to us is that it shows that confinement does not yield a typical gap equation, with M appearing on both sides, but rather that the mass arises directly from the string tension.

In this paper we restrict our investigation to QCD with two light flavors. Actually, the original motivation for a study of this sort was to apply it to technicolor or grand-unified theories, especially those with a small coefficient b in the β function ($\beta=-bg^3+\cdots$), coming from a near cancellation of fermions and scalars with gluons. Here, on the one hand, the running charge (proportional to b^{-1}) can be large without confinement; on the other hand, the vertex equation is very sensitive to changes in the fermion masses because of the near cancellation. We will return to the study of CSB in such theories on a later occasion.

Our paper is organized as follows. In Sec. II we derive in a heuristic way the coupled Schwinger-Dyson equations for the QCD vertices in the presence of quarks and argue that they successfully reproduce the renormalization-group (RG) results as well as the usual Feynman-diagram infrared divergences. Section III is devoted to a qualitative, analytical study of the properties of the vertex equations with particular emphasis on the role of the quarks. In Sec. IV we report the results from solving the vertex equations numerically. Finally, in Sec. V we incorporate our solutions into some standard versions of the fermion gap equation and study the compatibility of the whole program.

II. SCHWINGER-DYSON EQUATIONS FOR VERTICES

In QCD there are two fundamental proper vertices: The three-gluon vertex $\Gamma^{abc}_{\alpha\beta\gamma}$ and the quark-gluon vertex Γ^a_α. Each of these is gauge dependent and not a suitable candidate for further investigation. There are two nearly equivalent ways to modify these vertices into gauge-independent quantities. The first way [13,15], which is systematic and in principle exact, extracts from the S matrix or similar gauge-invariant quantity all graphs having a vertex structure plus pieces of other graphs which come from four- and higher-point functions. These pieces also have a vertex structure which arises from pinching out internal lines through elementary Ward identities; it has been thoroughly described and successfully applied to the calculation of the modified three-gluon vertex $\hat{\Gamma}^{abc}_{\alpha\beta\gamma}$ at one-loop level in Ref. [15]. Not only is $\hat{\Gamma}$ gauge invariant, it also obeys a Ward identity involving the gauge-invariant self-energies $\hat{\Pi}^{ab}_{\mu\nu}$ introduced in Ref. [13]. An

essential element of the construction of this or any other gauge-invariant vertex is the decomposition of the elementary three-gluon vertex into two parts, one of which generates a naive Ward identity on one (and only one) line in any convenient covariant gauge, and the other, called the pinch part, generates vertex contributions which cancel either with other gluon graphs or with ghost graphs. (A slightly different procedure, leading to the same results, applies for ghost-free gauges [13].) The decomposition of the bare vertex $\Gamma^{(0)}_{\alpha\beta\gamma}$ (we drop an overall structure-constant factor $i\epsilon_{abc}$) is, for the kinematics of Fig. 1,

$$\Gamma^{(0)}_{\mu\nu\alpha}=\hat{\Gamma}^{(0)}_{\mu\nu\alpha}+\xi^{-1}[k_\mu\Delta_{\alpha\nu}+(q+k)_\nu g_{\alpha\mu}] ,$$

$$\hat{\Gamma}^{(0)}_{\mu\nu\alpha}=-(2k+q)_\alpha g_{\mu\nu}+2q_\mu g_{\nu\alpha}-2q_\nu g_{\mu\alpha} \quad (2.1)$$

$$+(1-\xi^{-1})[k_\mu g_{\alpha\nu}+(q+k)_\nu g_{\alpha\mu}] . \quad (2.2)$$

Here ξ is a gauge parameter, defined in the bare gluon propagator:

$$d_{\mu\nu}(k)=\frac{1}{k^2}\left[-g_{\mu\nu}+(1+\xi)\frac{q_\mu q_\nu}{q^2}\right] . \quad (2.3)$$

One easily verifies that

$$q^\alpha\hat{\Gamma}^{(0)}_{\mu\nu\alpha}=d^{-1}_{\mu\nu}(q+k)-d^{-1}_{\mu\nu}(k) \quad (2.4)$$

and that the purely longitudinal pinch part [difference between $\Gamma^{(0)}$ and $\hat{\Gamma}^{(0)}$ in (2.1)] pinches out lines in a graph to which it is attached.

We can also use $\hat{\Gamma}^{(0)}$ to construct modified fermion-gluon vertices. The case where the fermion lines are on shell and the gluon line is off shell has been discussed in Ref. [32], and the details of the construction for all lines off shell will be given in a separate report by one of us (J.P.). It turns out that for our purposes it is enough, at the one-loop level, to use the graphs of Fig. 2 in the Feynman gauge, so that at this level we define

$$\hat{\Gamma}^{a}_{\alpha}=\frac{\tau^2 ig^2}{(2\pi)^4}\left[\left[\frac{1}{2}C_A+\frac{N}{d_f}C_f\right]\times\int\frac{d^4k\,\gamma^\rho S(p+q-k)_\alpha S(p-k)\gamma_\rho}{k^2}+\left[\frac{1}{2}C_A\right]\int d^4k\frac{\gamma^\mu S(p-k)\gamma^\nu\Gamma^F_{\mu\nu\alpha}}{k^2(k+q)^2}\right] , \quad (2.5)$$

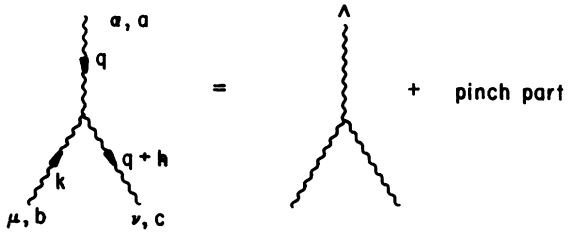

FIG. 1. Decomposition of the bare three-gluon vertex into a part $\hat{\Gamma}$ satisfying a Ward identity on the line marked with a Λ, plus a pinch part.

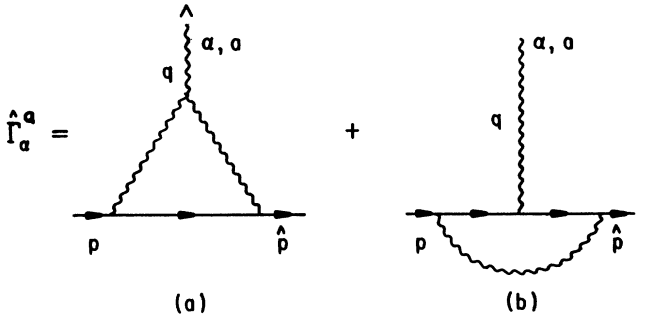

FIG. 2. One-loop graphs for the modified fermion vertex, using the gluon vertex of Fig. 1.

where τ^a is the fermion representation matrix, d_f is its dimension, C_f its Dynkin index, and $\Gamma^F_{\mu\nu\alpha}$ [15,32] is given by $\hat{\Gamma}^{(0)}_{\mu\nu\alpha}$ of (2.2) in the Feynman gauge $\xi=1$:

$$\Gamma^F_{\mu\nu\alpha}=2q_\mu g_{\nu\alpha}-2q_\nu g_{\mu\alpha}-(2k+q)_\alpha g_{\mu\nu} . \quad (2.6)$$

Note that the pinch part of the three-gluon vertex does not appear. It then follows the $\hat{\Gamma}^{a}_{\alpha}$ obeys the Ward identity

$$q^\alpha\hat{\Gamma}^{a}_{\alpha}=\tau^a[\Sigma(p)-\Sigma(\hat{p})] , \quad (2.7)$$

where $\Sigma(p)$ is the fermion self-energy in the Feynman gauge [actually, (2.7) holds in any gauge, as long as the hatted vertex in Fig. 2 is chosen as in (2.2)]. This same vertex, but for any ξ, was used in Ref. [32] and by Haeri [8] in his study of the fermion gap equation. As Haeri shows, the use of $\hat{\Gamma}^{a}_{\alpha}$ allows us to separate out, in a gauge-covariant way, the fermionic contributions to the running charge from other effects, and we will make use of this important feature in our study of the fermion gap equation.

There is a second approximate but heuristically useful approach to constructing vertices which are gauge invariant both on shell and in the UV region of momenta. One considers, instead of the usual proper vertices, half-proper vertices formed by multiplying each leg of a proper vertex by a momentum- and spin-dependent dimensionless factor which is extracted from the corresponding propagator. To illustrate, we write the fermion propagator in any gauge as

$$S(p)=\frac{Z_2(p)}{(\not{p}-M)} \quad (2.8)$$

and form the half-proper vertex by multiplying each proper-vertex leg by $Z_2^{1/2}(p)$, and each gluon leg by a factor $Z_3^{1/2}(q)$, where Z_3 is defined in terms of the gluon propagator by extracting the pole term in analogy to (2.8). It is easy to verify that the combination of renormalization constants which renormalizes a half-proper

vertex is gauge invariant, and that the UV behavior of this vertex as dictated by the RG is also gauge invariant. Moreover, all gauge-dependent terms in the on-shell half-proper vertex vanish. Even when dealing with the exactly gauge-invariant vertices $\hat{\Gamma}$, we find it convenient to use this half-proper formalism.

In principle, our goal should be to write down exactly the Schwinger-Dyson equations for the half-proper gauge-invariant vertices. In practice, this is far beyond our powers, and we will proceed by an extension of the approximations of Ref. [15]. First, we consider only the vertex equations truncated at one dressed loop, where all vertices and propagators are fully dressed. Truncation of the exact equations at the one-dressed-loop level is already a serious step, since it is not in principle gauge invariant even for group-invariant vertices. However, since we will not deal with the exact one-dressed-loop equations but rather approximations to them, it is easy to enforce gauge invariance.

Second, we will ignore all complications of spin, except for the powers of momentum that they induce. For example, $\hat{\Gamma}_{\alpha\beta\gamma}$ has a term of the form

$$\hat{\Gamma}_{\alpha\beta\gamma}(q_1,q_2,q_3)=[(q_1-q_2)_\gamma g_{\alpha\beta}+\text{c.p.}]\hat{\Gamma}_g \ , \tag{2.9}$$

where $\hat{\Gamma}_g$ is dimensionless. Similarly, a representative term in $\hat{\Gamma}_\alpha$ is of the form

$$\hat{\Gamma}_\alpha=\gamma_\alpha\hat{\Gamma}_f \ , \tag{2.10}$$

with $\hat{\Gamma}_f$ dimensionless; the gauge-independent [13] propagator is of the form

$$\hat{\Delta}_{\mu\nu}(q)=\left[-q_{\mu\nu}+\frac{q_\mu q_\nu}{q^2}\right]\frac{\hat{Z}_3(q)}{q^2-m^2}+\text{gauge terms} \ , \tag{2.11}$$

with m the gluon mass. Finally, we have already discussed the form of the fermion propagator (2.8). To indicate that we are in principle dealing with gauge-invariant propagators and vertices, we use $\hat{Z}_2$ instead of Z_2 as in (2.8).

Now introduce the half-proper vertices

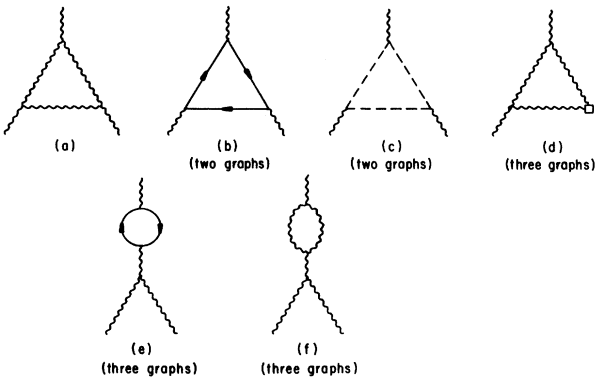

FIG. 3. One-loop contributions to the half-proper three-gluon vertex $\hat{G}$. The open square makes a special vertex formed from a box graph by pinching; see Ref. [15].

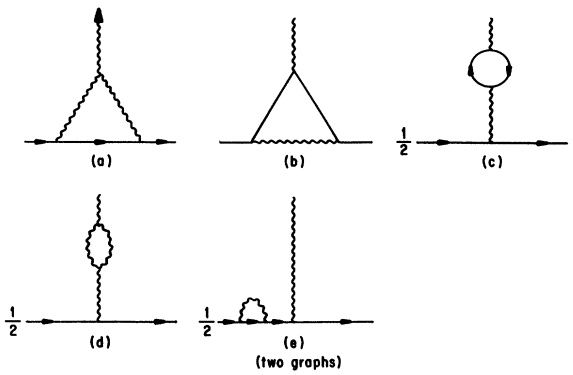

FIG. 4. One-loop graphs for the half-proper fermion vertex $\hat{F}$.

$$\hat{G}(q_1,q_2,q_3)=[\hat{Z}_3(q_1)\hat{Z}_3(q_2)\hat{Z}(q_3)]^{1/2}\Gamma_g \ , \tag{2.12}$$

$$\hat{F}(q_1,q_2,q_3)=[\hat{Z}_2(q_1)\hat{Z}_2(q_2)\hat{Z}(q_3)]^{1/2}\Gamma_f \ . \tag{2.13}$$

Following Ref. [15] we write one-loop equations which respect the dimensionless nature of $\hat{G}$ and $\hat{F}$, but which otherwise ignore spin. The (cubic) integral equations that $\hat{F}$ and $\hat{G}$ satisfy must each lead to the correct UV behavior as dictated by the RG; since this behavior is that $\hat{F}$ and $\hat{G}$ vanish the same way the usual running charge does, it is necessary that there be no inhomogeneous terms in these equations. Moreover, these equations must model the IR singularities which occur when m and/or $M=0$. Since potential IR divergences are very important in what follows, we are careful to construct equations which faithfully represent the one-loop divergences of perturbation theory which arise from the graphs of Fig. 3 (for $\hat{G}$) and Fig. 4 (for $\hat{F}$). A final approximation is that $\hat{F}$ and $\hat{G}$ will be taken to depend only on one momentum, presumably representative of any of the three momenta when these are all nearly equal. The final equations which we postulate, corresponding to the dressed-loop graphs of Fig. 5, are

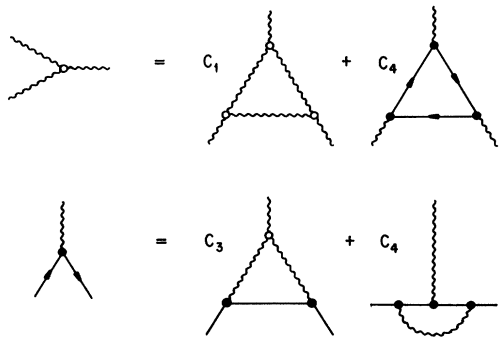

FIG. 5. One-dressed loop graph structure of the Schwinger-Dyson equations for $\hat{F}$ (solid circle) and $\hat{G}$ (open circle). The numbers C_i are given in the text.

$$\hat{G}(q)=\frac{b_1 g^2}{2\pi^2}\int\frac{d^4k\,\hat{G}^3(k)}{(k^2-m^2)[(k+q)^2-m^2]}-\frac{b_2 g^2}{2\pi^2}\int\frac{d^4\hat{F}^3(k)}{(k^2-M^2)[(k+q)^2-M^2]}\,, \tag{2.14}$$

$$\begin{aligned}\hat{F}(q)=&\frac{b_1 g^2}{2\pi^2}\int\frac{d^4k\,\hat{G}^2(k)\hat{F}(k)}{(k^2-m^2)[(k+q)^2-m^2]}-\frac{b_2 g^2}{2\pi^2}\int\frac{d^4k\,\hat{F}^3(k)}{(k^2-M^2)[(k+q)^2-M^2]}\\&+\frac{g^2}{2\pi^2}\left[\frac{1}{2}C_A+\frac{N}{d_f}C_f\right]\int\frac{d^4k\,k^2F^3(k)}{(k^2-M^2)[(k-q)^2-M^2][(k+q)^2-m^2]}\\&-\left[\frac{g^2}{2\pi^2}\right]\left[\frac{1}{2}C_A\right]\int\frac{d^4k\,k^2\hat{G}(k)\hat{F}^2(k)}{(k^2-m^2)[(k-q)^2-m^2][(k+q)^2-M^2]}-\left[\frac{g^2}{2\pi^2}\right]\frac{N}{d_f}C_f\int\frac{d^4k\,\hat{F}^3(k)}{(k^2-M^2)[(k+q)^2-m^2]}\,,\end{aligned} \tag{2.15}$$

where $b_1=11C_A/48\pi^2$ and $b_2-2n/48\pi^2$. C_A is the quadratic Casimir operator for the adjoint representation [for SU(N), $C_A=N$] and n is the number of quark families in the fundamental representation. Throughout this paper we will restrict ourselves to the case of two light quark flavors ($n=2$). It is now easy to see how the above equations in the UV reproduce the known renormalization-group results. Indeed, if we consider large values of q and neglect all masses, the last three terms of (2.15) add up to zero. Then $\hat{G}(q)=\hat{F}(q)$ is the solution of the system, which reduces to one equation, namely,

$$\hat{G}(q)=\frac{(b_1-b_2)g^2}{2\pi^2}\int d^4k\frac{\hat{G}^3(k)}{k^2(k+q)^2}\,. \tag{2.16}$$

This is of course, the same equation we found in Ref. [15], with b replaced by b_1-b_2, and its ultraviolet behavior is that of the usual running coupling. Of equal importance is to realize that Eqs. (2.14) and (2.15) reproduce correctly the infrared divergences of the one-loop Feynman diagrams, when the masses involved go to zero. This fact justifies in retrospect their particular form. To be precise, the first term in (2.14) reproduces the same divergences as graphs $3a+2c+3d+3f$ when $m\to 0$, whereas the second term reproduces the divergences of the graphs $3b+3e$ as $M\to 0$. Similarly, the first term in (2.15) reproduces the divergences of graph $4d$ as $m\to 0$ and the second reproduces the divergences of graphs $4c$ as $M\to 0$. The last three terms of (2.15) correspond to graphs $4a+4b+4e$. It is important to notice that they are infrared divergent if both m and M go to zero. Furthermore, as we already mentioned, their sum vanishes in the UV. This suggests a significant simplification to our program. Since we are mainly interested in the infrared behavior of the running coupling when $M\to 0$, but keeping $m\neq 0$ and of order Λ, we can neglect the last three terms of (2.15), as long as we consider quark masses smaller than the gluon masses. Under this restriction, omission of the aforementioned terms will not affect the behavior of the coupling in either the infrared or the ultraviolet regimes. With these terms omitted $\hat{F}(q)=\hat{G}(q)$ solves the system (2.14) and (2.15), which now becomes one single equation:

$$\hat{G}(q)=\frac{b_1 g^2}{2\pi^2}\int d^4k\frac{\hat{G}^3(k)}{(k^2-m^2)[(k+q)^2-m^2]}-\frac{b_2 g^2}{2\pi^2}\int d^4k\frac{\hat{G}^3(k)}{(k^2-M^2)[(k+q)^2-M^2]}\,. \tag{2.17}$$

Even in this simplified version, the integral equation (2.17) has a rich structure, which we will investigate in the next two sections, both analytically and numerically.

III. QUALITATIVE BEHAVIOR OF VERTEX MASS SINGULARITIES.

In Sec. II we presented drastically simplified vertex equations, based on capturing both the UV and IR behavior of the two equations. In this section, we make some qualitative remarks on just what this IR behavior is, both in connection with IR fermion mass singularities and in connection with a popular approximation in which the integral equation (2.17) is replaced by a different equation.

We can do the angular integrations in this equation, which then becomes

$$G(x)=\int_0^\infty dy\,G^3(y)[b_1K(x,y,m)-b_2K(x,y,M)]\,, \tag{3.1}$$

with the definitions

$$x=q^2,\quad G(x)=g\hat{G}(x), \tag{3.2}$$

$$\begin{aligned}K(x,y,\mu)=&\frac{y}{y+\mu^2}\\&\times\{(x+y+\mu^2)+[(x+y+\mu^2)^2-4xy]^{1/2}\}^{-1}\,,\end{aligned}$$

$$\mu=m \text{ or } M\,. \tag{3.3}$$

To gain some insight on the general properties of Eq. (3.1), we first consider its quarkless version, with $b_2=0$, already discussed in Ref. [15]. In the absence of quarks, (3.1) is scale-free, and clearly G depends on the single variable q/m only. Therefore, at this point, (3.1) either has a solution for *every* positive value of m, or it has no solution at all. We will now try to single out those solu-

tions of (3.1) which have physical interest by imposing certain physically motivated boundary conditions on the original integral equation. It will turn out that these boundary conditions will profoundly affect the general scaling properties of the solutions, as we explain below.

The most general solution $G(x)$ may vanish for finite x, including $x=0$, but such solutions are of no apparent physical interest. For example, if $G(0)=0$, we can hardly expect chiral-symmetry breaking to occur. So, we impose the condition that $G(x)>0$ for every finite positive x, with $G(\infty)=0$. With $G>0$ it is easy to show [see (3.9,10) below] that $G(x)$ is a strictly monotone decreasing function of x (but with quarks present this need not be true). Since the β function is essentially $x\dot{G}(x)$, this monotone decrease ensures that $\beta<0$.

The quarkless equation (3.1) is scale-free, so we temporarily measure all momenta in units of m. For any finite m, (3.1) at $x=0$ reads

$$G(0)=\frac{1}{2}b_1\int_0^\infty dy\frac{yG^3(y)}{(y+1)^2}\ . \tag{3.4}$$

The massless equation corresponds to replacing $y+1$ by y, and it has no acceptable solution since the massless version of (3.4) requires $G(0)=0$. There is a continuum of solutions to (3.1), distinguished by the value of G at some large ($y\gg1$) momentum, just as in perturbative QCD. We require

$$G(y_0)=\frac{1}{(b_1\ln y_0)^{1/2}}\ ,\quad y_0\gg1\ . \tag{3.5}$$

In effect, this normalization condition has introduced a renormalization point μ via $y_0=\mu^2/m^2$. The form (3.5) is demanded by the self-consistent large-x behavior $G(x)=(b_1\ln x)^{-1/2}$, which is easily established by using this form in the integral on the right-hand side (RHS) of (3.1).

In applying (3.5) to physical processes, we have in mind that μ is fixed (say, $\mu=100$ GeV), and that at this value of μ, $G(y_0)$ has a certain value ($G^2/4\pi=0.1$, say). Let us see what happens [15] when we decrease m, keeping μ and $G(y_0)$ fixed. Clearly, y_0 increases, and our numerics show that as y_0 increases, $G(0)$ increases [perhaps not surprising given the monotone decreasing property of $G(y)$]. Eventually, G is so large that the condition (3.4) on $G(0)$ cannot be maintained, because of the G^3 growth on the RHS. There is a largest value of y_0 (smallest value of m) for which (3.4) can be satisfied; there are no solutions for m less than the critical value m_c.

Conversely, for every $m>m_c$ (3.1) has a solution, as we exhibit in Sec. IV and in Ref. [1] by numerical techniques. Of course, as m becomes $O(\mu)$ the functional form of $G(y_0)$ is not that given in (3.5), which only applies at large y_0. These solutions satisfy (3.4), because even though $G(y)$ gets smaller for small y as m increases, the RHS can be as big as the LHS because of the large-y contribution, which is nearly divergent.

It must be clear now that the existence of the critical mass m_c in this scale-free equation comes about because the renormalization procedure of fixing μ and $G(\mu)$ introduces a second mass scale, through dimensional transmutation.

The imposition of the boundary condition (3.5) invalidates the original scaling argument that since G depends only on q/m only, it either has a solution for every m or no solution at all. Simply stated, one cannot satisfy *both* equation (3.1) *and* condition (3.5) for sufficiently small m. Since the presence of the boundary condition alters things so drastically, we suggest that the reader in what follows considers Eq. (3.1) implicitly accompanied by the boundary condition (3.5). It is only after the inclusion of the boundary condition (3.5) that the concept of the critical mass can make sense. At this point we must emphasize that we consider the previous arguments as suggestive at best, and they do not in any sense constitute a proof of the existence of a critical mass. However, earlier numerical work supports this qualitative picture. Indeed, when solved numerically, Eq. (3.1) *supplemented* with the boundary condition (3.5) has no nontrivial solutions for m less than a critical value m_c. This, it turns out, is not true (and correspondingly is not seen in our numerical work) for the fermion mass M, in spite of the similar appearance of the infrared divergences in (3.1) as m or $M\to0$. Instead, for sufficiently large gluon mass m, there is a solution to (3.1) for every positive M. For $M\lesssim m$ the vertex decreases as M decreases, as one might expect from the one-loop vertex, finally vanishing at logarithmic rate at $M=0$. On the other hand, for $M\gtrsim m$, $G(0)$ can be larger than without fermions, but only by a factor of $b_1(b_1-b_2)^{-1}$ which is nearly 1 for 2-flavor QCD. (In other gauge models this enhancement could be large, and lead to chiral-symmetry breakdown without confinement.)

Now consider adding the quarks, i.e., using the full equation (3.1), but always keeping $m>m_c$. With two mass scales, it is convenient to return to conventional momentum units, and at $x=0$ (3.1) reads

$$G(0)=\frac{1}{2}\int_0^\infty dy\,yG^3(y)\left[\frac{b_1}{(y+m^2)^2}-\frac{b_2}{(y+M^2)^2}\right]\ . \tag{3.6}$$

For $M^2\ll m^2$, the b_2 term is dangerous because it has an infrared divergence, but of opposite sign to the pure gluon case. We estimate the RHS of (3.6) by setting $G^3(y)=G^3(0)$ in the small-y ($y<1$) part of the b_2 integral, and observing that the remaining integral is roughly $(b_1-b_2)/b_1$ times the corresponding pure-gluon integral (3.4). Then, taking into account the b-coefficient scaling $b^{-1/2}$ as shown in (3.5) and required for consistency at large x in (3.1), we have roughly

$$G(0)\simeq-\frac{1}{2}b_2G^3(0)\ln\left[\frac{m^2}{M^2}\right]+\left[\frac{b_1}{b_1-b_2}\right]^{1/2}G_1(0)\ . \tag{3.7}$$

We expect that $G_1(0)$ will not be too different from the pure-gluon case, where it is $O(b_1^{-1/2})$.

For 2-flavor QCD $b_2\ll b_1$ ($b_2/b_1=4/33$) so the first

(negative) term on the RHS of (3.7) is not very important when $M^2 \gtrsim m^2$. But when $M^2 \ll m^2$ this first term becomes dominant; if we estimate both $G(0)$ and $G_1(0)$ by $b_1^{-1/2}$, the two terms on the RHS are of equal importance when

$$M \sim m e^{-b_1/b_2} . \tag{3.8}$$

For smaller M the vertex function is smaller than in the pure gluon case, eventually vanishing at $M=0$, but with no critical value of M below which there are no solutions.

In the opposite case of $M \gg m e^{-b_1/b_2}$ there is no problem finding numerical solutions, as reported in Sec. IV. Furthermore, unlike the quarkless case, where $G(x)$ is a monotonically decreasing function in the whole interval 0 to $+\infty$, when quarks are included $G(x)$ can display a maximum, if M is less than a critical value. Indeed, in the quarkless case ($b_2=0$) differentiating (3.1) with respect to x (denoted with overdot) gives us

$$\dot{G}(x) = b_1 \int_0^\infty dy\, G^3(y) \dot{K}(x,y,m) , \tag{3.9}$$

with

$$\dot{K}(x,y,m) = -\frac{y\left[\sqrt{(y+x+m^2)^2-4yx}+x+m^2-y\right]}{(y+m^2)\left[y+x+m^2+\sqrt{(y+x+m^2)^2-4yx}\,\right]^2\sqrt{(y+x+m^2)^2-4yx}} . \tag{3.10}$$

Obviously, since $\dot{K}(x,y,m) \leq 0$ for all values of x and y, $\dot{G}(x) < 0$ between 0 and ∞. The inclusion of fermions changes things since now

$$\dot{G}(x) = \int_0^\infty dy\, G^3(y)[b_1 \dot{K}(y,x,m) - b_2 \dot{K}(x,y,M)] , \tag{3.11}$$

and therefore if the values of M are small enough so that the nearly infrared divergent fermionic kernel becomes large enough to overcome the small b_2 factor ($b_2/b_1 = 4/33$), $\dot{G}(x)$ may actually change sign. Beyond this qualitative observation we can say no more. However, as we will see in Sec. IV, the numerical solutions of our equation display the features described above.

Finally, we conclude this section with some comments on the validity of a standard approximation frequently used when dealing with Schwinger-Dyson equations [33]. Let us for simplicity consider once again the quarkless case ($b_2=0$). In the absence of masses (3.1) can be immediately transformed into a differential equation just by differentiating twice with respect to x (see, for example, Ref. [15]):

$$\ddot{G}(x) + \frac{2\dot{G}}{x} = -\frac{b_1}{2}\frac{G^3(x)}{x^2} , \tag{3.12}$$

which can be solved very easily with standard numerical methods. As shown in Ref. [15], this equation necessarily leads to a singularity in the IR region, so it is tempting to regularize the inverse powers of x with a gluon mass. One commonly offered justification for such a regularization is the so-called θ-function approximation to the kernel of the integral equation (3.1):

$$\frac{1}{(k+q)^2-m^2} = \frac{1}{k^2-m^2}\theta(k^2-q^2) + \frac{1}{q^2-m^2}\theta(q^2-k^2) . \tag{3.13}$$

With such an approximation for the kernel and after performing the angular integration, instead of (3.1) we have

$$G(x) = \frac{b_1}{x+m^2}\int_0^x dy \frac{yG^3}{y+m^2} + b_1 \int_x^\infty dy \frac{yG^3}{(y+m^2)^2} . \tag{3.14}$$

By differentiating with respect to x twice, we arrive at the differential equation

$$\ddot{G} + \frac{2\dot{G}}{x+m^2} = -\frac{b_1}{2}\frac{xG^3}{(x+m^2)^3} , \tag{3.15}$$

which, of course, for $m^2=0$ reproduces (3.12).

This differential equation must be implemented with two boundary conditions. These are furnished by the integral equation (3.15) itself:

$$G(0) = b_1 \int_0^\infty dy \frac{yG^3(y)}{(y+m^2)^2} \tag{3.16}$$

and

$$\dot{G}(x) = -\frac{b_1}{(x+m^2)^2}\int_0^x dy \frac{yG^3}{y+m^2} , \tag{3.17}$$

$$\dot{G}(0) = 0 . \tag{3.18}$$

The θ-function approximation leads automatically to $\dot{G}(0)=0$ for all gluon masses m. Our experience [15] in solving the original integral equation (3.1) for $m \gg \Lambda$, where Λ is the RG mass, is that $\dot{G}(0)$ is small compared to m^{-2}, so the θ-function approximation works. But for $m \lesssim \Lambda$, the physically relevant case, $\dot{G}(0)$ is not small; indeed, it diverges like m^{-2} for small m, as one expects on dimensional grounds. Here, the θ-function approximation is definitely not good. From (3.9) and (3.10), the actual value of $\dot{G}(0)$ is

$$\dot{G}(0) = -b_1 m^2 \int_0^\infty \frac{dy\, yG^3}{(y+m^2)^4} , \tag{3.19}$$

which differs substantially in form from the approximate (3.17), and is never zero.

Adding fermionic terms does not change this situation.

IV. NUMERICAL RESULTS FOR THE VERTEX EQUATION

In this section we report the results from solving (3.1) numerically, show that these results are accurately approximated by physically motivated fitting functions, and extract from the results a nonperturbative β function. This β function reduces to the usual one for very small values of the coupling constant g.

The reader may be surprised to find that a homogeneous equation such as (3.1) has a RG, which one normally associates with the existence of a divergent inhomogeneous term which is used to compensate the divergences of the integral. We will see that the integral (3.1) converges, since $G\sim(\ln y)^{-1/2}$ at large y, so no inhomogeneous term is necessary. But it is necessary to introduce a mass scale for the argument of the logarithm, which we do in the usual way by replacing y by $y\Lambda^{-2}$, where Λ is the RG mass. The connection between Λ and the coupling constant g [introduced in the original form of the vertex equation (2.17)] at any mass scale μ is the usual one:

$$g\equiv g(\mu)=G(q^2=\mu^2)\ . \tag{4.1}$$

Before we proceed, we make an additional comment on a technical point we found to be essential for the stability of our numerical algorithm. Since the integration variable in Eq. (3.1) runs between 0 and ∞, its numerical treatment requires the introduction of an ultraviolet cutoff Λ_0 (not to be confused with the RG mass Λ), much larger than any other quark or gluon mass present. In doing so, particular care must be taken so that our equation will be supplemented with the appropriate boundary condition, which will compensate for the contributions between Λ_0 and ∞ that are omitted by cutting the integral off. To be specific, after introducing the UV cutoff Λ_0 (3.1) becomes

$$G(x)=\int_0^{\Lambda_0^2}dy\ G^3(y)[b_1K(x,y,m)-b_2K(x,y,M)] +(b_1-b_2)\int_{\Lambda_0^2}^{\infty}dy\ G^3(y)K(x,y)\ , \tag{4.2}$$

where we have safely neglected all masses compared to the integration variable $y>\Lambda_0^2$ in the second term in the RHS of (4.2). Let us call this term $S(\Lambda_0)$:

$$S(\Lambda_0)=(b_1-b_2)\int_{\Lambda_0^2}^{\infty}dy\frac{G^3(y)}{y+x+|y-x|} \tag{4.3}$$

and, for $x<\Lambda_0^2$,

$$S(\Lambda_0)=\left[\frac{b_1-b_2}{2}\right]\int_{\Lambda_0^2}^{\infty}dy\frac{G^3(y)}{y}\ . \tag{4.4}$$

Using the differential equation (3.13), which $G(x)$ satisfies for large values of x, we have

$$S(\Lambda_0)=-\int_{\Lambda_0^2}^{\infty}[y\ddot{G}(y)+2\dot{G}(y)]\ , \tag{4.5}$$

where the overdots indicate differentiation with respect to y. (4.5) can be integrated by parts, giving

$$S(\Lambda_0)=-(G+y\dot{G})\Big|_{\Lambda_0^2}^{\infty}\ . \tag{4.6}$$

This can be evaluated directly in terms of the explicit large-x solution for G given in Ref. [15]. In presenting the following two equations, we simplify the notation with the understanding that all dimensionful arguments of logarithms are to be divided by Λ^2. The large-x solution for G is

$$G(x)=\frac{1}{[(b_1-b_2)\ln x+\frac{3}{2}(b_1-b_2)\ln(\ln x)+\cdots]^{1/2}} \tag{4.7}$$

where the omitted terms are negligible for large x. Using this form for G in (4.6), we find

$$S(\Lambda_0)=[(b_1-b_2)\ln\Lambda_0^2+\tfrac{3}{2}(b_1-b_2)\ln(\ln\Lambda_0^2)]^{-1/2} -\tfrac{1}{2}[(b_1-b_2)\ln\Lambda_0^2+\tfrac{3}{2}(b_1-b_2)\ln(\ln\Lambda_0^2)]^{-1/2} \times\left[1+\frac{1}{\ln\Lambda_0^2}\right]\ . \tag{4.8}$$

This boundary condition is very stable against variations of Λ_0. In particular our results change only by less than 1% when Λ_0 undergoes changes of six orders of magnitude (10^3–10^9).

Note, by the way, that the solution (4.7) is precisely of the form given by the RG, although all but the leading term have somewhat different coefficients from those found by using the real β function of QCD. Further analysis of the all-orders β function implied by (4.7) is given in Ref. [15], where it is shown that the approximate β function, like the true QCD β function, has a factorially divergent power-series expansion in powers of g^2. This divergence is an artifact of the massless large-x expansion, and is cured [15] by the inclusion of masses [see (4.20) below].

Figure 6 shows numerical results for the running coupling $\alpha_s(q^2)$, defined by

$$\alpha_s(q^2)=G^2(q^2)/4\pi\ . \tag{4.9}$$

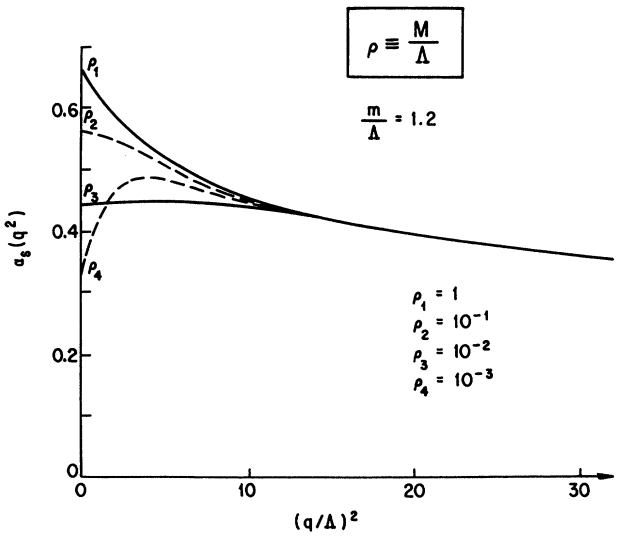

FIG. 6. Running coupling $\alpha_s(q^2)$ vs q^2 for various fermion masses and gluon mass fixed at $m=1.2\Lambda$.

In interpreting these results, it is very convenient to keep in mind a function which is both a physically motivated [13] approximation to $\alpha_s(q^2)$ for all q^2 and also a very reasonable fit to our numerical data, again for all q^2 including $q^2=0$. This function recognizes the role of mass generation in taming the logarithmic IR divergences which would be found by extrapolating (4.7) to the IR. It reads

$$G(q^2)\simeq\left[b_1\ln\left[\frac{q^2+\xi m^2}{\Lambda^2}\right]-b_2\ln\left[\frac{q^2+\xi M^2}{\Lambda^2}\right]\right]^{-1/2}. \tag{4.10}$$

Originally [13] it was suggested that $\xi=4$, representing the usual two-particle thresholds. We find that choosing $\xi=4.8$ fits the data of Fig. 6 to within 5% or better, over a range of fermion masses from $10^{-3}\Lambda$ to Λ. This fit to the vertex suggests that $m_c=\xi^{-1/2}\Lambda=0.46\Lambda$, so that for $m<m_c$ the vertex is singular (at $q=0$ G would be imaginary, unless M is very small). Actually, our numerical computations break down (for any quark mass M) at a larger critical mass, with $m_c\simeq 1.1\Lambda$, but are quite well behaved even for a slightly larger m, as Fig. 6 shows ($m=1.2\Lambda$). So we will adopt the value

$$m_c=1.1\Lambda \tag{4.11}$$

in what follows.

Let us now concentrate on the behavior of the strong coupling in the infrared, which is one of the central issues of our investigation. We will keep the discussion rather general, with b_1, b_2, and ξ unspecified—anticipating that the generic form of (4.10) may persist to more general cases—although of course we have only checked the SU(3) case with two families of fermions in the fundamental representation.

From (4.10) we have

$$\alpha_s(0)=\frac{1}{4\pi}\left[\frac{1}{b_1\ln(\xi m^2/\Lambda^2)-b_2\ln(\xi M^2/\Lambda^2)}\right]. \tag{4.12}$$

It is now clear from (4.12) that if M takes values smaller than $M_c=\Lambda\sqrt{\xi}$, its effect is to decrease the value of $\alpha_s(0)$. This is of course something we anticipated based on the qualitative discussion of the previous section. For the case at hand, $\xi=4.8$ and so $M_c\simeq 0.46\Lambda$ (see also Fig. 6). $\alpha_s(0)$ reaches its maximum value $\alpha_s^{\max}(0)$ when $m=M\simeq m_{\rm cr}\simeq 1.1\Lambda$. Substituting in (4.10), with $b_1=33/48\pi^2$ and $b_2=4/48\pi^2$, $\alpha_s^{\max}(0)\simeq 0.73$, a rather moderate value. We can already anticipate trouble with the fermion gap equation, which will (see Sec. V) require rather larger values of $\alpha_s(0)$. For small enough values of M, $\alpha_s(q^2)$ has a maximum at $q^2_{\max}\neq 0$:

$$q^2_{\max}=\xi\left[\frac{b_2m^2-b_1M^2}{b_1-b_2}\right] \tag{4.13}$$

as long as $M<m(b_2/b_1)^{1/2}$. The value of $\alpha_s(q^2_{\max})$ is

$$\alpha_s(q^2_{\max})=\left[4\pi\left[c+(b_1-b_2)\ln\frac{\xi(m^2-M^2)}{\Lambda^2}\right]\right]^{-1}, \tag{4.14}$$

with

$$c=b_1\ln\left[\frac{b_1}{b_1-b_2}\right]-b_2\ln\left[\frac{b_2}{b_1-b_2}\right]>0\ . \tag{4.15}$$

It follows that

$$\alpha_s(q^2_{\max})<\frac{1}{4\pi(b_1-b_2)\ln(\xi m^2/\Lambda^2)}, \tag{4.16}$$

which is just the quarkless vertex at $q^2=0$, multiplied by $b_1(b_1-b_2)^{-1}$. This latter number is greater than unity (but not by much for QCD), so fermions can slightly enhance the running coupling, but this is not a particularly important effect. What will be much more important for the fermion gap equation, studied in Sec. V, is that the maximum value of α_s is limited by the gluon mass.

We end our discussion with the derivation of the β function corresponding to the running coupling of (4.10). Given $G(q^2)$, this β function is of course superfluous but is of heuristic interest for comparison with previous constructs. Let us omit the quarks for the moment ($b_2=0$). Then

$$\beta=2q^2\frac{dG}{dq^2}=-b_1G^3\left[\frac{q^2}{q^2+\xi m^2}\right]. \tag{4.17}$$

Inverting (4.10) we have

$$q^2+\xi m^2=\Lambda^2\exp\left[-\frac{1}{b_1G^2}\right] \tag{4.18}$$

and so (4.17) becomes

$$\beta=-b_1G^3\left[1-\xi\left[\frac{m}{\Lambda}\right]^2\exp\left[-\frac{1}{b_1G^2}\right]\right]. \tag{4.19}$$

The explicit appearance of a factor $\exp(-1/b_1G^2)$ multiplying the contribution of the gluon mass term should be considered indicative of the nonperturbative nature of the entire approach. If we include quarks, the equivalent of (4.17) becomes

$$\beta=-G^3\left[b_1\frac{q^2}{q^2+\xi m^2}-b_2\frac{q^2}{q^2+\xi M^2}\right]. \tag{4.20}$$

Now we cannot invert (4.10) exactly but only approximately. Assuming that m and M are approximately of the same order, we find

$$\beta=-(b_1-b_2)G^3\left\{1-\xi\left[\frac{b_1(m/\Lambda)^2-b_2(M/\Lambda)^2}{b_1-b_2}\right]\times\exp\left[-\frac{1}{(b_1-b_2)G^2}\right]+O\left[\exp\left[-\frac{1}{(b_1-b_2)G^2}\right]\right]\right\}, \tag{4.21}$$

which, of course, for $b_2=0$ reduces to (4.19).

V. RUNNING COUPLING AND THE FERMION GAP EQUATION

In two previous sections we presented a detailed analysis of the running coupling that emerges from the study of a nonlinear vertex equation, valid in both the IR and the UV and involving all fundamental masses of the theory. Throughout this analysis these masses were more or less treated as free-adjustable parameters and nothing was said about the dynamical mechanisms generating them. In particular, the conventional wisdom for chiral-symmetry breaking through the study of gap equations for the quark self-energy is that mass generation becomes possible only if the strong coupling $\alpha_s(0)$ becomes larger than a critical value. For values of $\alpha_s(0)$ below this critical value, chiral symmetry is restored and the gap equation has only trivial solutions. This critical value for the coupling varies depending on the different approximations employed when writing the gap equation, but the common assumption is that, no matter what the critical value may be, the infrared dynamics of QCD will somehow manage to provide it. However, in the absence of any confinement effects, the infrared dynamics of the coupling are determined by an equation such as (3.1) and depend nontrivially on the values of the masses. In particular we argued that strict masslessness for either the gluons or the quarks would result in the breakdown of the vertex equation. Moreover, the actual value of $\alpha_s(0)$ is completely determined by the values of the masses present and some group-theoretical factors. So the question arises naturally whether these two dynamic pictures are compatible with each other.

Insofar as a study of the fermion gap equation alone is concerned, Haeri and Haeri [10] have recently and independently carried out an investigation quite similar to ours, using a massive gluon propagator and the original running coupling of Ref. [13]:

$$g^2(q)=\left[b\ln\left[\frac{q^2+4m^2}{\Lambda^2}\right]\right]^{-1} \tag{5.1}$$

and letting the gluon mass vary. This differs from our running coupling in (4.10) by having no fermion terms, and by having ξ of that equation equal to 4 instead of 4.8. Since these authors did not study the vertex equation, they could use the gluon mass as a free parameter, and what they found, with which we agree, is that it takes very small values of m, not much bigger than $\Lambda/2$, in order to generate a nonzero fermion mass M.

We will now analyze two different versions of the fermion gap equation: the one of Haeri and Haeri [10], and that of Atkinson and Johnson [6]. In both references the important issue of gauge covariance has been successfully addressed and gauge-technique-improved quark-quark-gluon vertices have been employed. The relevant equation from Ref. [6] is

$$M(x)=\frac{3C_f}{4\pi}\left[\frac{\alpha_1(x)}{x}\int_0^x dy\frac{yM(y)}{y+M^2(y)}+\int_x^\infty dy\frac{\alpha_1(y)M(y)}{y+M^2(y)}\right] \tag{5.2}$$

and, from Ref. [10],

$$M(x)=\frac{3C_f}{4\pi}\left[\alpha_2(x)\int_0^x dy\frac{yM(y)}{[y+M^2(y)]\{y+x+m^2+[(y+x+m^2)^2-4xy]^{1/2}\}}+\int_x^\infty dy\frac{y\alpha_2(y)M(y)}{[y+M^2(y)]\{y+x+m^2-[(y+x+m^2)^2-4xy]^{1/2}\}}\right]. \tag{5.3}$$

One major difference between the two equations is that in (5.3) the gluon propagators are massive. Moreover, in deriving (5.2), the usual θ-function approximation has been used to simplify the kernel whereas in (5.3) it has only been used for the running coupling, e.g.,

$$\alpha_2(p+k)\simeq\theta(p^2-k^2)\alpha_2(p)+\theta(k^2-p^2)\alpha_2(k)\ . \tag{5.4}$$

So by using these two different versions of the gap equation we reduce the risk that our conclusions will depend heavily on the validity of the θ approximation.

Furthermore, $\alpha_1(x)$ in (5.2) is given by

$$\alpha_1(x)=\frac{12\pi}{(32-2n)\ln[(\tau+x)/\Lambda^2]}\ , \tag{5.5}$$

where τ is an adjustable parameter, not explicitly associated with anything physical, whereas in (5.3) we have, for α_2,

$$\alpha_2(x)=\frac{12\pi}{(33-2n)\ln[(x+4m^2)/\Lambda^2]}\ , \tag{5.6}$$

m being the gluon mass. We have used our improved

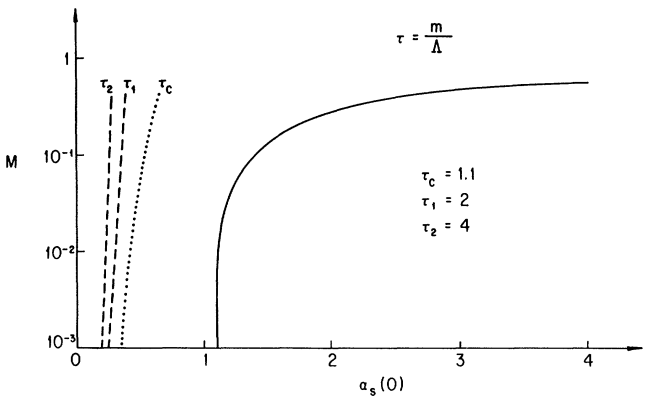

FIG. 7. Plot of the dynamical fermion masses M vs $\alpha_s(0)$ for various values of the gluon mass m. The solid curve comes from the fermion gap equation, and the other curves come from the vertex equation. The solid curve is calculated for $m=0$; any other value of m yields smaller values of M for a given $\alpha_s(0)$.

running coupling of (4.10),

$$\alpha_s(x)=\frac{12\pi}{33\ln[(x+\xi m^2)/\Lambda^2]-4\ln[(x+\xi M^2)/\Lambda^2]}\ , \tag{5.7}$$

with $\xi\simeq 4.8$, instead of the functions $\alpha_{1,2}$ in the gap equations, and solved them numerically. The results are shown in Fig. 7, where they are compared to analogous results from the vertex equation. Both this and the gap equation yield a relation between the fermion mass M and $\alpha_s(0)$, and the major result of our work is that these relations are incompatible for gluon masses $m>m_c=1.1\Lambda$. The gap equation is calculated with a massless gluon propagator, which gives the most optimistic case; if a massive gluon propagator is used, $\alpha_s(0)$ must be larger to achieve a given M. But (as shown in the figure) if m is increased in the vertex equation, $\alpha_s(0)$ decreases.

VI. CONCLUSIONS

As we have said before, the gap equation and the vertex equation are inconsistent because the vertex equation requires a reasonably large gluon mass in order to have a nonsingular solution, and this prevents the running coupling from getting large enough to drive chiral-symmetry breaking (CSB). While the vertex equation is a fairly crude model of the exact coupled nonlinear Schwinger-Dyson equations of QCD, the constraints which follow from it are quantitatively consistent with what we know about QCD via other considerations. For example, it has been argued [13] and confirmed [18] on the lattice that the gluon mass is about 500–600 MeV, comfortably above the critical value 1.1Λ, and a 500-MeV gluon mass gives $\alpha_s(0)\simeq 0.5$, as required by, e.g., quarkonium potentials [19]. Furthermore, the vertex equation yields the right UV asymptotics. We are confident that our vertex equation is as accurate in its domain as the various fermion gap equations [e.g., (5.2), (5.3)] are in theirs, and the reason for their mutual inconsistency is not because of inaccuracies of the models, but because confinement has not been taken into account.

This is not to say that it is always necessary that CSB be driven by confinement in non-Abelian gauge theories. What makes it necessary in QCD is that quark contribution to the β function, as measured by b_2, is rather small compared to the gluon contribution b_1. Clearly, as the number of flavors increases, or (as might happen in some other domain, such as grand-unified theories) for some other reason b_2 gets closer to the size of b_1, the running charge gets larger because the fermions tend to cancel the gluons in the denominator of the running charge [see, e.g., (4.10)]. If QCD had seven or eight flavors it would not be surprising that there would be CSB without confinement.

A theory with b_1- b_2 small may have a pitfall, however. In general, non-Abelian gauge theories will have confinement, as that both confinement mechanisms such as discussed in the introduction and single-gluon-exchange mechanisms such as discussed here will be operative. The latter leads to a gap equation schematically of the form $M=MK(M)$, where $K(M)$ depends on M through nonlinearities in the gap equation. Confinement, as discussed in the introduction, essentially yields $M=\sigma^{1/2}$, where σ is the string tension. The naive combination of mechanisms as in

$$M=\sigma^{1/2}+MK(M) \tag{6.1}$$

may show unusual sensitivities and behavior, depending on the derivative of K near the point M_0, where $K(M_0)=1$. For example, a very small slope makes the solution to (6.1) hypersensitive to the confinement term. The study [11] of non-Abelian gauge theories with small b_1-b_2 is interesting for many reasons, and CSB in such theories requires a separate investigation along the lines suggested in this paper, with the addition of confinement effects, if appropriate.

There is one important circumstance in QCD where confinement is missing and the results of this paper are (broadly speaking, but not in detail) applicable. That is at temperatures above the deconfinement transition, where by definition confining forces are absent. Our results suggest that this must also be the CSB phase-transition temperature, since we have seen that without confinement there is no CSB. To verify this in detail, of course, requires a study of the vertex and the gap equations at finite temperature. Such a study will be pursued in the near future.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant No. PHY/89-15286 and by the U.S. Department of Energy under Contract No. DE-AC02-76CH00016. One of us (J.P.) thanks P. Gondolo for his help with the numerical computations.

*Present address: Brookhaven National Laboratory, Upton, New York 11973.

[1] K. Johnson, M. Baker, and R. Willey, Phys. Rev. **136**, B111 (1964); **163**, B1899 (1967).
[2] For example, S. Adler and W. Bardeen, Phys. Rev. D **4**, 3095 (1971); T. Maskawa and H. Nakajima, Prog. Theor. Phys. **52**, 1326 (1974); **54**, 860 (1975); V. A. Miransky, Phys. Lett. **91B**, 421 (1980); P. J. Fomin, V. P. Gusynin, V. A. Miransky, and Yu. Sitenko, Riv. Nuovo Cimento **6**, 1 (1983).
[3] K. Lane, Phys. Rev. D **10**, 2605 (1974).
[4] K. Higashijima, Phys. Rev. D **29**, 1228 (1984).
[5] H. Pagels, Phys. Rev. D **19**, 3080 (1979).
[6] D. Atkinson and P. W. Johnson, Phys. Rev. D **37**, 2996 (1988).
[7] D. Atkinson and P. W. Johnson, Phys. Rev. D **37**, 2290 (1988).
[8] B. Haeri, Phys. Rev. D **38**, 3799 (1988).
[9] C. D. Roberts and B. H. J. McKellar, Phys. Rev. D **41**, 672 (1990).
[10] B. Haeri and M. B. Haeri, Phys. Rev. D **43**, 3732 (1991).
[11] For example, T. Appelquist, D. Carrier, L. C. R. Wijewardhana, and W. Zheng, Phys. Rev. Lett. **60**, 1114 (1988); T. Appelquist, K. Lane, and U. Mahanta, *ibid.* **61**, 1553 (1988).
[12] For example, W. A. Bardeen, C. N. Leung, and S. T. Love, Nucl. Phys. **B323**, 493 (1989); K.-I. Kondo, H. Mino, T. Takeuchi, and L. C. R. Wijewardhana, in *TeV Physics,* Proceedings of the 12th Johns Hopkins Workshop on Current Problems in Particle Theory, Baltimore, Maryland, 1988, edited by G. Domokos and S. Kovesi-Domokos (World Scientific, Singapore, 1988).
[13] J. M. Cornwall, Phys. Rev. D **26**, 1453 (1982).
[14] V. A. Miransky, V. P. Gusynin, and Yu. Sitenko, Phys. Lett. **100B**, 157 (1981); V. A. Miransky and P. J. Fomin, *ibid.* **105B**, 387 (1981).
[15] J. M. Cornwall and J. Papavassiliou, Phys. Rev. D **40**, 3474 (1989).
[16] E. Bogomol'nyi, V. A. Faddeev, and L. N. Lipatov, in *Physics Reviews,* edited by I. M. Khalatnikov, Soviet Scientific Reviews, Vol. 2 (Harwood, Chur, Switzerland, 1980), Sec. A, p. 247.
[17] J. E. King, Phys. Rev. D **27**, 1821 (1983).
[18] C. Bernard, Phys. Lett. **108B**, 431 (1982); Nucl. Phys. **B219**, 341 (1983); J. E. Mandula and M. Ogilvie, Phys. Lett. B **185**, 127 (1987); R. Gupta, G. Guralnik, G. Gilcup, A. Patel, S. R. Sharpe, and T. Warnock, Phys. Rev. D **36**, 2813 (1987); K. Yee, C. Bernard, and A. Soni, in *Lattice '90,* Proceedings, Tallahassee, Florida, 1990, edited by U. M. Heller, A. D. Kennedy, and S. Sanielevici [Nucl. Phys. B (Proc. Suppl.) (in press)].
[19] E. Eichten, K. Gottfried, T. Kinoshita, J. Kogut, K. D. Lane, and T.-M. Yan, Phys. Rev. Lett. **34**, 369 (1975); E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T.-M. Yan, Phys. Rev. D **17**, 3090 (1978); **21**, 203 (1980).
[20] H.-Q. Ding, Phys. Rev. D **42**, 2350 (1990).
[21] A. Casher, Phys. Lett. **83B**, 395 (1979).
[22] J. F. Donoghue and K. Johnson, Phys. Rev. D **21**, 1975 (1980).
[23] J. M. Cornwall, Phys. Rev. D **22**, 1452 (1980).
[24] J. Kogut, M. Stone, H. W. Wyld, J. Shigemitsu, S. H. Shenker, and D. K. Sinclair, Phys. Rev. Lett. **48**, 1140 (1982).
[25] C. Bernard, Ref. [18]; J. M. Cornwall, in *Workshop on Non-Perturbative QCD,* edited by K. A. Milton and M. A. Samuel (Birkhaüser, Boston, 1983), p. 119; J. M. Cornwall and A. Soni, Phys. Lett. **120B**, 431 (1983).
[26] Bernard (Ref. [18]).
[27] Cornwall (Ref. [25]).
[28] Reviewed in J. M. Cornwall, Comments Nucl. Part. Phys. **15**, 223 (1986).
[29] J. M. Cornwall, Nucl. Phys. **B157**, 392 (1979).
[30] J. M. Cornwall, in *Gauge Theories, Massive Neutrinos, and Proton Decay,* edited by A. Perlmutter (Plenum, New York, 1981), p. 141.
[31] This cancellation only in color-singlet states is no accident. The infrared singularity in M or $V(r)$ can be removed [23] by a singular gauge transformation, and it is only for gauge-invariant color-singlet states that the two singularities can cancel.
[32] J. Papavassiliou, Phys. Rev. D **41**, 3179 (1990).
[33] Comments in a similar vein have been made by H. Munczek and D. McKay, Phys. Rev. D **42**, 3548 (1990).