2

SLAC-PUB-2372
December .979
(T/E)

ELEMENTS OF QUANTUM CHROMODYNAMICS*

J. D. Bjorken

Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

Presented

at

the

Summer Institute on Particle Physics,

at

Stanford Linear Accelerator Center, California

July 9-20, 1979

*Work supported by the Department of Energy under contract number DE-AC03-76SFG0515.

TABLE OF CONTENTS

Chapter

Appendix

ACKNOWLEDGEMENT

Before beginning, I wish to acknowledge the invaluable and generous help of Marvin Weinstein in introducing me to much of the material covered in these lectures. While there are many shortcomings of these notes, they are not to be blamed on him; they would have been much poorer without his help and insights.

1. INTRODUCTION

Quantum chromodynamics should need little introduction, since it already permeates almost all descriptions of strong interaction phenomen nowadays. We recount here very briefly the basic motivation, beginning with some factual evidence:

1. Quarks of fractional charge and three colors seem to be require as constituents of hadrons in order to understand the spectrum of hadron and their resonances.

2. The ratio of electron-quark to neutrino-quark deep inelastic scattering argues strongly for fractional charge of the quarks.

3. In addition to the spectroscopic evidence, the observed width of the decay $\pi^0 \to 2\gamma$ and the large cross section for $e^+e^- \to$ hadrons is successfully understood provided there are 3 colors of quarks.

4. The color-symmetry should be exact (or very nearly so); otherwise we would expect additional low-lying color non-singlet hadron state states for which there is no empirical evidence.

Based on this evidence, an analogy between color and charge is an attractive one. Just as the conserved charge is closely related to the electromagnetic force, one may search for a strong force related in a similar way to the conserved color quantum-numbers. The most immediate answer to this, and the one most similar to quantum electrodynamics (QED is quantum chromodynamics (QCD). The beginning of the analogy between QED and QCD is exhibited in Table 1.

In that table the bold-faced quantities are 3 × 3 matrices acting on column vectors of quark fields

TABLE 1

	QED	QCD
Conserved quantum numbers	charge	3 colors
Symmetry group	U(1)	SU(3)
Transformation property of source (electron or quark)	$e \to e^{iQ\theta} e$	$q_i \to \left(e^{(i/2)\underset{\sim}{\lambda}\cdot\underset{\sim}{\omega}}\right)_{ij} q_j$; $i,j = 1,2,3$
Current density	$\bar{e}\,\gamma_\mu\, e$	$\frac{1}{2}\,\bar{q}\,\gamma_\mu\,\underset{\sim}{\lambda}^A\, q$; $A = 1,2,\ldots,8$
Quanta of force coupled to currents	massless photon	eight massless gluons
Field-variables (potentials)	$A_\mu(x)$	$\underset{\sim}{A}_\mu(x) = \sum_{A=1}^{8} \frac{1}{2}\,\underset{\sim}{\lambda}_A\, A^A_\mu(x)$
Gauge invariant substitution (when acting on quarks)	$\frac{\partial}{\partial x_\mu} \to \frac{\partial}{\partial x_\mu} - ieA^\mu(x)$	$\frac{\partial}{\partial x_\mu} \to \frac{\partial}{\partial x_\mu} - ie\underset{\sim}{A}^\mu(x)$

$$q(x) = \begin{pmatrix} q_1(x) \\ q_2(x) \\ q_3(x) \end{pmatrix} \tag{1.1}$$

The matrices λ_A are the 8 independent 3×3 hermitian traceless matrices of Gell-Mann, and are exhibited in Section 3.

The comparison in Table 1 could go on and on. That in fact will be the case throughout these lectures. For now it suffices to say that QCD in the weak coupling limit can be formulated in a way very similar to QED. We may ask whether, once having that formulation in hand, it produces any useful results. The answer is yes; it appears to provide a consistent description of those hadron processes that depend on short distances only, as well as provide some justification for the applicability of parton-model ideas to hadron processes. The list of these includes the value of $R(e^+e^- \to \text{hadrons})$ and the approximate scaling behavior of deep-inelastic lepton-nucleon structure functions. In addition the deviations from scaling exhibit the qualitative behavior expected from the theory. Some claim the significance is precise and quantitative, but we shall not raise that question here.

An important property of QCD that distinguishes it from QED is "asymptotic freedom." In QED, the effective charge at short distances becomes larger as a consequence of vacuum polarization. If for QED we write in momentum space the force between static charges as

$$V(q^2) = \frac{4\pi\alpha(q^2)}{q^2} \qquad (1.2)$$

then

$$\frac{1}{\alpha(q^2)} \cong \frac{1}{\alpha(m_e^2)} - \frac{1}{3\pi}\log\frac{q^2}{m_e^2} \qquad (1.3)$$

Thus at sufficiently short distances (but in practice ridiculously short), QED becomes a strong-coupling theory. As we shall discuss in more detail later, in QCD the vacuum-polarization has the opposite sign; the corresponding equation is

$$\frac{1}{\alpha(q^2)} = \frac{1}{\alpha(M^2)} + \frac{(33-2N_f)}{12\pi}\log\frac{q^2}{M^2} \qquad (1.4)$$

where N_f (~3 or 4) is the relevant number of quark flavors. Thus at sufficiently large distances (large compared to 10^{-14} cm) QCD becomes a strong-coupling theory. This is both good and bad news: good news because something non-perturbative is needed at large distances to provide a mechanism for confinement of color: neither the gluons nor quarks which appear in the field equations appear in the spectrum of hadrons. What exists is only a hypothesis that only color singlet combinations of quarks and gluons can exist as isolated particles. The bad news is just that we don't know how to calculate with — and even to formulate — the theory on a large distance scale.

It is therefore a serious question whether QCD is a theory at all. Let us compare the situation with QED. There one can follow textbook approaches to the subject — specifically two particular textbooks,[1] hereafter to be known as Book I and Book II. The Book I approach uses the classical equations of motion and common sense to motivate rules for Feynman-diagrams. The emphasis is on learn-by-doing and intuition, with a relatively casual attitude toward a strict systematic logical development. I think most of contemporary perturbative QCD does not get far beyond this level. Beyond the relatively rock-solid predictions to which we already alluded, there do not exist clear rules which divide phenomena dependent on the non-perturbative, confining part of the theory from phenomena dependent on the purely perturbative aspect.

In QED, the Book II approach is the strict logical one of canonical quantization of a classical field theory. In the presence of interactions, but in the weak-coupling regime, the LSZ formalism can be used to relate asymptotic states of physical, isolated electrons and photons to the field variables. However, the formalism relies rather heavily on these asymptotic scattering states — in other words, on large distances. Therefore this approach appears to be closed in QCD. Actually as we shall see, the situation is not quite that bad. However, I think perturbative QCD is much more restricted than QED, for reasons that hopefully become clearer as we proceed.

These lectures will concentrate for the most part on the Book II approach. This is not meant to imply disparagement of the Book I method; Stan Brodsky will most ably cover that.[2] Nor is it meant to downgrade other approaches, in particular those based on path-integral formulations.[3] Indeed, use of path-integral techniques has thus far been the most successful mode of attack on the difficult mathematical problems posed by QCD. The rules for diagrams are most efficiently derived in that way (especially in correctly accounting for the subtleties associated with Faddeev-Popov ghosts). Also the demonstration of confinement in strong-coupling lattice QCD and the studies of instantons are best done within the path-integral formalism. Our rationale for avoiding it here is simply to look at the subject in a slightly different way [after all, some problems, such as the hydrogen atom, are more difficult using path-integrals], as well as to spare the less theoretically oriented reader the unavoidable preliminary technology needed to set up the path-integral formalism. In any case, the problem confronting everyone is difficult; all attacks should be brought to bear.

We shall classify the subject-matter into three stages of increasing complexity. In the first we consider "pure" QCD in the absence of any fermions or other sources. The QED analogue is the (trivial) theory of free non-interacting photons. For QCD this is presumably the (nontrivial) theory of interacting gluonium (color-singlet bound glue) states. The second stage allows the introduction only of superheavy quarks in addition to the pure glue, with emphasis on the non-relativistic limit of the quark motion. In QED this is not much more than the theory of the Coulomb interaction — in other words, nothing but all of chemistry. In QCD, this much should already allow study of basic questions of confinement; e.g., the nature of the static potential between heavy quarks at large distances. Only in the third stage will we introduce the light quarks u, d, s. The presence of copious vacuum polarization and pair creation should modify the structure of the theory in a major way. Nevertheless it can be hoped that these modifications have less to do with the existence of confinement, and more to do with the nature of the spectrum of confined hadrons.

These lectures will be organized as follows. In Section 2, after making some assumptions about the nature of QCD, we shall describe what we think the solution of the theory looks like, first for Stage I, and then for Stages II and III. This will not even be at Book I level — call it Book Zero. For the most part we will not question whether it is the solution; that problem is too hard. But if it weren't, it isn't clear we would ever be very interested in QCD as a theory to be applied to the real world.

In Section 3, we discuss canonical quantization of QCD with use of a Hamiltonian formulation (in $A_0 = 0$ gauge). We shall encounter formal problems such as gauge-ambiguities and topological classifications of the states (θ-vacua, instantons, and all that) not met in QED. This will be attempted at a descriptive, relatively painless level, details being left as an appendix to these lectures. Sections 5, 6, and 7 describe in more detail the properties of Stages I, II, and III of QCD as we sketched above. In Section 8 we briefly sketch and compare various approaches to the confinement problem. In Section 9 we very briefly discuss more radical alternatives to QCD, such as the string model or the Pati-Salam program. Section 10 is devoted to conclusions and to an assessment of the experimental situation. Two appendices are devoted to a more technical exposition. We urge that it be studied in parallel with the main text.

2. THE SOLUTION TO QCD

In this section we shall be optimists and look through rose-colored glasses at what the solution to the theory should qualitatively look like if all goes well. In doing this we make the following set of assumptions.

1. At distances small compared to 10^{-14} cm, the theory is just that of unconfined gluons and quarks interacting with each other via a coupling α_{strong} of small strength: $\alpha_{strong} \lesssim 0.2$.

2. Only color-singlet particle states exist in isolation and none of them are massless at any stage.[4]

3. All confinement effects are "soft," i.e., characterized by a momentum scale $\lesssim 1$ GeV.

4. "Naive" parton-model ideas may be used to connect what goes on at short distances to hadron structure.

Now let us advance to Stage I and ask what the theory is like. At short distances the only quanta are color-octet spin-1 gluons. At large distances we must form composites of these quanta to form color singlets. The details of how to do this depend upon questions of confinement. Nevertheless we can make some educated guesses, based upon what happens to up and down quarks. At short distances these quarks have color and negligible mass (a few MeV, according to current-algebra ideas[5]). But in a hadron the quarks can be considered to have a nonvanishing mass (~300 MeV) and to move non-relativistically. If this is also so for gluons,[6] then the S-wave states can be classified rather easily; their wave functions must be gauge-invariant and color-singlet. The simplest are two-gluon bound states. They are all SU(3) singlet and C-even states. C-odd states can be formed from 3 gluons (decay products of ψ and Υ are a familiar example) as well as C-even.

Two color-electric fields might bind to $J = 0^+$ and 2^+ states, while two color-magnetic fields might bind to another distinct 0^+ and 2^+ hyperfine multiplet. An electric gluon also might bind to a magnetic gluon; in this case the parity is reversed and the triplet is 0^-, 1^+, 2^-. (The 0^- state may be regarded as suspect inasmuch as the operator which creates it is $\mathrm{Tr}\,\vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}}$, which is a total divergence and which plays an important role in instanton phenomena and the U(1) problem;[7] see Section 6 and Appendix G.)

Just as for ordinary hadrons, these states should possess a composite structure and have a rich spectrum of excited states. They should scatter from each other like hadrons generally do. The big question is evidently their typical mass. The only mass in the theory is the mass scale at which perturbation theory breaks down; by hypothesis ~1 GeV. Estimates from the MIT bag model have given masses of 1.0-1.5 GeV for the lowest lying gluonium states. But, to say the least, this is an uncertain business.[8] There is a school of thought (which I don't understand) that puts gluonium masses considerably higher. We shall return to the pheno .nology of such objects in Section 4, but now choose to quickly move on to Stage II.

In Stage II, we introduce heavy color-triplet quarks such as charm or bottom quarks. Actually for this idealization superheavy quarks Q with mass $\gtrsim$100 GeV would serve the purpose best, so that the size of their bound-state wave functions would clearly be small compared to the confinement radius. The properties of such superheavy quarkonium would be expecte. to be well-calculated perturbatively. The "binding energy" calculated from perturbation theory would be hydrogenic, $-\alpha_{st}^2 M$ (~1 GeV), and the size $(\alpha_{st} M^{-1}) \sim 10^{-15}$ cm.

Now let us imagine scattering an electron from this heavy quarkonium system at very large momentum transfer — say, $Q^2 \sim 10^4$ GeV^2. The struck quark will move semi-relativistically ($\gamma \sim 2$) away from its partner. Because of its large inertia, nothing to do with the confinement mechanism can immediately stop it. By hypothesis (specifically, assumption 3 made in the beginning of this section) the rate of momentum transfer to gluon degrees of freedom, potential energy, etc. cannot exceed a few GeV per fermi of travel. Thus after some time (say 10^{-22} sec., which is a long time), the heavy quark and antiquark will be separated by a large distance (say, $\gtrsim 10^{-11}$ cm). However by hypothesis (assumption 2) the quarks are not themselves free because they are color triplets, not color singlets. Nor will any finite number of (octet) gluons locally dressing or screening the struck quark do any good.[9] Pair-creation of superheavy quarks might be invoked, but this is a very unlikely process because of the large rest mass of the quarks Q. Indeed if somehow a virtual $Q\bar{Q}$ pair could be created the final state would consist of two heavy quarkonia. But the total mass would have to be $\gtrsim 4M_Q$, and the total energy of the system need not be that large.

We are thus unable to avoid the conclusion that during the separation of Q and $\bar{Q}$ some spoor of gluon degrees of freedom were left behind in the space between them. The $Q\bar{Q}$ system is only able to be considered as one complex extended (color-singlet) system and not two isolated ones. What region of space is likely to be affected? We may consider three possibilities: (a) very big, (b) string like, and (c) minimal, as shown in Figure 1. The very big system looks inefficient (too much stored energy) and also appears to depend upon degrees of freedom of long wavelength. By hypothesis (assumption 2) there are no massless color-singlet modes, and configurations of such large spatial extension (even though they are not in isolation) make more serious the problem of absence of massless modes. While this is rather feeble hand waving, we are led to consider more favorably case (b). Here a string of fixed thickness connects the Q and $\bar{Q}$. Evidently the diameter of the string should be related to the confining scale and be between 10^{-13} and 10^{-14} cm. The energy per unit length should be constant, and can be estimated from "old" hadron spectroscopy or from the charmonium potential. In either case an energy per unit length of 1 GeV per fermi is a reasonable guess.

The minimal case (c) also supposes a string configuration, but allows increasing constriction as the string gets longer and longer. But such a constriction seems to violate our starting hypothesis. Given a constriction of size small compared to the confining radius, momenta large compared to the confinement scale will be involved in the confinement mechanism. This violates our assumption 3.

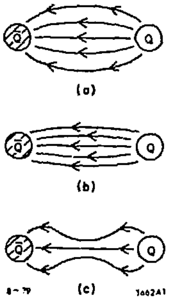


Fig. 1. Three views of the color field between widely separated heavy quarks.

Therefore we find solution (b) the preferred one and infer that (in Stage II) a string composed of gluon degrees of freedom with characteristic thickness of order the confinement scale and an energy per unit length ~1 GeV/f connects widely separated superheavy quarks. Furthermore this string is universal; it depends only on the color triality of the quark source and not on specific color representations (e.g., 6, 15, 24, etc.) of the sources. To see this, first suppose that the superheavy quark were a color octet. Then upon separation of Q and $\overline{Q}$ there will be no string at all: the color of the quark can be locally screened by an octet gluon. In other words there should be color-singlet states of Q (or $\overline{Q}$) bound to a gluon. Now consider a quark in a different representation of nonvanishing triality such as 6. A string connecting a 6 and $\overline{6}$ would a priori be expected to have a different energy per unit length than for a string connected 3 and $\overline{3}$ — probably larger, because there is more color-flux. If the energy per unit length is larger, then it is energetically favorable for the sources to be partially screened by gluons to form an effective source which is 3 or $\overline{3}$. For example $\overline{3}$ is contained in $6 \otimes 8$. In a similar way, any color representation can be reduced to 1, 3, or $\overline{3}$ by multiplication by a suitable number of 8's. Thus either there is no string or else it is coupled to an effective source made of 3 and $\overline{3}$.

Now let us return to the original problem of dynamics of the $Q\overline{Q}$ system. Evidently, the first approximation is that they move semi-classically in a confining linear potential. But how does the system decay? The two mechanisms available are radiation of gluonium and, for $E > 4M_Q$, dissociation into two quarkonium states. Each of these processes can be further subdivided into soft and hard processes. We discuss the hard processes first. The Q and $\overline{Q}$ are in a bound state with very little damping and therefore pass by each other twice each period of the motion. At each passage they may, with finite but small probability, undergo a hard collision such as

$$\begin{aligned} Q + \overline{Q} &\to Q + \overline{Q} + g \\ Q + \overline{Q} &\to Q + \overline{Q} + Q + \overline{Q} \end{aligned} \tag{2.1}$$

with g denoting a hard (pointlike) gluon with high $p_\perp$. These being processes dependent on short distances only, they can in principle be calculated via perturbative QCD. The first lowers the $Q\overline{Q}$ subenergy by a large amount at the cost of production of a jet of gluonium quanta. The second leaves two lower-energy $Q\overline{Q}$ quarkonium systems. Inasmuch as there is a finite probability <1 per period of revolution for such hard processes, the width for decay via them will be a finite fraction <1 of the level spacing. (In fact, in 1 + 1 dimensional QCD, decay into quarkonium pairs is the only mechanism available and specific calculations support this conclusion.[10])

Turning to "soft" processes, we can imagine $Q\overline{Q}$ pair creation by the string itself, provided it is long enough to contain internal energy greater than the rest mass of the pair; i.e.,

$$TL \geq 2M_Q$$

where the string tension T is, again, the energy per unit length. The probability of such a long string breaking into a smallish $Q\overline{Q}$ system would be expected a priori to be extremely small. The QED analogy is electron-positron pair creation in a constant electric field $\vec{E}$. This was calculated long ago by Heisenberg and Euler;[11] the answer is[12]

$$\frac{dP}{dt} \simeq V\, e^{-\sqrt{\pi/\alpha}(M^2/E)} \tag{2.2}$$

Converting to our units, the probability per unit time would be

$$\Gamma \sim LT\, e^{-\sqrt{\frac{\pi}{\alpha_s}\frac{A}{T}}\,M^2} \tag{2.3}$$

where A is the area of the string, and we have made the identification of E^2 (energy density) with Γ/A (string tension per unit area). Thus the probability of soft breaking of the string is exponentially small because of the large quark mass. Notice however that there is every reason to expect that when the light quarks are introduced this mechanism will be important.

Finally we come to the most important mechanism — soft radiation of gluonium. By this we mean radiation of gluonium in the direction of motion of the $\overline{Q}$: it is simply bremsstrahlung. Because gluonium has a non-perturbative, extended structure its coupling constant to heavy quarks is uncertain. But there is not much reason[13] to expect it to be extremely small. It follows that as the Q moves away from $\overline{Q}$, the rate of conversion of quark energy into the gluonium radiation should be comparable to the rate of energy delivered into the lengthening string. Thus per period the energy lost to gluonium radiation is a finite fraction of the total excitation energy of the system. We conclude that the widths of the excited states are broad compared to the level spacings. The widths may even be a finite fraction of the mass of the system. Hence the discrete levels we maintained up to this point dissolve into a continuum. Actually the details are not too important, because there will clearly be quite a change when we go to Stage III. Suffice it to say that in Stage II, heavy quark pair production is never of great importance and the central mechanism of energy storage in excited quarkonia is in the universal string, while the central mechanism of energy dissipation is in soft gluonium radiation.

Before leaving Stage II, we should mention that the ground state quarkonia $Q\overline{Q}$ can annihilate into 2 or 3 hard gluons, via a process calculable in perturbation theory. This, along with the hard-gluon bremsstrahlung in $Q\overline{Q}$ collisions, provides us with an argument that gluonium states must exist. After all, one might try to entertain the idea that pure QCD is trivial; i.e., while pointlike gluons exist at short distances, all they do is make strings but no finite-mass (as opposed to infinite-mass) particles to populate an LSZ-type Hilbert-space. However, by the above hard processes the quark sources produce at short distances distinct gluon quanta carrying away energy and directed momenta. Energy conservation demands this be materialized into asymptotic states of finite energy. Given our assumption 3 of soft confining forces, it is hard to come up with a scenario of the dynamics which avoids these gluon jets being composed of a collection of gluonium states.

The existence of strings in Stage II also has implications for Stage I: there should exist in pure QCD metastable closed strings of large circumference — a kind of soliton. Such strings will shrink as they emit gluonia, eventually merging into the quantum gluonium spectrum.[14]

Finally we go to Stage III, where light quarks are introduced. A light quark q is one whose mass (as measured by short-distance probes) is small compared to the confinement scale. Thus u, d, s, may be considered light, and charm is borderline, but probably better considered heavy: pair creation of $c\bar{c}$ in hadron processes is relatively rare. The new features emergent with inclusion of light quarks are of course the spectrum of ordinary hadrons, and the dissolution of the universal string. There is no longer any rationale for string production. The color of a heavy quark Q can be screened locally by production of $q\bar{q}$ pairs — a process which is energetically cheap. Furthermore even were a string to be made, it would rapidly break into many pieces through the Heisenberg-Euler pair-creation mechanism we already discussed.

We may ask what happens to the gluonia. They can now decay into ordinary mesons made of $q\bar{q}$ (cf., Figure 2). There is an important issue of how much Zweig-rule suppression of these decays is operational here. If there is enough for the width to be small compared to 100 MeV, the gluonia should remain distinct members of the family of hadron resonances. Although the issue seems to involve a certain amount of witchcraft, consensus (but <u>not</u> unanimous) seems to be that in fact the states ought to be narrow. We return to this question again in Section 5. Finally, how light quarks effect confinement in collision processes — especially deep-inelastic processes — has been discussed in detail extensively in the past, and we shall not cover this ground again here.[15]

We have given very little attention to Stage III, which is, after all, real life. Why should Stages I and II be relevant at all? The reason lies in the belief (!) that the phenomenon of color confinement is a universal one Examination of the perhaps more limited and tractable issues embodied in Stages I and II might shed enough light on the nature of confinement so that, even though Stage III may be more complicated to

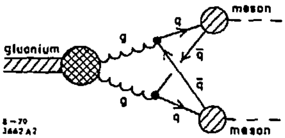


Fig. 2. Mechanism for gluonium decay into ordinary mesons.

manage mathematically, the notion of color-confinement will no longer be an implausible or mysterious one.

It may be that confinement is not universal, e.g., in Stages I and II gluons and heavy quarks are not confined, and only after light quarks are introduced does confinement emerge. This would entail a more specific view of confinement dynamics, and might well be an encouragement to search for mechanisms of confinement which do not depend upon an underlying non-Abelian gauge theory.

3. QUANTIZATION: INTRODUCING THE MICROUNIVERSE

We now turn to the formulation of the theory. [Those who are faint-hearted and frightened by simple equations are invited to rejoin us below Equation (3.43).] Up to a certain point, this is a parroting of the canonical quantization of QED. Whenever one is in doubt he should retreat to QED by replacing the triplet of quark fields

$$q(x) = \begin{pmatrix} q_1(x) \\ q_2(x) \\ q_3(x) \end{pmatrix} \tag{3.1}$$

by a single electron field $\psi(x)$ and by replacing the octet of gluon fields, described by a 3 × 3 traceless matrix[16]

$$\underset{\sim}{A}_\mu(x) = \frac{1}{2} \sum_{A=1}^{8} \underset{\sim}{\lambda}^A A_\mu^A \tag{3.2}$$

by the electromagnetic field. The steps for constructing the theory are:

1. Start with the free Dirac equation(s) for the 3 quarks

$$(i\not{\partial} - m)\, q(x) = \left(i\gamma_\mu \frac{\partial}{\partial x_\mu} - m\right) q(x) = 0 \tag{3.3}$$

2. <u>Make the gauge invariant substitution</u>

$$\frac{\partial}{\partial x_\mu} \to D^\mu \equiv \frac{\partial}{\partial x_\mu} + ie A^\mu(x) \tag{3.4}$$

which leads to

$$(i\nabla - eA - m)\, q(x) \equiv (i\not{D} - m)\, q(x) = 0 \tag{3.5}$$

Here e will be a <u>strong</u> coupling constant e_s. (More about its normalization later.)

3. <u>Look at local gauge transformations, analogous to those in QED</u>

$$q(x) = S(x)\; q'(x) \tag{3.6}$$

[$S(x)$ is a unitary 3 × 3 matrix that depends upon space and time]

and find out the transformation law that A must satisfy in order to keep the equation form-invariant. A few lines of algebra show that it is

$$A_\mu = S A'_\mu S^{-1} - \frac{i}{e} S \frac{\partial S^{-1}}{\partial x^\mu} = S A'_\mu S^{-1} + \frac{i}{e}\left(\frac{\partial S}{\partial x^\mu}\right) S^{-1} \tag{3.7}$$

This may look a little unfamiliar; in QED

$$S = e^{-ie\Lambda(x)} \tag{3.8}$$

and evidently commutes with A because the "matrices" are 1 × 1; hence in QED

$$A_\mu = A'_\mu + \frac{\partial \Lambda}{\partial x^\mu} \tag{3.9}$$

(In QCD there will be somewhat more emphasis on the actual gauge transformations $S(x)$ and less on their generators $\Lambda(x)$.)

4. <u>Build the gauge fields $F_{\mu\nu}$</u>

There are smooth ways and clumsy ways of doing this. A smooth way is to notice that in QED

$$\left[D_\mu , D_\nu\right]\phi(x) = -ie\ F_{\mu\nu}\ \phi(x) \tag{3.10}$$

for any function ϕ, and with the covariant derivative

$$iD^\mu = i\frac{\partial}{\partial x_\mu} - eA^\mu \tag{3.11}$$

already defined in Equation (3.4). The same device works here, leading to the definition

$$\underset{\sim}{F}_{\mu\nu} = \frac{i}{e}\left[\underset{\sim}{D}_\mu,\underset{\sim}{D}_\nu\right] = \partial_\nu\underset{\sim}{A}_\mu - \partial_\mu\underset{\sim}{A}_\nu + ie\left[\underset{\sim}{A}_\nu,\underset{\sim}{A}_\mu\right] \tag{3.12}$$

The extra term quadratic in the $\underset{\sim}{A}$'s coming from the non-commuting 3×3 matrices causes all the extra grief in QCD. This will make the Maxwell equations nonlinear: the gauge fields, which carry color, couple to themselves as well as to matter fields.

The fields $\underset{\sim}{F}_{\mu\nu}$, unlike the $\underset{\sim}{A}_\mu$, do not both translate and rotate under a gauge-transformation, they only rotate. For example from Equation (3.12) it follows that

$$\underset{\sim}{F}_{\mu\nu} = \underset{\sim}{S}\ \underset{\sim}{F}'_{\mu\nu}\ \underset{\sim}{S}^{-1} \tag{3.13}$$

The easy way to check this is to notice that the covariant derivative also only rotates

$$\underset{\sim}{D}_\mu = \underset{\sim}{S}\ \underset{\sim}{D}'_\mu\ \underset{\sim}{S}^{-1} \tag{3.14}$$

i.e., for all functions ϕ,

$$\left(i\partial_\mu - e\underset{\sim}{A}_\mu\right)\phi(x) = \underset{\sim}{S}(x)\ \left(i\partial_\mu - e\underset{\sim}{A}'_\mu\right)\underset{\sim}{S}^{-1}(x)\ \phi(x) \tag{3.15}$$

Then we can again use Equation (3.12) to establish Equation (3.13).

5. Write down the "Maxwell Equations"

To do this, notice that in clear analogy with QED the interaction should be

$$\mathrm{Tr}\, \underset{\sim}{J}_\mu \underset{\sim}{A}^\mu = \bar{q}\gamma_\mu \underset{\sim}{A}^\mu q = \frac{1}{2}\sum_{A=1}^{8} \bar{q}\gamma_\mu \underset{\sim}{\lambda}^A q\, \underset{\sim}{A}_A^\mu \tag{3.16}$$

which serves to define the 3 × 3 matrix $\underset{\sim}{J}_\mu$. Under the gauge transformation, Equation (3.6), it is not hard to show[17] that J_μ will rotate

$$\underset{\sim}{J}_\mu = \frac{1}{2}\sum_{A=1}^{8} \left(\bar{q}\gamma_\mu \underset{\sim}{\lambda}^A q\right) \underset{\sim}{\lambda}^A = \underset{\sim}{S}\, \underset{\sim}{J}_\mu' \, \underset{\sim}{S}^{-1} \tag{3.17}$$

Thus, to be gauge-invariant, the Maxwell equations must behave in the same way. This is again ensured by constructing them via use of the covariant divergence:

$$\left[\underset{\sim}{D}_\mu , \underset{\sim}{F}^{\mu\nu}\right] = \left[\partial_\mu + ie\, \underset{\sim}{A}_\mu , \underset{\sim}{F}^{\mu\nu}\right] = e\underset{\sim}{J}^\nu \tag{3.18}$$

These equations along with the Dirac equation can be derived from a Lagrangian,[18] but we bypass all that and take directly Equations (3.5), (3.12) and (3.18) to define the dynamics of the theory.

So far, we have mainly seen similarities of QCD to QED. As we already noted, the big difference between QED and QCD is that there is a term quadratic in $\underset{\sim}{A}_\mu$ in $\underset{\sim}{F}_{\mu\nu}$. Along with the term in $\underset{\sim}{D}_\mu$ proportional to $\underset{\sim}{A}_\mu$, this means that the left-hand side of the Maxwell equations have terms quadratic and cubic in $\underset{\sim}{A}_\mu$. This causes all the headaches in QCD (but, one hopes, also the seeds of the confinement phenomenon). The physics of this is that the gluon fields themselves possess color and therefore act as an additional source of color field.

We now must replace the classical variables with operators (in Heisenberg picture) and quantize. To implement this we will go to a Hamiltonian formalism. Before doing this, it is useful to

6. Choose a gauge (for us, $\underset{\sim}{A}_0 = 0$)

We do this with an eye toward the next steps, namely identifying canonical variables and carrying out a canonical quantization procedure. Before proceeding we remark that this is not the only path to take at this point. For small e we may just follow Book I, identify Feynman rules, and start calculating. A more formal but very efficient way of doing that is to get the rules via the path-integral formalism. For covariant formalisms, one runs into technical complications — the so-called Faddeev-Popov ghosts.[19] The canonical formalisms avoid these and amount to choosing one space-time component of the potentials $\underset{\sim}{A}_\mu$ to vanish. There are three basic choices: either $\underset{\sim}{A}_0 = 0$ (temporal or canonical gauge), $\underset{\sim}{A}_3 = 0$ (axial gauge), or $\underset{\sim}{A}_0 + \underset{\sim}{A}_3 = 0$ (null-plane or light-cone gauge). The latter two, which pick out a space axis, are convenient for collinear collision processes—namely, the parton-model type of QCD application. We shall not be so interested in that kind of thing at this point, but rather in large-distance questions associated with confinement in pure QCD, and then in the properties of the theory with additional heavy quark sources added. Therefore, $\underset{\sim}{A}_0 = 0$ gauge is convenient, especially since this keeps $\underset{\sim}{D}_0 = \partial/\partial t$ clean and helps to isolate the real dynamics from the phony dynamics of time-dependent gauge-transformations.

To see that this choice is always possible, return to Equation (3.7) and set $\underset{\sim}{A}_0' = 0$. This implies we must find an $\underset{\sim}{S}$ such

$$\underset{\sim}{A}_0 = \frac{i}{e} \frac{\partial \underset{\sim}{S}}{\partial x_0} \underset{\sim}{S}^{-1} \tag{3.19}$$

or

$$\frac{\partial \underset{\sim}{S}}{\partial t} = -ie \, \underset{\sim}{A}_0 \, \underset{\sim}{S} \tag{3.20}$$

For gauge potentials which vanish at $t = -\infty$, this is a Schrodinger equation and can be solved formally

$$\underset{\sim}{S} = T \, e^{-ie \int_{-\infty}^{t} \underset{\sim}{A}_0(\vec{x}, t') \, dt'} \tag{3.21}$$

with T the time-ordering symbol. (Remember $\underset{\sim}{A}_0$ is a 3×3 matrix; it doesn't commute with itself at different times.) Now

7. <u>Write out the equations of motion in $A_0 = 0$ gauge</u>:

Dirac Equation

$$i \frac{\partial q}{\partial t} = \left\{ \vec{\alpha} \cdot (\vec{p} - e\vec{\underset{\sim}{A}}) + \beta m \right\} q \tag{3.22}$$

"Ampere's Law":

$$\frac{\partial \underset{\sim}{E}^i}{\partial t} = (\vec{\underset{\sim}{D}} \times \vec{\underset{\sim}{B}})^i - eJ^i \equiv (\vec{\nabla} \times \vec{\underset{\sim}{B}})^i - ie\epsilon_{ijk}\left[\underset{\sim}{A}^j, \underset{\sim}{B}^k\right] - eJ^i \tag{3.23}$$

Definition of B:

$$\vec{\underset{\sim}{B}}^i = (\vec{\underset{\sim}{D}} \times \vec{\underset{\sim}{A}})^i \equiv (\vec{\nabla} \times \vec{\underset{\sim}{A}})^i - ie\epsilon_{ijk}\left[\underset{\sim}{A}^j, \underset{\sim}{A}^k\right] \tag{3.24}$$

Definition of E:

$$\vec{\underset{\sim}{E}} = -\frac{\partial \vec{\underset{\sim}{A}}}{\partial t} \tag{3.25}$$

"Gauss' Law":

$$\underset{\sim}{D} \cdot \underset{\sim}{E} \equiv \vec{\nabla} \cdot \vec{\underset{\sim}{E}} - ie\left[\underset{\sim}{A}^i, \underset{\sim}{E}^i\right] = e\underset{\sim}{J}_0 \tag{3.26}$$

These look structurally like the QED Maxwell's equations, as written on countless T-shirts. There are only three distinctions (other than the ubiquitous 3 × 3 matrices). Two are the nonlinear terms in Ampere's Law and in the definition of $\underline{B}$, and the third is the presence of the covariant divergence $\vec{\underline{D}}$ in Gauss' law. Notice that Gauss' Law is an equation of constraint, not an equation of motion. This will be important in the subsequent interpretation of the theory.

With such a similarity to QED it should be no surprise that the above equations of motion can be derived from a Hamiltonian which is essentially the same as in QED. This leads us to the next step:

8. Construct a Hamiltonian and identify canonical coordinates:

The result is an utterly unexceptional analogue of the QED Hamiltonian (in this gauge):

$$H = \int d^3x\,\mathcal{H} = Tr\int d^3x\left[\underline{E}^2(x)+\underline{B}^2(x)\right] + \int d^3x\; q^\dagger\left\{\vec{\alpha}\cdot(\vec{p}-e\vec{\underline{A}})+\beta m\right\}q$$

$$= \frac{1}{2}\sum_{A=1}^{8}\int d^3x\left[E_A^2(x)+B_A^2(x)\right] + \int d^3x\; q^\dagger\left\{\vec{\alpha}\cdot(\vec{p}-e\vec{\underline{A}})+\beta m\right\}q \qquad (3.27)$$

Just as in QED, the canonical coordinates are the $A(x)_i^A$ (cf. Equation (3.2)), and the momenta conjugate to $A(x)_i^A$ are the $E(x)_A^j$. Because the magnetic field

$$\underline{B}^i(x) = \epsilon_{ijk}\left[\underline{D}_j(x)\,,\,\underline{A}^k(x)\right] \qquad (3.28)$$

has a term quadratic in $\underline{A}$, the Hamiltonian has terms cubic and quartic in $\underline{A}$. The Hamilton equations of motion are

$$\frac{\partial H}{\partial p} = \dot{q} \Longleftrightarrow \frac{\partial \mathscr{H}}{\partial E^i_A} = E^i_A = \frac{\partial A^A_i}{\partial t} = -\frac{\partial A^i_A}{\partial t}$$

$$\frac{\partial H}{\partial q} = -\dot{p} \Longleftrightarrow \frac{\partial \mathscr{H}}{\partial A^A_i} = -\frac{\partial E^i_A}{\partial t} = -(\vec{D} \times \vec{B})^i_A + eJ^i_A = -\frac{\partial E^i_A}{\partial t} \qquad (3.29)$$

(This second equation requires a little care in the integrations by parts; the covariant gradient $\vec{\underset{\sim}{D}}$ can be integrated by parts, just like $\vec{\nabla}$.)

Let us ignore for a while the Hamiltonian of the fermions. Let us furthermore temporarily imagine replacing the integral $\int d^3x$ by a sum over a million coordinates x_i. (That should be fine-grained enough for a lot of purposes.) Then each $A^A_i(x)$ is a canonical coordinate q^α and each $E^i_A(x)$ is its conjugate momentum p^α. The Hamiltonian of this somewhat mutilated version of pure QCD is nothing more than a good old Schrodinger Hamiltonian

$$H = \sum_{\alpha=1}^{24\times 10^6} \frac{1}{2} p_\alpha^2 + V\left(q_1,\ldots,q_{24\times 10^6}\right) \qquad (3.30)$$

of a particle in a 24-million dimensional space moving in a static potential. The potential (the $\underset{\sim}{B}^2$ term) is non-negative and at most quartic in the coordinates. It is important not to forget this homely analogy, especially when the going gets rougher (as it regrettably will). Finally

9. Impose canonical commutation relations on the conjugate variables $\underset{\sim}{A}(x)$ and $\underset{\sim}{E}(x)$. It is time (in fact somewhat overdue) to make clear our conventions of notation and normalization. The 3×3 matrices $\underset{\sim}{A}(x)_i$ are expanded in terms of 8 independent canonical fields which are coefficient of Gell-Mann's λ-matrices:

$$\underset{\sim}{A}_i(x) = \sum_{A=1}^{8} \frac{1}{2} A_i^A(x) \, \underset{\sim}{\lambda}^A \tag{3.31}$$

where

$$\underset{\sim}{\lambda}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \underset{\sim}{\lambda}_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \underset{\sim}{\lambda}_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

$$\underset{\sim}{\lambda}_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \underset{\sim}{\lambda}_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad \underset{\sim}{\lambda}_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

$$\underset{\sim}{\lambda}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \underset{\sim}{\lambda}_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{3.32}$$

Thus

$$\mathrm{Tr}\, \underset{\sim}{\lambda}_A \underset{\sim}{\lambda}_B = 2\delta_{AB} \tag{3.33}$$

with

$$\left[\underset{\sim}{\lambda}_A , \underset{\sim}{\lambda}_B\right] = 2if_{ABC}\, \underset{\sim}{\lambda}_C \tag{3.34}$$

defining the structure-constants of the SU(3) algebra. There is an analogous definition for $\underset{\sim}{E}(x)$, and the canonical equal time commutation relations are taken to be[20]

$$\left[E_A^i(\vec{x},t) \, , \, A_j^B(\vec{y},t)\right] = -i\delta_{ij}\, \delta^{AB}\, \delta^3(\vec{x}-\vec{y}) \tag{3.35}$$

These give a consistent canonical formulation and lead with no difficulty to interpreting two out of our of the Maxwell-equations in terms of the Hamiltonian equations of motion

$$\frac{\partial A_j^A}{\partial t} = -i\left[H, \, A_j^A(x)\right] = E_A^j \qquad i\left[H, \, E_A^j(x)\right] = \frac{\partial A_j^A}{\partial t} \tag{3.36}$$

Finally, before leaving this introduction to formal quantization of QCD, we should add one more item to the list:

10. Check the consistency of the quantization procedure with Lorentz covariance

We shall not do this here; the only substantive problem lies with Lorentz-boosts. There as in QED, the generator should include a gauge-transformation of the fields designed to restore $A_0 = 0$ gauge in the new frame. The formal covariance of temporal-gauge quantization has been checked,[21] although I have not found a direct check along the lines used in Book II.

The astute reader familiar with Chapter 14 of Book II will notice that we have not been very good parrots: there are some differences in what is done there and what we do here. In Book II, the longitudinal degrees of freedom of the electromagnetic potential $\vec{A}(x)$ are eliminated from the beginning by a choice of gauge, leaving only the transverse field to describe the two physical polarization states of the photon. In what is done above, longitudinal photons (or gluons) are quantized as well. What does this mean? It appears we have overachieved and introduced too many coordinates into the dynamics. The key to understanding this lies in Gauss' law, which we here interpret as an equation of constraint on allowed solutions of the Great Big Schrodinger Equation. That is, not all solutions of

$$H\Psi_\varepsilon\left(q_1,\ldots,q_{24\times 10^6}\right) = \varepsilon\Psi_\varepsilon\left(q_1,\ldots,q_{24\times 10^6}\right) \tag{3.37}$$

are deemed physically acceptable-- only those for which

$$\left\{\vec{\underset{\sim}{D}}\cdot\vec{\underset{\sim}{E}}(x) - e\underset{\sim}{J}_0(x)\right\}\Psi_\varepsilon = 0 \tag{3.38}$$

are allowed physical states. Note that (in QCD) this is a set of 8×10^6 equations; one for each color and coordinate x! The meaning of all this is somewhat clearer when one identifies (cf. Appendix A for details) the Gauss' law operator in the above equation as the generator of time-independent gauge-transformations. Equation (3.38) means that the Great Big Wave-function Ψ_E is required to be invariant under time-independent gauge transformations. Not all solutions of the Great Big Schrodinger Equation will have this property. If we think of these gauge-transformations as "rotations" of the coordinates it means Ψ_E must be "S-wave" in 8×10^6 of the 24×10^6 coordinates, if it is to be physically acceptable. Notice also that because H is invariant under time-independent gauge transformations, it commutes with the Gauss' Law operator,[22] so these 8×10^6 coordinates are symmetry degrees of freedom.

What should one do about this? One way, appropriate to QED, is to identify the (in that case) 10^6 symmetry coordinates and their conjugate momenta and explicitly excise them from the Hamiltonian and therefore from the formalism. Because of the linearity of QED this is easy to do. One breaks up the field into longitudinal and transverse pieces

$$\vec{A} = \vec{A}_L + \vec{A}_T$$
$$\vec{E} = \vec{E}_L + \vec{E}_T \tag{3.39}$$

and writes

$$H = \int d^3x \, \frac{1}{2}\left[E_L^2(x) + E_T^2(x) + (\vec{\nabla} \times \vec{A}_T)^2\right] + \int d^3x \, e^\dagger(x) \left| \vec{\alpha} \cdot (\vec{p} - e\vec{A}_L - e\vec{A}_T) + \beta m \right| e(x) \tag{3.40}$$

Gauss' Law allows one to eliminate $\vec{E}_L$

$$\vec{E}_L(x) = e\vec{\nabla}\frac{1}{\nabla^2}J_0(x) = -e\vec{\nabla}_x\int d^3y\,\frac{1}{4\pi|x-y|}J_0(y) \tag{3.41}$$

in favor of the instantaneous Coulomb-interaction

$$\frac{1}{2}\int d^3x\,E_L^2(x) = \frac{e^2}{2}\int d^3x\,d^3y\,J_0(x)\,\frac{1}{4\pi|x-y|}J_0(y) \tag{3.42}$$

Then the only vestige of longitudinal degrees of freedom left is the $\vec{\alpha}\cdot\vec{A}_L$ term in the Dirac Hamiltonian. A phase transformation on the Dirac field easily removes this as well.

In QCD it is much more difficult to eliminate the "trivial" symmetry variables, because of the nonlinear terms in H which couple together longitudinal and transverse modes of the fields. Gribov[23] has demonstrated specific difficulties which block a generalization of the QED Coulomb-gauge quantization procedure. These are sketched in Appendix A; suffice it to say that things go along reasonably in parallel with QED, until the point at which the inverse operation $1/\nabla^2$ (in other words, instantaneous r^{-1} potentials) appeared in the formalism. The corresponding operator emergent in QCD is $1/\vec{\nabla}\cdot\vec{\underset{\sim}{D}}$. But Gribov showed that there exist gauge potentials $\vec{A}(x)$ for which homogenous solutions exist

$$\vec{\nabla}\cdot\vec{\underset{\sim}{D}}(A)\,\phi(x) = 0 \tag{3.43}$$

so that $\vec{\nabla}\cdot\vec{\underset{\sim}{D}}$ is not invertible. This problem has deep ramifications and is difficult if not impossible to elude by alternative choices of gauge-fixing. At present there exists <u>no</u> fully satisfactory formulation of QCD in terms of only physical degrees of freedom.[24]

Given this situation, an alternative is just to leave in the extra degrees of freedom. After all, the quantization of the theory is satisfactory and the problem is only one of frustration in not properly coping with a gigantic residual symmetry. And it may be no worse to deal with 24×10^6 coordinates than 16×10^6. The problem is that explicit invariance under time-independent gauge transformations must be maintained in every subsequent step in the development of the theory. Otherwise the unphysical solutions of the Great Big Schrodinger Equation will mix with the physical ones, and there is no assurance that this does not create some kind of nonsense in the resultant physics.

We emphasize that this difficulty only occurs for sufficiently large gauge-fields, much larger (roughly by a factor α_{st}^{-1}) than commonly found in the weak-coupling limit. Thus weak-coupling, short-distance applications should hopefully not be affected by the gauge-fixing ambiguity. In particular it will not be seen in any finite order of perturbation theory.

Let us now summarize where things stand. We have seen that it is possible to formulate (in $\underline{A}_0 = 0$ gauge) a straightforward quantization of QCD in a way analogous to QED. Nevertheless there arose some problems, which we here enumerate:

a. In the presence of strong fields the gauge cannot be completely fixed in a satisfactory manner: unphysical degrees of freedom are quantized and a careful selection of the physical subset of solutions to the Great Big Schrodinger Equation must be made, namely those consistent with Gauss' Law.

b. Even assuming weak coupling, the basic quanta in the formalism are colored quarks and gluons, which do not (or at least <u>should</u> not) appear as asymptotic states in real life. Thus there is no obviously satisfactory way of introducing and constructing an S-matrix as done in QED,[25] without first resolving the confinement question. It seems necessary to solve QCD in order to formulate it.

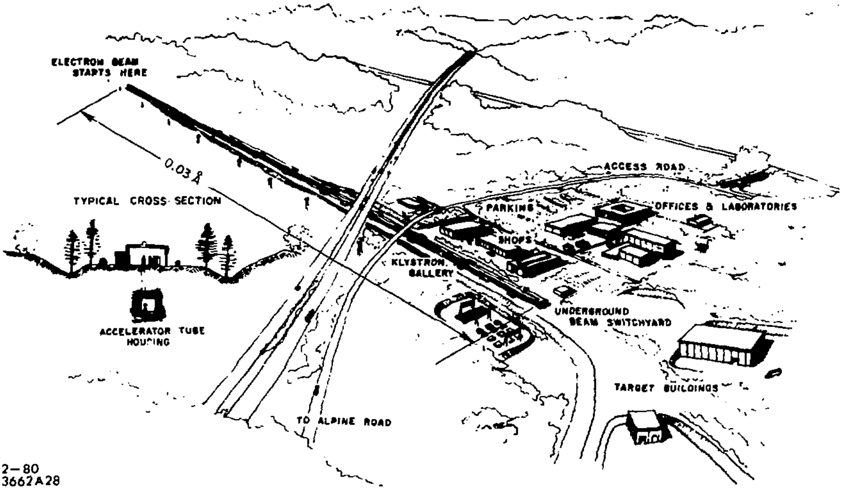
c. Asymptotic freedom implies that although at short distances the coupling constant may be chosen (consistently) to be small, it becomes large at large distances. But then the nonlinearities of the theory become pernicious.

In addition to the above three problems, which are no doubt interconnected, there is a fourth which is to be discussed in Section 4 (in order to spare the reader too much of a dose of formalism all at one blow):

d. The instanton phenomenon complicates the structure even of the QCD vacuum. Effects of small instantons can be controlled, but what happens at larger distances is not in good control.

All the above complication is minimized if we could somehow restrict everything about the formulation to short-distances. It is possible to do this: we can just quantize the theory in a sufficiently tiny box. While this obviously leaves something to be desired, it does provide solutions to the above problems: perturbation-theory is universally applicable; the coupling-constant can be taken as always small, and the dangerous large instantons don't fit in the box. Quarks and gluons are the physical quanta and <u>are</u> unconfined, provided the confinement radius is large compared to the size of the box. Gauge-fixing ambiguities do not arise because the field strengths relevant to typical scattering processes are too small.

But is this really satisfactory? Have we destroyed QCD in order to save it? After all, we do not fit into such a box, and what is a quantum theory if it does not include us as observers? I think this aspect can in principle be addressed in a reasonably satisfactory way. Let us imagine that there exists a world out there built out of superheavy charged fermions which we call femtofermions and which interact with each other by exchange of a different octet of femtogluons carrying femtocolor. The mass-scale of this world is taken to be 10^{15} times larger than our own; hence by dimensional analysis the distance-scale is $\sim 10^{-15}$ that of our own (hence the femto-prefix). The femtoquarks will bind into femtonucleons and femtomesons. If we give the femtoquarks electric charges $\pm 2/3$, $\pm 1/3$ in the usual way, and introduce a femtoelectron of mass 5×10^{14} MeV, we can make femto-nuclei, femto-atoms, and so on, up to — you guessed it — femtophysicists. Femtophysicists will build accelerators, magnets (with field-integral $\int B \, d\ell \sim 10^{15}$ times ours), and the like, and do experiments on not only the constituents of femtomatter, but also the constituents of ordinary matter. How? Figure 3 shows a way: build an accelerator (the one shown, a remarkable technical accomplishment of the femtophysicists, is ~.03 Å in length) which can create a photon beam: ordinary quarks and leptons can be pair-produced and then separated from femtomatter beams. Thus femtophysicists in principle could scatter ordinary quarks and gluons from each other and measure the relevant S-matrix elements. They could even communicate the results to us via electromagnetic radiation. (However femtophysicists would have to be patient souls: one bit of information transmitted per ten femtodays corresponds to ~ one bit per nanosecond to us.)



2—80
3662A28

Fig. 3. Femto-SLAC.

Could actually such a femtouniverse exist? If one were nearby we would certainly know it. Just the black-body radiation from a femto-earth would consist of 10^{53} photons per second, each of energy ~20 TeV. Be sure to avoid the femto-sun. However, none of this is realistic: we have ignored gravity. Gravity is not a scale invariant interaction; the gravitational force between a femtoproton and femtoelectron is $\sim 10^{-9}$ the electromagnetic force. The gravitational force between "macroscopic pieces of femtomatter will be crushingly large.

But we digress. While our femtouniverse is not completely realistic, it should indicate that even with a tiny quantization-volume, we can think in reasonably physical terms about scattering processes, now involving quarks and gluons — and not gluonia, quarkonia, or hadrons. There is one major problem of principle: a typical beam pipe at femto-SLAC has a diameter $\lesssim 10^{-14}$ cm. The beams will be even smaller; hence they have an unavoidably large $\langle p_\perp \rangle$, large compared to the confinement-scale of a few hundred MeV. Thus quarks and gluons in the beam will have large momentum uncertainty — we can consider them as virtual, or off-shell. There will also be infrared-divergence problems in calculation of S-matrix elements similar to those encountered in the QED of massless electrons. In that case Lee and Nauenberg[26] showed that the conventionally defined S-matrix depends upon the quantization volume. Another way of saying this might be that the outcome of physical scattering experiments depends upon the details of the preparation of initial beams as well as the properties of the final-state detection apparatus. Kinoshita as well as Lee and Nauenberg,[26] have given formal recipes on how to avoid these infrared problems by summation over groups of initial as well as final states. I am not sure these problems are fully under control in QCD. But here we shall assume they are. Specifically, we assume that in the femtouniverse

a. Asymptotic fields and states of "physical" colored quarks and gluons exist in the way described by LSZ.[25]

b. These fields can be formally related to the fields appearing in the equations of motion by a U-matrix (cf., Book II, Chapters 16-17) which can be calculated perturbatively by a diagrammatic expansion.

While we have not explicity done all this, I am confident that the net result of such a program would be the rules for diagrams obtained by Book I or path-integral methods.

What good is all this? Perhaps we can provide a satisfactory theory for femtophysicists, but does that help us directly? While I am not sure of the answer, I think it is yes. First of all, it should be the case that any process which in principle can be measured by femtophysicists as well as by us can be calculated perturbatively. Let us look at some good examples:

a. ($e^+e^- \to$ hadrons). To lowest order in α, this one is fine. Femtophysicists should be able to make femto PEP. $R(e^+e^- \to$ hadrons) appears to be a very clean QCD calculation.

b. $\sigma(\gamma\gamma \to$ hadrons). For this one we should restrict our attention to virtual photons. There may still be a problem because of presence of large longitudinal distances at high energies in this process (cf., Figure 4). The distance-scale is $\Delta z \sim E_\gamma/m^2_{q\bar{q}}$ where $m^2_{q\bar{q}}$ ($\gg 1$ GeV^2) is the mass of the virtual fermion pair created by the γ-ray.

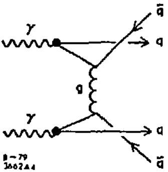


Fig. 4. Large longitudinal distances in $\gamma-\gamma$ scattering.

c. ($e + q \rightarrow e + q$ + gluons). For large x, this process depends on distances smaller than the size of the box. For small x, large longitudinal distances are again important, and infrared effects associated with the initial beams may need to be taken into account.

d. ($q + \bar{q} \rightarrow \mu^+ + \mu^-$ + gluons). The Drell-Yan process[27,28] appears to be applicable — at least for the total cross section (where one measures only the attenuation of the initial quark beams). The $p_\perp$ distribution of dileptons appears also to be an acceptable process, <u>but only at large</u> $p_\perp$.

Processes involving hadrons must of course take into account the relationship (if any) between the quark and gluon beams of the femtophysicist and the equivalent quark and gluon beams of the parton model. For sufficiently coarse-grained measurements involving hadrons it is plausible that the two kinds of incident beams are equivalent. In the case of the deep-inelastic structure functions relevant to electroproduction and neutrino interactions, one can go further and derive the Q^2-dependence of the moments of the structure functions without introducing the parton model. One need only rely on the Wilson operator-product expansion.[29] However the basis of that expansion is in fact an analysis to aribtrary order of perturbation theory. We have only been able to control perturbation theory by restricting ourselves to life in the femtouniverse. But in the femtouniverse the Q^2-dependence (at very large Q^2) of electroproduction moments does not depend in any way on the composition of the quark beam. Thus it is extremely reasonable to generalize this behavior to the proton as well. Maybe it is even a rigorous statement, but I don't know enough about rigor to be able to tell.

A somewhat stronger version of this idea is widely used in applying perturbative QCD to physical processes. As we already indicated, most perturbative calculations in QCD will run into difficulty with infrared divergences associated with collinear emission of gluons or quarks. This is the kind of problem dealt with by Kinoshita and by Lee and Nauenberg[26] for massless QED, and involve macroscopic distances (where confinement effects will creep in). However one can define quantities which are free of infrared singularities and which are calculable in QCD perturbation theory. A prototypical example is the angular distribution and angular correlations of quark and gluon energy produced in e^+e^- annihilation.[30] It is an observable in the femtouniverse which is independent of details of apparatus or beams. It is tempting (and probably correct) to assert that this prediction is also true for the hadrons measured by ordinary physicists as well. If our assumption that the confinement mechanism is "soft" is true, this result will follow. But there is an additional assumption involved, namely that between the scale of the femtouniverse and the scale of ordinary macroscopic measurements, the jet of quarks and gluons is not significantly defocused or self-focused. Nevertheless, this assumption is very reasonable. Generalizing, we may postulate that whenever one finds and calculates perturbatively a QCD process which is infrared finite, the answer is correct. In other words, if the calculation is not obviously wrong, it's right. But it is possible that this postulate, while not obviously wrong, may not be right.

So far we have used the femtouniverse as a guide and criterion for the applicability of QCD perturbation theory ideas. In addition, the femtouniverse may be of use in attacking non-perturbative questions as well. Let us accept that we can understand in principle everything about the theory in a box of dimension 10^{-14} cm. Now stack together a million such boxes, with the fields coupled at the boundaries in the way appropriate for the full theory. Then the theory in the resulting cube of dimension 10^{-12} cm should already provide a description of confinement: free gluons and quarks should begin to disappear as asymptotic states, and gluonia, quarkonia, and ordinary hadrons should appear: they easily fit into a box of that size. The problem is to identify and properly couple together the crucial degrees of freedom present in the small boxes. The spirit of such an approach is very close to lattice-calculations: we return to that approach in Section 8.

A clue as to how one might proceed comes from closer study of the structure of the theory in the small box. This has been examined in some detail, although some points are not yet fully understood.[31] Everything appears to be almost the same as in QED. Of course, it is important to use Fourier series, not Fourier integrals. With use of periodic boundary conditions, the gluon state of lowest non-vanishing momentum has momentum $k = 2\pi/V^{1/3} >> 1$ GeV. As in QED, longitudinal gluons of $\vec{k} \neq 0$ can be eliminated by use of Coulomb-gauge. The only mode of the potential $A_i^A(x)$ which seems to behave in a way different from QED is the constant $\vec{k} = 0$ mode, which we label as the vacuum mode. The Hamiltonian for this vacuum mode, ignoring couplings to degrees of freedom with $\vec{k} \neq 0$, can be written

$$H = \sum_{\substack{i=1,2,3\\ A=1,\ldots,8}} \left(-\frac{1}{2}\frac{\partial^2}{\partial Q_i^{A2}}\right) + \frac{e^2}{2V}\,\epsilon_{ijk}\,\epsilon_{imn}\,f^{ABC}\,f^{AMN}\,Q_j^B\,Q_k^C\,Q_m^M\,Q_n^N \tag{3.44}$$

where

$$A_i^A(x) = \frac{1}{\sqrt{V}}\,Q_i^A + \sum_{k\neq 0}\frac{1}{\sqrt{V(2k)}}\left[a_i^A(k,\epsilon)\,e^{i\vec{k}\cdot\vec{x}} + \text{h.c.}\right]$$

$$E_A^i = \frac{1}{\sqrt{V}}\,\pi_A^i + \ldots \tag{3.45}$$

with π_i^A and Q_i^A canonically conjugate pairs; i.e.,

$$\pi_A^i = -i\,\frac{\partial}{\partial Q_i^A} \tag{3.46}$$

The problem of finding the eigenvalues and eigenfunctions of the Hamiltonian is equivalent to solving the Schrodinger equation of a single non-relativistic particle moving in a (not very symmetric) quartic potential in a 24-dimensional space. It turns out that the "particle" is confined, i.e., the wave functions are normalizable. The Hamiltonian has an energy-spectrum with many low-lying levels. Properties of the solution are as follows:

1. In low-lying states

$$\langle A^2\rangle = \frac{1}{V}\,\langle Q^2\rangle \approx (\text{Const})\,\frac{V^{-2/3}}{(e^2)^{1/3}} \tag{3.47}$$

and the dispersion around the mean value appears to be small.

2. In general the level spacings of excited states are in proportion to $(e^2/V)^{1/3}$. This is to be compared with the lowest-lying gluon modes, whose level spacings are $\sim 2\pi/V^{1/3}$. Thus the energy-spacings are

small. By the uncertainty relation $\Delta t\ \Delta E \geq 1$, this means the period of the motion is large; the motion is slow.

3. The lowest-lying set of levels appears to be a rotational band, i.e., the polarization-vector of the vacuum field tumbles in space. The energy should be given by

$$E_J \sim \frac{1}{V} \frac{J(J+1)}{\langle A^2 \rangle} \tag{3.48}$$

The effect of these vacuum rotor-modes on the physics in the femtouniverse is obscure but probably minimal. This is because of the dense level spacing of the rotor-modes. Very little energy is needed to excite rotor states, and their motions should not interfere with the dynamics of the gluons and quarks occurring on a much higher energy scale. [For instance, we do not worry much about excitation of photon states during a hard collision of a high energy hadron with a piece of matter.] However, these vacuum modes may, upon coupling of a million femtouniverses together, contribute to the long-wavelength aspects of the physics at the larger distance scale. Hence it is possible that they may be of relevance to an understanding of the confinement issue.

4. INSTANTONS

There is yet another complication present in QCD not encountered in QED. The states of the theory may be classified according to the topological structure of the gauge potentials. In our quantization procedure — especially in the femtouniverse — we considered only weak gauge-fields $\underset{\sim}{A}_\mu$ — and only pure gauge transformations which can be reached from the identity transformation continuously, i.e., by a product of infinitesimal transformations. However, there exist pure gauge potentials $\vec{A}(x)$ (potentials for which $\vec{E} = \vec{B} = 0$) which cannot be reached continuously from $\vec{A} = 0$ without pulling the gauge potential in from the boundary of the quantization volume. An example of this is a spherical "gauge bubble" (Figure 5) given by

$$e\vec{A}(x) = iU \vec{\nabla} U^{-1} \quad \text{with} \quad U = e^{-i\vec{\tau}\cdot\hat{r}f(r)} \equiv e^{-i\vec{\tau}\cdot\vec{r}f(r)/r} \qquad (4.1)$$

where $\vec{\tau} = (\tau_1,\tau_2,\tau_3) \equiv (\lambda_1,\lambda_2,\lambda_3)$, and with $f(r)$ increasing from zero at $r = 0$ to a multiple of π at $r = \infty$. If $f(r) = 0$, then $U = 1$ and if $f(r) = n\pi$, $U = (-1)^n$. Thus $\vec{A}$ is nontrivial only in the region where $\partial f/\partial r$ is nonvanishing. This region can be taken as a shell of arbitrary radius R. Continuous gauge transformation can change the shape of $f(r)$, modify R, and distort the bubble. But they cannot untangle the topological twists which come from coupling internal-symmetry generators $\vec{\tau}$ to space coordinates $\vec{r}$.

It turns out[32] that the amount of topological twist in the gauge-bubble can be characterized by a "topological quantum number":

$$N = -\frac{ie^3}{24\pi^2} \operatorname{Tr}\int d^3x \; \epsilon_{ijk} A_i(x) \, A_j(x) \, A_k(x) \qquad (4.2)$$

This formula is valid if and only if $\vec{A}$ is "pure gauge." The quantity N is invariant under continuous time-independent gauge transformations which vanish at the boundary of the quantization volume. Furthermore for A's which vanish on the boundary, N can take only integer values. We have neither motivated this construction nor demonstrated these results; that is business relegated to Appendix B. Here it suffices to

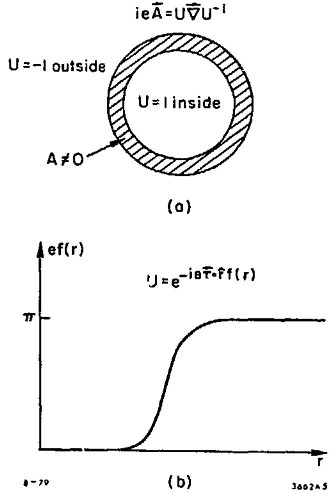


Fig. 5. A gauge-bubble of pure gauge-potential ($\vec{E} = \vec{B} = 0$).

know that there is a way of characterizing and classifying the gauge-bubbles. If we start (perturbatively) from any weak-field configuration connected smoothly to $\underset{\sim}{A} = 0$, the topological quantum number N is zero. Thus it would seem that we might be able to restrict our attention to those N = 0 configurations, since the only way to reach others is by bringing the gauge field in from the boundary of the quantization volume, and boundary-effects (for sufficiently large universes!) shouldn't matter. However, this is not the case; there exists a dynamical communication through quantum-mechanical tunneling which is a volume effect. To see how this comes about, it is necessary to again think of QCD in terms of the Great Big Schrodinger Equation with $\underset{\sim}{A}(x)$ the coordinates and $\underset{\sim}{E}(x)$ the momenta. (It is important to recall here Equation (3.30) and the accompanying admonition.) A gauge bubble with $N \neq 0$ is a configuration with classically zero energy since $V(A) = \frac{1}{2}\int B^2 d^3x = 0$ for a "pure gauge." If we go directly toward $\underset{\sim}{A} = 0$ by scaling down $\vec{\underset{\sim}{A}}$, i.e., by letting $\underset{\sim}{A}(x) \to f(t)\vec{\underset{\sim}{A}}(x)$, $0 \le f \le 1$, the magnetic field no longer vanishes because the intermediate configurations are not "pure gauge." (This can be checked by direct calculation: $\underset{\sim}{B}_i = ief(1-f)\left[\underset{\sim}{A}_j, \underset{\sim}{A}_k\right]\epsilon_{ijk}$.) Hence there is a potential barrier between configurations with $N \neq 0$ and $N = 0$; the potential energy for all paths in A-space which smoothly connect $\underset{\sim}{A} = 0$ with the gauge bubble is non-vanishing.[33]

Now if we start with a (physical) solution $\Psi_0(A)$ of the Great Big Schrodinger Equation restricted to $\underset{\sim}{A}$'s with $N = 0$, it would appear to be possible for the wave function to leak through the barrier. This turns out to be the case: the wave function $\Psi_0(A)$ can tunnel into $N = \pm 1$ configurations and onward.

Schematically, the situation looks like Figure 6; the potential is (in a sense) periodic. The important tunneling couplings will evidently be nearest-neighbor with $|\Delta N| = 1$. It should not be surprising to anyone slightly familiar with solid-state physics that the energy eigenfunctions will be Bloch-waves

$$\Psi_\theta(\underset{\sim}{A}) = \sum_{N=-\infty}^{\infty} e^{iN\theta}\, \Psi_N(\underset{\sim}{A}) \tag{4.3}$$

This is known as the θ-basis. Physical observables will depend upon θ, but must depend on it only to the extent that the tunneling amplitudes are non-vanishing. Therefore, it is important to estimate the tunneling. For small e, one must tunnel through a thick barrier (cf., Equation (4.2) which implies $|e \cdot \bar{A} \cdot L| \gtrsim 1$, with L the size of the region in space in which $\underset{\sim}{A}(x)$ is non-vanishing), and a semiclassical estimate is appropriate. The rule for obtaining this amplitude can be abstracted from experience with nonrelativistic quantum-mechanics, and some details of how to get it are in Appendix B. The recipe is as follows. To find the amplitude:

1. Change t to it (i.e., go to Euclidean metric).

2. Find a classical solution of the Euclidean equations of motion (the QCD version of Maxwell equations) which as $t \to -\infty$ reduces to $\underset{\sim}{A} = 0$ and as $t \to +\infty$ reduces to a gauge bubble.

3. Calculate the action $S = \int_{-\infty}^{\infty} L\, dt$ of this classical solution. Then the tunneling amplitude is $\approx e^{-S}$.

The required classical solution was found by Belavin, Polyakov, Schwartz and Tyupkin.[34] There are, in fact, a class of such solutions characterized by

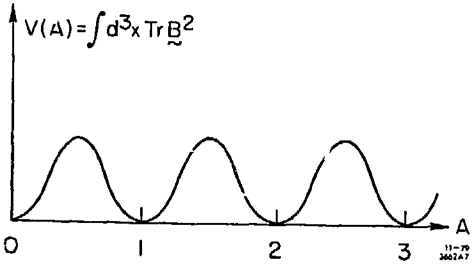


Fig. 6. Effective potential energy in field space as a function of winding-number N.

a. Size. The fields $\underset{\sim}{F}_{\mu\nu}$ are concentrated in a characteristic 4-dimensional volume of size λ, where λ can be chosen at will. For $\rho = \sqrt{|t|^2 + r^2} >> \lambda$ the fields $F_{\mu\nu}$ fall off as ρ^{-2}, are proportional to e^{-1} (recall $|eA\lambda| \sim 1$), are self-dual ($\vec{\underset{\sim}{E}} = \pm\vec{\underset{\sim}{B}}$), and have a rather complicated angular dependence.

b. Orientation in internal space: the $\vec{\tau}$'s in the gauge bubble (and fields $\underset{\sim}{F}_{\mu\nu}$) can be chosen from the 8 λ-matrices in many ways.

c. Location in space-time. The center of the configuration is arbitrary. These classical solutions are the instantons. The gauge-bubbles are <u>not</u> instantons. Instantons are in some sense a dynamical agent responsible for their creation.

Because the field strengths F, from dimensional arguments, are $\sim 1/e\lambda^2$ inside the instanton, and because the action S of QCD is

$$\frac{1}{2}\,\mathrm{Tr}\int d^4x\, \underset{\sim}{F}_{\mu\nu}\, \underset{\sim}{F}^{\mu\nu} \sim \frac{1}{e^2} \int_{\rho = |x| \leq \lambda} d^4x\, \frac{1}{\lambda^4} \sim \frac{1}{e^2} \tag{4.4}$$

the tunneling amplitude is small. The correct numerical calculation gives

$$\text{Tunneling amplitude} \sim e^{-8\pi^2/e^2} = e^{-2\pi/\alpha_s} \tag{4.5}$$

The effect is nonanalytic in α_s and could not be seen in an ordinary perturbation expansion.

Because an instanton can be located anywhere in space-time, we can have gauge-bubbles continuously created and destroyed all over the place and all the time. This is what makes their effects relevant: the coupling of the N-vacuua is a volume, not a surface effect. For example, not only does the energy $\langle H \rangle$ of a θ-vacuum depend upon θ but also the energy-density. The formula is[35,36]

$$\langle H \rangle \sim -V \cos\theta \left(\frac{8\pi^2}{e^2}\right)^6 \int \frac{d\lambda}{\lambda^5}\, e^{-8\pi^2/e^2(\lambda)} \tag{4.6}$$

The predominant exponential of the tunneling amplitude is embellished by various factors. The factor $\cos\theta$ reflects the periodicity of the θ-vacuua; i.e., the conjugate operator N has integer eigenvalues (cf., Equation (4.3)). The factor $(8\pi^2/e^2)^6$ arises because of the large number of degrees of freedom possible for an instanton (scale-size, group orientation, location) and is a measure of the "entropy" of this gas of instantons in space-time. The weight factor λ^{-5} for the integration over scale sizes λ is determined by dimensional analysis. Finally, the value of e^2 to be used should match the scale of the instanton — although already here the rigor of the analysis begins to unwind somewhat.

If we restrict ourselves to the femtouniverse, it is easy to see that the most important instantons are those which fit snugly inside the box. This is because the running-coupling constant Equation (1.4) is

$$\frac{8\pi}{e^2(\lambda)} = \frac{2\pi}{\alpha_s(\lambda)} \cong \frac{2\pi}{\alpha_s(M^2)} + \frac{33 - 2N_f}{6} \log \lambda^2 M^2 \tag{4.7}$$

with M a very large mass scale ($\gg$ 1 GeV) where we normalize α_s ($\ll 1$). The coupling constant α_s increases as the scale λ increases, and although the growth of α_s is only logarithmic, its effect is large because of its presence in the exponential. Thus the tunneling factor

$$e^{-8\pi^2/e^2} \sim e^{2\pi/\alpha_s(M^2)}\,(\lambda M)^{11 - \frac{2}{3}N_f} \tag{4.8}$$

depends like the 8th power of the scale-size and dominates over the $d\lambda\ \lambda^{-5}$ factor. In a femtouniverse the vacuum energy is (disregarding numerical factors $\lesssim 10$) from Equation (4.6) is then found to be

$$E \sim -\frac{1}{V^{1/3}} \cos\theta \left[\frac{2\pi}{\alpha_s(M^2)}\right]^6 \left[1 - \frac{33-2N_f}{12\pi}\, \alpha_s(M^2) \log M^2 V^{2/3}\right]^6 (V^{1/3}M)^{11-\frac{2}{3}N_f} \tag{4.9}$$

and also depends upon the distance scale $V^{1/3}$ to a large power ~8. The onset of instanton-related effects is consequently extremely abrupt; in the perturbative regime an increase of distance scale by a factor 2 increases the instanton effects by a factor ~500. It is also regrettable but true that as the tunneling amplitude gets large, the semi-classical approximation used to compute it breaks down.[37] Thus it is not much of an overstatement to say that whenever an instanton calculation can be believed, the result is physically unimportant. For example, there have been estimates of the QCD modification to the colliding-beam R coming from couplings of the quark pairs to instantons.[38] They find a contribution which behaves as s^{-8}, with the crucial value of $\sqrt{s}$ occurring at about 1 GeV. A similar situation also exists for deep inelastic scattering.

While the instanton does not upset in any way perturbative QCD, it does play an important role in two other areas. One has to do with CP violation in the strong interactions and another has to do with chiral symmetry breaking and the resolution of the "U(1) problem" — the problem of the origin of the large mass of the η and/or η' mesons. We shall return to these questions in Section 7. There is also the possibility that instantons play a significant role in the understanding of the confinement question. This will be considered in Section 8.

5. PURE QCD: GLUONIUM PHENOMENOLOGY

We have argued in Section 2 that pure QCD (i.e., Stage I) is the theory of interactions of massive, color-singlet (and SU(3) flavor-singlet) bound states of gluons, and that it is unlikely that these states disappear when light quarks are introduced. Instead they decay (with not extraordinarily large width) into open channels of ordinary mesons. It is an important test of QCD to find decisive evidence for or against the existence of these states. In order to do that, some theoretical guidance is clearly of use. It is surprising to me, given the degree of hubris extant that QCD is the correct theory of strong interactions, how little discussion[39,40,41] there has been of this problem. We approach the question in three stages: (a) the gluonium spectrum, (b) decay schemes, and (c) production mechanisms.

a. Spectrum: Properties of bound states of two gluons--in particular their mass—can be motivated in several ways. First of all,[41] we can expect that local operators built of products of the gluon fields and having appropriate quantum numbers will create gluonium states, just as local products of quark fields (e.g., $\bar{q}_i\gamma_5 q_j$; $\bar{q}_i\gamma_\mu q_j$, etc.) create meson states when applied to the vacuum. The simplest such operators are $\sum_{A=1}^{8} F^A_{\mu\nu}(x)F^A_{\alpha\beta}(x)$, which should create gluonia of various spins and parities from the vacuum. Consider the spin-zero operator $F^2(x) = \mathrm{Tr}F_{\mu\nu}(x)F^{\mu\nu}(x)$ for simplicity and look at its two-point-function:

$$D^+(x) = \int d^4x\, e^{iq\cdot x}\langle 0|F^2(x)F^2(0)|0\rangle \qquad (5.1)$$

(An object very similar to this measures $\sigma(\nu + \bar{\nu} \to \text{hadrons})$ through single-graviton annihilation).

At short distances (large q^2), QCD perturbation theory allows a computation:

$$D^+(q^2) \approx (\text{const})\ q^4\ \theta(q^2) \qquad q^2 \to \infty \tag{5.2}$$

At some larger distance scale the perturbation expansion breaks down, and the spectrum cuts off among a presumably discrete collection of gluonium states (Figure 7). We have two choices: either the gluonium mass is "reasonable" (1-2 GeV, say) or else the gluonium mass is large and QCD perturbation theory breaks down at a surprisingly large momentum scale (>> 2 GeV). We shall concentrate our attention here on the first possibility. This option is somewhat reinforced by calculations by Jaffe and Johnson in the context of the MIT bag-model.[42] In the bag model, the vacuum in the region occupied by quarks and gluons (the bag) is modified and has, by hypothesis, a different energy density. The kinetic energy of the quarks or gluons in the bag provides a pressure which balances the pressure created by the (true) vacuum at the walls of the bag. Jaffe and Johnson found masses ~1-1.5 GeV, although nowadays they have less confidence in that estimate.[43] Possible s-wave configurations of two and three gluons are exhibited in Table 5-1, along with local operators which create them. The $\phi_{ijk}(J,M)$ are the spin wave-functions which couple the fields into a state of definite J and M. This classification might be called the "naive gluonium model."[40] It is at least as naive as the naive nonrelativistic quark spectroscopy, naive parton model, naive Drell-Yan formula, etc.

There is nothing very definite that one can say regarding the masses--and especially level spacings--of these candidate states (not all of which

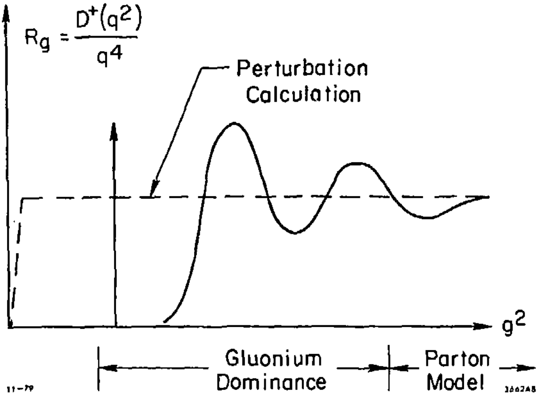


Fig. 7. Schematic of gluonium spectral function, normalized to its high energy behavior.

TABLE 5-1

J^{PC}	Field Operator
0^{++} 2^{++}	$E^A_i E^A_j \phi^{ij}(J,M)$
0^{-+} 1^{-+} 2^{-+}	$E^A_i B^A_j \phi^{ij}(J,M)$
0^{++} 2^{++}	$B^A_i B^A_j \phi^{ij}(J,M)$
0^{-+} 1^{--} 3^{--}	$E^A_i E^B_j E^C_k f^{ABC} \epsilon_{ijk}$ $E^A_i E^B_j E^C_k d^{ABC} \phi_{ijk}(J,M)$
0^{++} 1^{++} 2^{++}	$E^A_i E^B_j B^C_k f^{ABC} \phi_{ijk}(J,M)$
1^{+-} 2^{+-} 3^{+-}	$E^A_i E^B_j B^C_k d^{ABC} \phi_{ijk}(J,M)$
0^{-+} 1^{-+} 2^{-+}	$E^A_i B^B_j B^C_k f^{ABC} \phi_{ijk}(J,M)$
1^{--} 2^{--} 3^{--}	$E^A_i B^B_j B^C_k d^{ABC} \phi_{ijk}(J,M)$
0^{++} 1^{+-} 3^{+-}	$B^A_i B^B_j B^C_k f^{ABC} \epsilon_{ijk}$ $B^A_i B^B_j B^C_k d^{ABC} \phi_{ijk}(J,M)$

need exist at narrow resonances). The question of hyperfine splittings is for example problematical. The splittings might be considerably larger than the typical few hundred MeV for hadrons, because of the larger constituent spin.

Another approach to gluonium properties is to try to connect them up to properties of the Pomeranchuk trajectory,[44] inasmuch as two-gluon, t-channel exchange has vacuum quantum numbers and a "naive" intercept at $J = 1$. Again, the recurrences of that trajectory would tend to lie in the 1-2 GeV mass range. The spins and parities most naturally suggested are 1^{--}, 2^{++}, 3^{--}, etc.

A somewhat related attack is to assume that OZI-violating processes are mediated by gluonia. The systematics of these processes (e.g., $\psi \to$ hadrons, $\psi' \to \psi\pi\pi$) might then provide some insight into masses or couplings of the gluonia. This has been studied rather extensively.[45] Nevertheless, while some progress is made in correlating the data, no clear picture of the gluonium spectra emerges.

b. Decay modes: The decay modes of gluonia are of course sensitive to their mass. The masses are in turn sensitive to, among other things, the hyperfine splittings within a multiplet, which we emphasized might be larger than for systems built from quarks, because of the larger spin of the gluon. QCD tends to put antiparallel spin alignment (in color singlet states) lowest in mass (e.g., $m_\pi < m_\rho$; $m_N < m_\Delta$, etc.) so we might expect the lowest J in a multiplet to lie lowest. We shall, however, remain unprejudiced for awhile, and classify decays of gluonia into two-body channels of the s-wave and p-wave $q\bar{q}$ meson states. This may be a credible guide for gluonium states with mass $\lesssim 1.5$-2 GeV. The final-state mesons we consider are the $^1S_0$ pseudoscalar (P) and $^3S_1$ vector (V) nonets, along with the (C-odd) $^1P_1$ (A') and (C-even) $^3P_{0,1,2}$(S, A, T) nonets. The only two-body channels we allow are those which contain no more than one P-wave meson. The two mesons in the final state must be coupled to an SU(3) singlet. Table 5-2 gives a listing of the meson-states available which are consistent with SU(3) symmetry. (We assume no singlet-octet mixing for $^1S_0$ and $^1P_1$ mesons, and ideal mixing for ^3S and ^3P mesons.) The resultant J^{PC} of the final states are also recorded for $\ell = 0$ and some $\ell = 1$ configurations. These two tables can be used together to search out likely decay modes of the gluonium candidates we have identified. The reader is urged to do this for himself and in that way reach the proper state of enlightenment. This writer cannot claim to have reached that state himself, and offers here only a motley collection of observations:

1. Phase-shift analysis in $\pi\pi$ and $K\bar{K}$ channels should be a good place to look for narrow 0^{++} and 2^{++} low-mass gluonium channels. It was once argued that the S^* is a gluonium state,[42] and the $^3P_0 q\bar{q}$ state lies higher—perhaps to be identified with the very broad ϵ "resonance" at ~1300 MeV.

2. The 0^{-+} $\vec{E}\cdot\vec{B}$ state has the same quantum numbers as the η' and may be mixed with it. (The decay $\eta' \to \gamma\gamma$, however, is well-accounted for in terms of just the $q\bar{q}$ component.) This gluonium channel may be especially tricky[46] because this channel has the same quantum numbers as the divergence of that axial current which is nonvanishing because of instanton effects (cf. Section VII). However, other than η', the only other low-mass open channel is $\pi\delta$.

3. Many of the states have the rather undistinguished 3π, 4π, 5π channels open to them; these offer relatively little hope for a novel search method. Among the (low-spin) gluonia we can put in this category are the 1^{-+}, 2^{-+}, 2^{++}, 0^{--}, 1^{--}, 1^{++}, 1^{+-} and 2^{+-}.

TABLE 5-2

Final Mesons	J^{PC}(s-wave)	J^{PC}(p-wave)	Components
PP	0^{++}	—	$\pi\pi$, $K\bar{K}$, $\eta\eta$, $\eta'\eta'$
PV	1^{+-}	0^{--}, 1^{--}, 2^{--}	$\pi\rho$, KK^*, ($\eta+\eta$, $\omega+\phi$)
PA'	1^{--}	0^{+-}, 1^{+-}, 2^{+-}	πB, KQ,
PS	0^{-+}	1^{++}	πS, $K\kappa$, ($\eta+\eta'$, $S^*+\varepsilon$?)
PA	1^{-+}	0^{++}, 1^{++}, 2^{++}	πA_1, KQ, ($\eta+\eta'$, D)
PT	2^{-+}	1^{++}, 2^{++}, 3^{++}	πA_2, KK^{**}, ($\eta+\eta'$, f+f')
VV	0^{++}, 2^{++}	etc.	$\rho\rho$, K^*K^*, $\omega\omega$, $\phi\phi$
VA'	0^{-+}, 1^{-+}, 2^{-+}	etc.	ρB, K^*B, ?
VS	1^{--}	0^{+-}, 1^{+-}, 2^{+-}	$\rho\delta$, $K^*\kappa$, ($\omega+\phi$, $S^*+\varepsilon$?)
VA	0^{--}, 1^{--}, 2^{--}	etc.	ρA_1, K^*Q, ($\omega+\phi$, D?)
VT	1^{--}, 2^{--}, 3^{--}	etc.	ρA_2, K^*K^{**}, ωf, $\phi f'$

<u>Notation:</u>

P:	$^1S_0$ (0^{-+})	π, K, η, η'
A':	$^1P_1$ (1^{+-})	B, Q, ?, ?
V:	$^3S_1$ (1^{--})	ρ, K^*, ω, ϕ
S:	$^3P_0$ (0^{++})	δ, κ, S^*, ε
A:	$^3P_1$ (1^{++})	A1, Q, ?, D
T:	$^3P_2$ (2^{++})	A2, K^{**}, f, f'

4. The 1^{--} states are candidates for resonance production in e^+e^- storage rings. However, the width into e^+e^- may be very small, and in any event hard to calculate. Furthermore, such states are rather far down the list (3 bound gluons) and may be fairly massive (or for that

matter nonexistent). The activity at DCI and ADONE should evidently be watched with this in mind.

5. Because the gluonia are SU(3) singlets, decays into mesons containing strange quarks may not be suppressed as much as is customary in "old" physics.

6. The tabulation we have made allows decays for all gluonium candidates except for the last of the list--the 3^{+-} state of three magnetic gluons. Decay into p-wave A_1 or A_2 channels are open in that case.

Before moving onto questions of production-mechanisms, we reiterate that this catalogue is not a list of predicted gluonium states. Some, but not necessarily all of the channels listed should support discrete gluonium states. Were QCD to be a true theory, it could provide successful predictions of the spectrum of gluonia comparable to its successful postdictions of the spectrum of hadrons.

c. Production mechanisms: Half the momentum of an energetic proton is not in quark-momentum and is presumably carried by gluon degrees of freedom. It follows that there ought to be a sizable production cross section for gluonia in hadron-hadron collisions. Why hasn't it been found? A possible answer is that there are so many different states of comparable mass and so many decay modes for each that for any given channel σB is very small. Another possible answer is that there hasn't been a sufficiently vigorous search.

Beyond the self-evident searches in hadron collisions and in photoproduction, other methods with more specificity would be especially useful. The physics of onia is an especially attractive possibility.

In the QCD framework, it is manifest that heavy onium only couples to ordinary hadrons via gluons.[47] For example in the process $e^+e^- \to \Psi + X$ or $e^+e^- \to \Upsilon + X$ (Figure 8a), the system X could be rich in gluonium states. A related idea uses diffractive electroproduction (Figure 8b)

$$\mu + \text{nucleus} \to \mu + \Psi + \text{gluonium} + \text{nucleus} \qquad (5.3)$$

Neither of these processes is blessed with a big cross section. Probably the most attractive option is to resonantly produce Ψ (or perhaps Υ) and look at the radiative decay[48]

$$\begin{aligned} &\Psi \to \gamma + X \\ &\Upsilon \to \gamma + X \end{aligned} \qquad (5.4)$$

The principal decay mechanism for Ψ or Υ is supposed to be 3-gluon annihilation (Figure 9a) with subsequent materialization into hadrons. The radiative decay should therefore proceed via a virtual γgg channel (Figure 9b), ideal for formation of gluonium as the state X. Perturbative QCD (not to be trusted much[49] at the Ψ mass-scale) predicts[48]

$$\frac{\Gamma(\Psi \to \gamma gg)}{\Gamma(\Psi \to ggg)} \sim 10\% \qquad (5.5)$$

To good approximation the gluons (or γ-rays) are uniformly distributed in the Dalitz triangle; this immediately leads to an energy spectrum of the γ peaked (Figure 10a) at the high end (higher order radiative corrections round off the upper end; we don't worry about that here). Thus the mass spectrum of the recoiling system X—which hopefully contains the gluonia—is peaked (Figure 10b) at the low energy end. In this region we would expect the gluonium resonances to dominate and turn the spectrum into

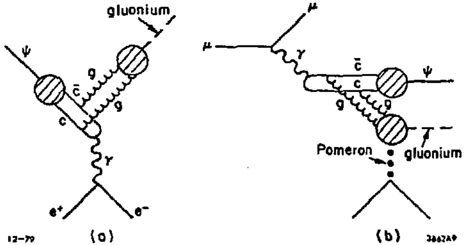


Fig. 8. Mechanisms for gluonium production: (a) $e^+e^- \to \psi g$, and (b) $\mu N \to \mu\psi g N$.

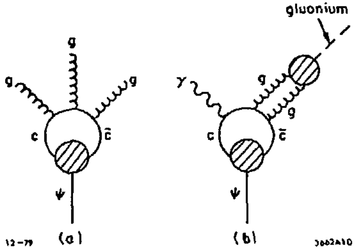


Fig. 9. Gluon (a) and gluonium (b) production in ψ (or ν) decays.

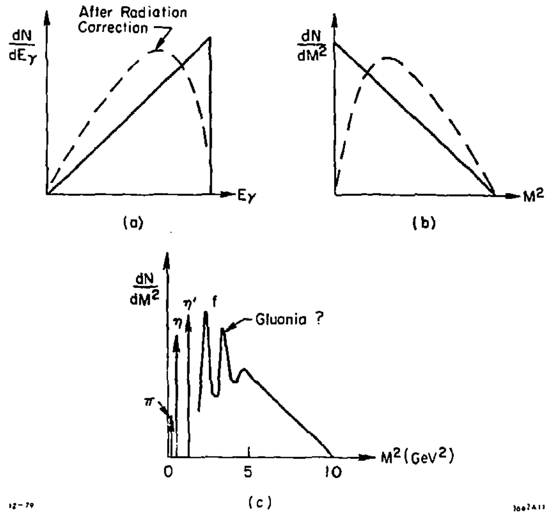


Fig. 10. Spectra in (a) γ-ray energy and (b), (c), recoiling mass for the process $\psi \to \gamma + x$.

something like Figure 10c. There is not yet much known about the inclusive single γ spectrum from Ψ-decays. The decays

$$\Psi \rightarrow \gamma\eta$$
$$\gamma\eta'$$
$$\gamma f$$

have been measured. If they dominate the mass spectrum $M^2 \leq 2\ GeV^2$ and are dual to the perturbative (parton) estimate they should account for ~40% of the decays $\Psi \rightarrow \gamma + X$. The measured exclusive channels actually account for ~1/2% of the Ψ-decays, implying by this bookkeeping a branching-ratio ~1% for all the radiative decays.

A study of these exclusive channels from a QCD viewpoint has been made by Krammer who looks in detail at the angular correlations in the γf channel.[50] Assuming mediation by two transverse gluons (Figure 11) coupled to a $^3P_2$ $q\bar{q}$ system, the helicity state of the $J = 2$ $\pi\pi$ system (it is predominately[51] ±2) is predicted and compared with the data on the $\pi\pi$ angular correlations. The agreement with QCD is good and appears to be nontrivial.[52]

If these exclusive channels are really mediated by two gluons, then there should be other analogous channels which are reasonably well predicted by the quark-model. In particular, a brief look at Table 5-2 lets one accumulate the following list.

$$\begin{array}{ll} \Psi \rightarrow \pi'\gamma & A1\gamma \\ \quad\ \ \pi\gamma & E\gamma \\ \quad\ \ \delta\gamma & A2\gamma \\ \quad\ \ S^*\gamma & f'\gamma \end{array} \qquad (5.6)$$

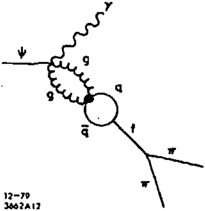


Fig. 11. Mechanism for the decay chain $\psi \to \gamma f \to \gamma\pi\pi$.

But an even more interesting question is what lies beyond. If narrow gluonium states dominate in the region from $M = 1.4$ GeV to $M = 2$ GeV, they should provide ~30% of all radiative decay modes. The γ-ray energies are 1 GeV, and probably badly buried in contamination from π^0 decays. A 2% γ-ray energy resolution corresponds to a resolution in gluonium mass of order 30 MeV. It may be unrealistic to try to resolve any gluonium lines by measurement of the recoil γ-rays alone--even using Crystal Ball--and reduction of background by looking for exclusive gluonium decay channels may be needed. Here one might try for some of those involving neutral decays, e.g. η. But it will be difficult. A scenario appropriate for the Crystal Ball might be

$$\begin{array}{lll} \Psi \to \gamma + \text{gluonium} & & 1/2\% \\ \quad \hookrightarrow \eta + \eta & & 1\% \\ \qquad\quad \hookrightarrow \gamma\gamma \;\; \hookrightarrow \gamma\gamma & & 38\% \times 38\% = 14\% \end{array} \qquad (5.7)$$

The net signal is ~7 events/10^6 decays, even with a rather generous branching ratio assumed for the ηη decay-channel.

If the mass-scale for gluonia is "surprisingly" large, say ~3 GeV, then it is necessary to utilize the Υ decays. For the Υ, the QCD prediction is[48]

$$\frac{\Gamma(\Upsilon \to \gamma + X)}{\Gamma(\Upsilon \to \text{all})} \sim 3\% \qquad (5.8)$$

with again the same spectral shape for the recoil-mass distribution. For a 3 GeV gluonium system recoiling against the γ, the γ-ray will have 90% of the endpoint energy, namely $E_\gamma \sim 4.3$ GeV. A 2% γ-ray energy-resolution now yields a 300 MeV mass-resolution. Again final-state information on the gluonium decay-products probably has to be invoked. Because of the large mass, this will be even more difficult than in the previous case.

Finally, if clean gluon-jets are isolated in Υ-decays, decays of even more massive onia, or in hard processes such as $e^+e^- \to q + \bar{q} + g$, it is plausible that the "leading particle" in such a jet will often be gluonium. Thus in such kinematic regions it is especially appropriate to search for the decay channels enumerated in Table 5-2.

6. INCLUSION OF HEAVY SOURCES: ONIUM PHENOMENOLOGY

In Stage II, we modify the pure QCD of gluonium by introduction of heavy-quark sources. In theory (cf. Section 2) a universal linear confining potential is expected between such heavy sources, provided they have non-vanishing color-triality. But beyond this theoretical issue, Stage II is of interest as a description of the phenomenology of charmonium, bottomonium, toponium, or any other even heavier onia which might eventually be observed.

We start with the idealized situation of very heavy quarks, Q, whose binding can be described accurately in terms of the short distance r^{-1} potential. The static interaction energy of Q and $\bar{Q}$ in a color-singlet state can be read off from the expression for the quark-current source of the color field, Equation (3.16), inasmuch as the subsequent quantization procedure followed accurately (to lowest order in α_s) that of QED:

$$H = \sum_{A=1}^{8} \frac{e^2}{4}\left(\lambda_1^A \cdot \lambda_2^A\right)\frac{1}{4\pi|\vec{r}_1 - \vec{r}_2|} = \frac{e^2\left(\lambda_1 + \lambda_2\right)^2 - \lambda_1^2 - \lambda_2^2}{32\pi|\vec{r}_1 - \vec{r}_2|} = \frac{4}{3}\frac{\alpha_s}{4\pi r} \quad (6.1)$$

where we have used the facts that $(\lambda_1 + \lambda_2)^2 = 0$ in a color-singlet state, and that the 3 × 3 matrix $\sum_A (\lambda_1^A)^2$ is color-invariant, hence a multiple of the unit matrix. Thus, using the normalization in Equation (3.33)

$$\sum_{A=1}^{8} (\lambda_1^A)^2 = \frac{1}{3} \operatorname{Tr} \sum_{A=1}^{8} (\lambda_1^A)^2 = \frac{1}{3} \cdot 8 \cdot 2 = \frac{16}{3} \tag{6.2}$$

Just about everything about this system is analogous to QED. The binding-energy of superheavy onium of mass $M \sim 2M_Q$ is thus

$$E_B = -\left(\frac{4}{3}\right)^2 \frac{\alpha_s^2}{2} \left(\frac{M_Q}{2}\right) = -\frac{2}{9} \alpha_s^2 M \tag{6.3}$$

and the level-spacings are hydrogenic. Thus the 2S and 1S levels are split by an amount

$$\Delta E_{2S-1S} = \frac{1}{6} \alpha_s^2 M \tag{6.4}$$

The Bohr-radius of the onium is

$$r \approx \frac{3}{4} \cdot \frac{1}{\alpha_s} \cdot \frac{4}{M} = \frac{3}{\alpha_s M} \tag{6.5}$$

Thus pure Coulombic "size" of bottomonium (taking $\alpha_s \geq 0.2$) is $\lesssim$0.3f, not especially small.

The value of α_s, just as in QED, depends upon momentum transfer as a consequence of vacuum-polarization effects. The dependence on distance-scale has already been cited in Equation (1.4), and it is appropriate to here discuss this contribution in more detail. A diagrammatic analysis can be carried out;[53] in Coulomb-gauge the vertex and self-energy insertions (Figure 12) are unimportant. All that happens is a cancellation of divergent renormalization constants Z_1 and Z_2 just as in QED. The

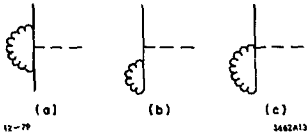


Fig. 12. Vertex and self-energy insertions for the Coulomb energy.

corrections come from the three diagrams in Figure 13. Only the first (Figure 13a) coming from fermion-loops, exists in QED. It tends to increase the charge at short distances because of vacuum polarizability. The calculation is just like QED vacuum polarization except for the group factors. Because the polarization loop is diagonal and independent of the color label A = 1, 2,...,8, we may choose A = 3, which measures the color-isospin of the quarks. Since the coupling of $A^3(x)$ to quark sources is

$$\frac{e}{2}\lambda_3 = \frac{e}{2}\tau_3 = eT_3 \tag{6.6}$$

the modification from QED is just to make the replacement

$$e^3_{QED} \longleftrightarrow e_s^2 \sum_{\substack{\text{colors}\\ \text{of quarks}}} T_3^2 = e_s^2\left[\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 0\right] = \frac{1}{2}e_s^2 \tag{6.7}$$

In QED, we have

$$\frac{e^2}{q^2} \to \frac{e_o^2}{q^2}\left[1 - \varepsilon^{(a)}(q^2)\right] \tag{6.8}$$

$$\varepsilon^{(a)}(q^2) = \frac{\alpha_o}{3\pi}\log\frac{\Lambda^2}{q^2} \tag{6.9}$$

Thus in QCD we get from this contribution

$$\varepsilon^{(a)} = \frac{\alpha_o}{6\pi} N_f \log\frac{\Lambda^2}{q^2} \tag{6.10}$$

or after renormalization and summation of the geometric series

$$\frac{1}{\alpha(q^2)} = \frac{1}{\alpha(m^2)} + \frac{N_f}{6\pi}\log\frac{m^2}{q^2} \tag{6.11}$$

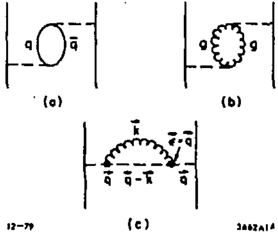


Fig. 13. Vacuum polarization corrections to the Coulomb interaction: (a) quark pairs, (b) transverse gluon pairs, and (c) correction to instantaneous interaction via coupling to transverse gluons.

The second contribution (Figure 13b) from pair production of transverse gluons by the Coulomb field is physically similar, and in fact has the same sign as the first contribution. The coupling of gluons to the Coulomb-field is through the convection-current, with polarization vectors $\vec{\varepsilon}$ and $\vec{\varepsilon}'$ dotted together (Figure 14a)

$$J_0^{\text{scattering}} \sim (E + E')\ \vec{\varepsilon} \cdot \vec{\varepsilon}' \tag{6.12}$$

For the crossed process, there is a minus-sign (Figure 14b)

$$\begin{aligned} J_0^{\text{pair}} &\sim (E - E')\ \vec{\varepsilon} \cdot \vec{\varepsilon}' \\ &\approx \frac{\vec{k} \cdot \vec{q}}{k}\, \vec{\varepsilon} \cdot \vec{\varepsilon}' \end{aligned} \tag{6.13}$$

showing that there is an explicit $|\vec{q}|^2$ factor emerging in the vacuum polarization.

Since the polarization vectors can be safely dotted together and do not enter the guts of the calculation, the structure of the vacuum polarization amplitude is the same as for a pair of spin-zero gluons, which in turn is 1/4 of that for spin-1/2 quanta (given the same charge of the source). (Recall that $R_{e^+e^-} = 1/4$ for a pair of charged spinless point particles.) Hence there are only counting-factors to consider. From this source we therefore get

$$\varepsilon^{(b)} = \Bigl(\sum_{\text{gluons}} T_3^2\Bigr) \cdot \underbrace{2}_{\text{polarization sum}} \cdot \underbrace{\frac{1}{4}}_{\text{spin zero vs. spin 1/2}} \cdot \underbrace{\frac{\alpha_0}{3\pi} \log \frac{\Lambda^2}{q^2}}_{\text{spin 1/2 result}} \tag{6.14}$$

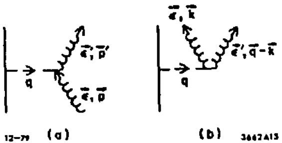


Fig. 14. Scattering (a) and pair-production (b) of transverse gluons by a Coulomb field.

The gluon isospin sum is

$$\sum_{\text{pairs}} T_3^2 = (1)^2 + 2(\tfrac{1}{2})^2 = \tfrac{3}{2} \tag{6.15}$$

and upon again summing the geometric series, we get

$$\frac{1}{\alpha(q^2)} = \frac{1}{\alpha(m^2)} + \left(\frac{N_f}{6\pi} + \frac{1}{4\pi}\right) \log \frac{m^2}{q^2} \tag{6.16}$$

again the "wrong" sign for asymptotic freedom.

The third contribution (Figure 13c) does not exist in QED and comes from the coupling of the Coulomb field of the quarks to the vacuum fluctuations of the (transverse) gauge quanta. It has the important sign-change leading to asymptotic freedom, i.e. the decrease of the effective coupling at short distances. It is in fact much larger than the other two and dominates everything. The physics of the third term is not 100% transparent[54] but is roughly as follows: the Coulomb field created by the quarks itself carries color and therefore interacts with the vacuum fluctuations of the (transverse) color fields. This jitters the Coulomb field and rounds off the short-distance singularity. Whatever the validity of this hand-waving, the calculation follows directly from the formula for "Gauss' Law," expanded out to order e^2. One gets (ignoring $\vec{A} \cdot \vec{E}_T = \vec{A}_T \cdot \vec{E}_T$ term, which is already accounted for) from Equation (3.26)

$$\vec{\nabla} \cdot \vec{E}_L^A - ief^{ABC} \vec{A}_B \cdot \vec{E}_L^C = J_0^A \tag{6.17}$$

Upon inverting this for E_L

$$\vec{E}_L = \frac{1}{\nabla^2} \vec{\nabla} \left(J_0 + ief\, \vec{A} \cdot \vec{\nabla} \frac{1}{\nabla^2} J_0 - e^2 f\, \vec{A} \cdot \vec{\nabla} \frac{1}{\nabla^2} f\, \vec{A} \cdot \vec{\nabla} \frac{1}{\nabla^2} J_0 + \ldots \right) \tag{6.18}$$

Now we can put in plane waves for $\vec{A}$ and, in the last term, contract them together to obtain a vacuum-polarization contribution. After routine Fourier transformation (Figure 13c) the structure of the term is

$$\vec{E}_L^3 = \frac{\vec{q}}{q^2}\left[J_0^3(q) + \sum_{A,B} e^2 (f^{3AB})^2 \sum_{\epsilon} \int \frac{d^3k}{2|k|(2\pi)^3} \frac{\vec{\epsilon}\cdot(\vec{k}-\vec{q})\vec{\epsilon}\cdot\vec{q}}{|\vec{k}-\vec{q}|^2\, q^2} J_0^3(q)\right] \quad (6.19)$$

The Clebsch-Gordan is this time twice as big, because E_T and E_L are not identical degrees of freedom

$$\sum_{A,B} (f^{3AB})^2 = 2(1)^2 + 4(\tfrac{1}{2})^2 = 3 \quad (6.20)$$

Using $\vec{\epsilon}\cdot\vec{k} = 0$, and going to the zero-frequency limit, we get a correction

$$\epsilon^{(c)} \approx -4\pi\alpha_s \cdot 3 \cdot \frac{2}{3}|\vec{q}|^2 \cdot \int_q^{\Lambda} \frac{4\pi k^2 dk}{16\pi^3 k^3} \cdot \frac{1}{|\vec{q}|^2} \approx -\frac{\alpha_s}{\pi} \log \frac{\Lambda^2}{q^2} \quad (6.21)$$

In computing the "electrostatic" energy

$$\frac{1}{2}\int d^3x \sum_A (\vec{E}_L^A)^2 = \frac{1}{2}\int d^3x \; J_0 \frac{1}{\nabla^2} J_0 + \dots \quad (6.22)$$

we get another such term from the other E_L, and a <u>third identical</u> term from contracting the O(e) contributions from each of the E_L factors. Thus the entire contribution from this source to the effect ve charge is <u>three</u> times what we have written down. While not at all obvious, the geometric sum does survive these combinatoric complications[55] and the net change in coupling-constant from this last source is

$$\epsilon^{(c)} = -\frac{3\alpha_o}{\pi} \log \frac{\Lambda^2}{q^2} \quad (6.23)$$

We now can add together the three contributions. Writing in momentum-space, the interaction energy as

$$V(q) = -\frac{16\pi}{3}\frac{\alpha_s(M^2)}{|\vec{q}|^2}\frac{1}{[1+\varepsilon(q)]}$$

we have

$$\varepsilon^{(a)} = -\frac{\alpha_s(M^2)}{6\pi}\sum_{\text{flavors}}\log\frac{M^2}{|q|^2} = -\frac{\alpha_s(M^2)}{6\pi}N_f(q^2)\log\frac{M^2}{q^2} \quad (6.11)$$

$$\varepsilon^{(b)} = -\frac{\alpha_s(M^2)}{4\pi}\log\frac{M^2}{|q|^2} \quad (6.16)$$

$$\varepsilon^{(c)} = +\frac{3\alpha_s(M^2)}{\pi}\log\frac{M^2}{|q|^2} \quad (6.23)$$

where in $\varepsilon^{(a)}$ the sum over fermion flavors goes only over those fermions whose masses are small compared to the momentum scale of interest. Summing up the three contributions leads to the aforementioned formula for the running coupling constant $\alpha(q^2)$:

$$\frac{1}{\alpha(q^2)} = \frac{1}{\alpha(M^2)} + \left(\frac{N_f}{6\pi} + \frac{1}{4\pi} - \frac{3}{\pi}\right)\log\frac{M^2}{q^2} = \frac{1}{\alpha(M^2)} - \frac{(33-2N_f)}{12\pi}\log\frac{M^2}{q^2} \quad (6.24)$$

We see "asymptotic freedom" is controlled by the modifications to "Gauss' Law" due to the vacuum fluctuations of the transverse fields in interaction with the Coulomb field of the sources.

Notice the definition of the coupling constant $\alpha_s(M^2)$ is such that at $|q|^2 = M^2$, the higher-order corrections vanish. Here there is some difference with the traditional QED renormalization program,[56] which defines the physical charge in terms of the force at large distances (as $q^2 \to 0$). In QCD, weak coupling is at short distances only, and there is no unique choice of mass-scale at which to normalize. This leads to arbitrariness in definition of the coupling constant; this freedom can lead to confusion in the interpretation of the theory. One must choose a normalization at some very short-distance scale M (a good choice is $M \sim M_W$) where perturbation-theory is manifestly good. At this scale we may for example choose the definition such that the vacuum-polarization corrections to the static potential vanish. Then, by convention, the coupling-constant at this mass scale is defined to be

$$\alpha_s(M^2) \equiv \frac{12\pi}{(33 - 2N_f(M^2))} \cdot \frac{1}{\log \frac{M^2}{\Lambda^2}} \tag{6.25}$$

The dimensionless number $\alpha_s(M^2)$ is traded in for the parameter Λ which is an energy-scale. The convenience of this definition rests in the fact that, for all values of q^2 for which the perturbative calculation, Equation (6.24), is accurate (and only for such values) one has

$$\alpha_s(q^2) \approx \alpha_s^{(0)}(q^2) = \frac{12\pi}{[33 - 2N_f(q^2)]\log \frac{q^2}{\Lambda^2}} \tag{6.26}$$

Thus the expression is form-invariant; Λ does not explicitly depend upon the choice of the scale at which the coupling constant is defined— at least to this order of approximation.

What is the significance of the parameter Λ? It is evidently a rough measure of the point at which perturbation-theory breaks down. However, when $q^2 \sim \Lambda^2$, there is no justification left for believing the form of Equation (6.26). Also, while the parameter Λ is claimed to be measured reasonably accurately (let us say, between 300 and 700 MeV), this does not mean that we have accurate knowledge of the coupling-constant. For example, Figure 15 shows the spread in $\alpha_s(q^2)$ allowed by the uncertainties in knowledge of Λ. We may conclude that it makes little sense to even consider perturbation-theory below $Q^2 \sim 1\ \mathrm{GeV}^2$, and that Q^2 in excess of 5 to 10 GeV^2 are needed to bring α_s down to a small enough value to trust quantitative results.

Actually the slippery nature of coupling-constant renormalization, when combined with the proven[57] renormalizability of QCD, becomes a powerful tool for summing higher orders of perturbation-theory. This is the "renormalization-group" method. The application to the behavior of the QCD running-coupling constant is just like that of QED. Without <u>any</u> explicit diagrammatic calculation, the Q^2-dependence resulting from summation of leading and next-to-leading logarithms can be determined. In particular one justifies that the leading approximation is the summation of the simple vacuum-polarization bubbles into a geometric series, a result which is not so obvious in QCD as in QED. A further result of the renormalization-group calculations is that if one starts with a small coupling constant $\alpha_s(M^2)$ defined at a short distance scale M and tries to accurately extend it back to large distances, the convergence is rather slow. The formula for α_s becomes, after summation of next-to-leading logs[58]:

$$\alpha_s(q^2) = \alpha_s(M^2) + \sum_{\substack{n\geq m \\ n>1}} C_{nm}\, \alpha_s(M^2)^n \left(\log \frac{M^2}{q^2}\right)^m \tag{6.27}$$

or

$$\alpha_s^{-1}(q^2) = \alpha_s^{-1}(M^2) - \frac{(33-2N_f)}{12\pi} \log \frac{M^2}{q^2} - \frac{3(153-19N_f)}{2\pi(33-2N_f)^2} \log \frac{\alpha_s(q^2)}{\alpha_s(M^2)} + \dots$$

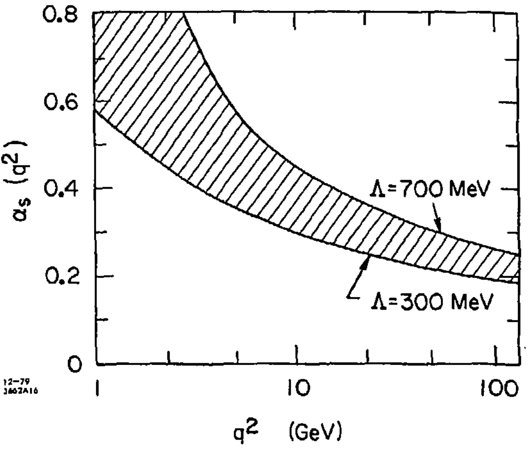


Fig. 15. Q^2 dependence of the running coupling constant for $\Lambda = 500 \pm 200$ MeV.

with $\alpha_s^{(0)}$ the "standard form," Equation (6.26), for the running coupling constant. As q^2 gets small, the correction becomes important when the running coupling-constant gets large. For $q^2 \gg M^2$ the correction-term is never large, despite that fact that

$$\frac{\alpha_s(q^2) - \alpha_s^{(0)}(q^2)}{\alpha_s^{(0)}(q^2)} \sim \frac{\log \log q^2}{\log q^2} \tag{6.28}$$

giving a very slow rate of convergence.

Let us now return to the question of onium bound states. The question of higher-order corrections to (and even existence of) the potential has been extensively studied. For the most part, the development runs in parallel with the treatment of QED bound states. There is one somewhat new feature[59] which appears in high orders, and that is that the interaction-energy has terms $\alpha_s^3 \ln \alpha_s$. These come about from diagrams as in Figure 16 where the intermediate (transverse) gluon is very soft, and produces an infrared divergence. The infrared divergence is cut off by the energy-splitting of the intermediate (color-octet!) $Q\overline{Q}$ configuration from the color-singlet $Q\overline{Q}$ bound state. But beyond this complication there is no signal of confinement other than the behavior of the running coupling-constant.

How does all this compare with the properties of the Ψ and Υ systems ? Most of the spectroscopic facts of those systems rest on a nonrelativistic model of quark motion and are not very specific to QCD. Given the assumption of nonrelativistic motion, Quigg, Rosner, and Thacker[60] have used the level-spacings of the Ψ and Υ states, along with the wave-functions at the origin $|\Psi(0)|^2$ (measured via the leptonic decay widths) to

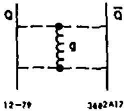


Fig. 16. H-diagram contribution to the Coulomb energy.

reconstruct the form of the potential. They find (Figure 17) a rather good linear potential at moderate distances, with some tendencies toward a 1/r singularity at short distances. The slope of the linear potential (string-tension?) is 1 GeV/f^2 in agreement with what is inferred from the string model of ordinary hadrons. John Richardson[61] has turned things around and taken the QCD form, Equation (6.26) and made a very simple ad hoc change which turns the potential at large distances into a linear potential. He writes, in momentum space

$$V(q) = \frac{4}{3}\,\frac{12\pi}{(33 - 2N_f)q^2 \log\left(1 + \frac{q^2}{\Lambda^2}\right)} + \frac{16\pi\Lambda^2}{(33 - 2N_f)q^4} \tag{6.29}$$

A q^{-4} behavior as $q^2 \to 0$ corresponds to a linearly rising potential. In terms of a string model, the coefficient is related to the string-tension T as follows:

$$T = \frac{1}{2}\, q^2\alpha(q^2)\Big|_{q^2\to 0} \cong \frac{8\pi\Lambda^2}{33 - 6} \approx 0.25 \pm 0.15 \text{ GeV}^2 \sim 1 \text{ GeV/fermi} \tag{6.30}$$

Whether or not this has any fundamental basis, this potential does well in describing the level-structure of charmonium. Other calculations for even heavier onia have been made,[62] as shown in Figure 18. One sees that the pure QCD Coulomb potential is not applicable until the onium mass exceeds ~ 100 GeV.

Does QCD have much to say about fine and hyperfine-structure? Here the situation is not very clear for several reasons. First of all, the detailed origin of the linear component of the potential is theoretically unclear; hence it is not reasonable to expect spin-dependent refinements

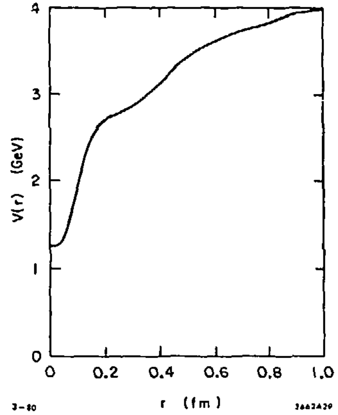


Fig. 17. Typical charmonium potential as reconstructed by Thacker, Quigg, and Rosner (we choose m_c = 1.2 GeV, and E_0 = 3.8 GeV).

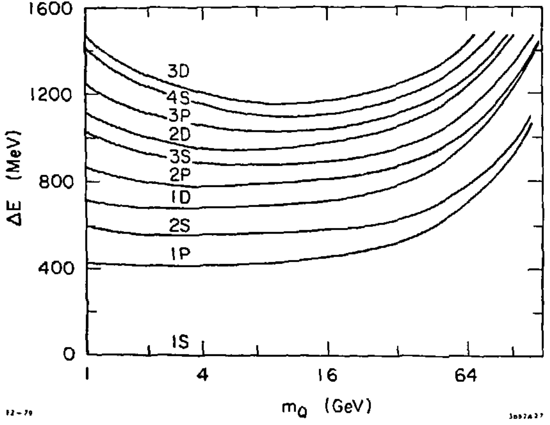


Fig. 18. Onium levels as a function of mass. From Ref. 50.

to it to be under control. Perhaps the most reliable (and also most successful) QCD application is the pattern of hyperfine splittings. These are typically short-distance in nature and operative in s-states, so that one may hope that a perturbative estimate is at least qualitatively adequate. The most interesting result from QCD is that the signs of the splitting come out correctly both for $q\bar{q}$ and qqq systems. Thus for $q\bar{q}$ systems, the relative signs between charge and spin couplings are opposite to QED, because of the $\underset{\sim}{\lambda}_1 \cdot \underset{\sim}{\lambda}_2$ factor, and $^3S_1$ lies above $^1S_0$. This pattern is universal among the quarks; ρ, ω, ϕ lie above π, η, η' and D^* lies above D. Thus Ψ had better lie above η_c, although the magnitude of the splitting is not certain.

For qq systems, the agreement persists. The QCD prediction is that Δ (1236) lies above the nucleon. This can be seen fairly easily; in going from a $^1S_0$ e^+e^- to a $^1S_0$ e^+e^+ in QED, the sign of the spin-dependent force changes sign; hence is attractive. However, in QCD, there is still the $\underset{\sim}{\lambda}_1 \cdot \underset{\sim}{\lambda}_2$ color factor which will be negative in $\underset{\sim}{\bar{3}}$ and positive in $\underset{\sim}{6}$. Thus the energy between two quarks goes as

$$H_{spin} \sim -(\vec{\sigma}_1 \cdot \vec{\sigma}_2)(\underset{\sim}{\lambda}_1 \cdot \underset{\sim}{\lambda}_2) \quad (6.31)$$

Then in the Δ, where each quark pair is in a spin-triplet state with $(\sigma_i \cdot \sigma_j) = +1$

$$\langle\Delta|H_{spin}|\Delta\rangle = -\langle\lambda_1 \cdot \lambda_2 + \lambda_2 \cdot \lambda_3 + \lambda_3 \cdot \lambda_1\rangle = +\frac{3}{2}\underset{\sim}{\lambda}_1^2 \quad (6.32)$$

and there is net repulsion which is optimally strong. Thus N must lie below Δ.

The situation regarding fine-structure splittings is confused. A Breit-type hamiltonian built in analogy to QED predicts quite a large amount of spin-orbit splitting among ordinary hadron multiplets. However, there is relatively little observed. The problem is most acute in the P-wave 70, L = 1 baryon states, where there is remarkably little spin-orbit coupling found. There is some success in organizing the $^3P_1$ states of the charmonium system using semiphenomenological potentials, but the picture is not a completely simple one.[65]

There are many detailed studies of the level splittings from both a QCD and directly phenomenological point of view.[63] The above discussion only scratches the surface. From the present point of view, many of the analyses are not fully relevant because they incorporate the old evidence for the η_c at 2.8 GeV as seen by DASP. A rather thorough study of baryon splittings has been made by Karl and Isgur,[64] who account for a remarkable amount of data on 70, L = 1 and even on 56, L = 2 from a semiphenomenological QCD starting-point.

We next turn to transitions between onium levels. In Stage II, many of the low-lying levels are stable with respect to strong interactions, because the only open channels contain (probably relatively massive) gluonium states. Nevertheless, we may induce transitions (e.g., in the femtouniverse) by slowly varying external gluon fields. Under this circumstance, a multipole-expansion is suggestive. This has received quite a bit of study.[66] Effective local color-singlet operators, e.g., as written in Table 5-1 (gluonium field operators), will have calculable couplings, including transition-couplings, to the onia. This is still a long way from estimating hadronic widths for onium transitions. For that one needs to know the spectrum of gluonia and their decay widths into ordinary hadrons. However, some information on relative widths can be gleaned by analysis of angular-momentum barrier factors present in the various transition channels. For example the $^3P_1$ χ (3510) state is forbidden to decay into two gluons, and operators higher order in α_s need to be invoked to induce the decay. It follows that its hadronic width would be expected to be smaller than for the $^3P_{0,2}$ χ-states. This seems to be the case experimentally.

The annihilation channels, where the Q-value is large compared to the confinement scale, offer a circumstance where perturbative QCD should be applicable (onium can happily annihilate even when confined in the femtouniverse; thus its decay-width should be calculable perturbatively). The width for $\Upsilon \to ggg$ can be stolen from the 1949 (!) QED calculation by Ore and Powell[67] for the 3-photon annihilation of positronium.

$$\Gamma(\Upsilon \to ggg) = \frac{5}{18}\left(\frac{M_\Upsilon}{2m_c}\right) \cdot (\text{WFC}) \cdot \Gamma(^3S_1 \text{ positronium} \to \gamma\gamma\gamma) \cdot \left(\frac{\alpha_s}{\alpha}\right)^3 \qquad (6.33)$$

The factor (WFC) is a wave-function correction originating from the fact that the onium wave function is not purely Coulombic. The numerical factor comes from the color wave function.

There is some ambiguity in what value of $\alpha_s(q^2)$ to use in the above formula. We may contemplate three choices for q^2:

a. $q^2 = m_\Upsilon^2$: This is the time scale over which the annihilation process takes place.

b. $q^2 = 1/9\ m_\Upsilon^2$. This corresponds to a distance scale equal to the Compton wavelength of each gluon.

c. $q^2 = \alpha_s^2 m_\Upsilon^2$. This corresponds roughly to the size of the onium bound state (the initial state should fit into the femtouniverse).

This gives a spread in possible q^2 values from $\sim 3\ GeV^2$ to $\sim 100\ GeV^2$, and a corresponding uncertainty in α_s (from Figure 15) of a factor $\lesssim 4$. Thus the theoretical width for the T is (conservatively) uncertain to a factor $\lesssim 60$. That is, any experimental value within that range could be rationalized as being in agreement with QCD.

Given this situation, a reasonable way to proceed is to fit the value of α_s to the observed T width of ~ 100 keV. In this way one arrives at a value

$$\alpha_s(T \rightarrow 3g) \sim 0.2 \tag{6.34}$$

This implies, from Figure 15, a Q^2 value $\sim 10\ GeV^2$, which is quite acceptable. We can also infer from simple dimensional considerations that the Q^2-value for Ψ-decay should be about 10 times smaller. Hence, were perturbative QCD ideas to be acceptable for the $\Psi \rightarrow ggg$ decay, we should have

$$\frac{\Gamma(\Psi \rightarrow 3g)}{\Gamma(\Psi \rightarrow \mu\mu)} \approx \frac{\Gamma(T \rightarrow 3g)}{4\Gamma(T \rightarrow \mu\mu)} \cdot \left[\frac{\alpha_s(Q_\Psi^2)}{\alpha_s(Q_T^2)}\right]^3 \tag{6.35}$$

The leptonic branching ratio for the T is still poorly known. But putting in numbers for these widths, one finds

$$\frac{\alpha_s(Q_\Psi^2)}{\alpha_s(Q_T^2)} \approx 1.0 \pm 0.2 \tag{6.36}$$

Even qualitatively this is not what is expected. From Figure 15, there should be a decrease in α_s of 50% to a factor 2 in going from the Ψ system to the T system.

The easiest conclusion to draw from this is that the mass of the Υ is much too low to allow quantitative applications of perturbative QCD. This should be no surprise inasmuch as each gluon carries an average of ~1 GeV of momentum. In addition, recently calculated radiative corrections[49] turn out to be remarkably large, even for the Υ system.

A perhaps better test of perturbative QCD is to look for the 3-jet final state in Υ-decay, and also the radiative process

$$\Upsilon \rightarrow gg\gamma \tag{6.37}$$

The branching ratio for this process is estimated to be ~3%. These questions are discussed in detail by Stan Brodsky.[2]

7. INCLUSION OF LIGHT FERMIONS

Thus far we have not faced directly the real-life situation present when the light fermions are included in the theory. We have already indicated the basic changes which occur. The most important is the appearance of the ordinary hadrons. Qualitatively there is not much of an additional conceptual problem. The problem is a quantitative one: can we understand the classification and the mass-spectrum--especially of the excited resonant states--and their couplings to each other? This question is a very big and difficult one to handle theoretically, and will not be directly attacked here. It would seem necessary to have a rather firm control of the mechanism of confinement before one had firm control of such questions.

Along with the introduction of the light quarks goes the disappearance of the string, which can break due to (Heisenberg-Euler) pair creation.[11] Actually, the lifetime of a piece of string is somewhat uncertain. If mesons of large J can be considered as quarks rotating about each other and connected by a piece of string, then their lifetimes ($\Gamma \approx$ 10-100 MeV) give some measure of the string lifetime. The relatively large widths of charmonium states which lie above $D\bar{D}$ threshold are another indicator. But in any case the linear potential becomes complex (absorptive) as r increases, and we should not expect that the concept continues to make much sense when, say, Re V(r) >> 1 GeV. This is because Im V(r) grows with r (linearly?, quadratically?) as well as Re V(r).

Introduction of light quarks will also modify the properties of the gluonia discussed in Section 4. Gluonia may mix with the ordinary mesons, and will also decay into meson channels. However, as already mentioned, these are not expected to be large effects. Eloquent argumentation for this has been given by Witten[58] on the basis of the 1/N expansion, where N is the N of SU(N). One hopes N = 3 may be "large"; as $N \to \infty$ one can argue that gluonia decouple from quarks (as well as from each other!). While probably most theorists expect gluonia to survive as distinguishable states even in the presence of mixing with ordinary mesons, a group from ITEP[69] has challenged this view. Their argument is based on the QCD sum-rules for the charmonium system and a generalization to the gluonia. The QCD sum rules are interesting in their own right. Weighted integrals of the colliding-beam cross section are related to matrix elements of various local operators via the Wilson operator-product expansion. The structure is,

$$\int_0^\infty \frac{ds}{s^N} R(s) = \begin{array}{l}\text{sums of vacuum matrix}\\ \text{elements of local operators.}\end{array} \qquad (7.1)$$

The right-hand side is calculable in terms of short-distance properties of QCD--the more so the smaller the value of N. With some mathematical trickery they can sum things up and also write

$$\int_0^\infty ds\, e^{-s/M^2}\, R(s) = \text{another sum of vacuum matrix elements of local operators.} \qquad (7.2)$$

When they consider separately the R_i associated with currents of u, d and s and c' ... M appropriately, the sum is saturated by ρ, ω, and ϕ. The result is a successful calculation of masses and widths of these resonances in terms of short-distance parameters of QCD!! They then turn[70] to processes such as discussed in the beginning of Chapter V, created by local color-singlet gauge-invariant operators such as $\mathrm{Tr}(\underset{\sim}{E}^2 - \underset{\sim}{B}^2)$ or $\mathrm{Tr}\ \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}}$. The corresponding sum rules are constructed analogously to Equation (7.2). The right-hand side can be evaluated from the information already obtained from the previous sum rules. They find that the left-hand side can be saturated by known mesonic states such as ε or η' and there is no need for additional distinct gluonium states. This does not of course prove such states do not exist; however, it does undermine the notion that inclusion of light fermions does not significantly modify the properties of the gluonia, since without the fermions present there would be no alternative but to saturate the QCD sum rules with gluonium states.

In the context of QCD, some of the most interesting aspects of the presence of light quarks have to do with questions of symmetry--from the approximate chiral flavor symmetry SU(3) ⊗ SU(3) of current algebra to the discrete symmetries of C, P and T. The issues involved are quite subtle, and we shall concentrate our attention in this section on them. The first issues have to do with chiral symmetry. Introduction of the fermions u, d, s,... adds to the QED Hamiltonian a term

$$H_{strong} = H_{Pure\ QCD} + \sum_i \psi_i^\dagger \{\vec{\alpha} \cdot (\vec{p} - \vec{e}A)\delta_{ij} + \beta m_{ij}\} \psi_j \qquad (7.3)$$

where the mass-matrix is usually presumed to originate from the Higgs-mechanism of the electroweak interaction. Because the Higgs sector is not expected to respect internal symmetries, it follows that m_{ij} need not be diagonal. In fact the mass matrix may even contain γ_5. Nevertheless, because of the presumed diagonal and flavor-independent nature of the kinetic-energy and $\vec{j} \cdot \vec{A}$ terms, one can redefine the fields ψ_i in such a way that the mass term in H_{strong} is diagonal and γ_5 free. This is an important feature: it implies (or would seem to imply) that, despite P, C, and/or T violating effects in the Higgs sector (or whatever) responsible for quark mass generation, such symmetry violation will not find its way into the strong interactions. This desirable result will be tempered somewhat in what follows: the Tr $\vec{E} \cdot \vec{B}$ surface-term does induce CP violating effects in strong interactions via instantons. We shall come to this later on.

In the absence of the mass terms (a reasonable approximation for u, d, s), the Hamiltonian is invariant under independent rotations (in flavor space) of the left- and right-handed fermion fields

$$\psi_{iL}(x) \to U_{ij}^{(L)} \psi_{jL}(x)$$
$$\psi_{iR}(x) \to U_{ij}^{(R)} \psi_{jR}(x) \qquad (7.4)$$

leading to a chiral $U(3)_L \otimes U(3)_R$ symmetry. This is broken only by the "small" mass terms. Eighteen vector and axial currents can be constructed which in the limit $m_u, m_d, m_s \to 0$ are formally conserved by the equations of motion. However, one of them, the 9th flavor-singlet axial current, is not conserved because the short-distance ultraviolet divergences of the theory do not allow it. This phenomenon is the triangle-anomaly: the divergent graphs of Figure 19 do not allow the shifts of origin in momentum space required to obtain the formal vanishing of the divergence of this axial current.[71] Defining

$$J_{\mu 5} = \sum_{i=u,d,\ldots} \bar{q}_i \gamma_5 \gamma_\mu q_i \tag{7.5}$$

the result is

$$\frac{\partial J_{\mu 5}}{\partial x_\mu} = \frac{e_s^2 n_f}{4\pi^2} \mathrm{Tr}\, \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}} + \begin{array}{l}\text{terms vanishing}\\ \text{as } m_i \to 0\end{array} \tag{7.6}$$

where n_f is the relevant number of flavors (3 for u, d, s). The $\mathrm{Tr}\, \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}}$ factor should already be familiar from Section 4 on instantons and from Appendix B. There we found that the term $\mathrm{Tr}\, \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}}$ itself is a total divergence

$$\frac{e^2}{4\pi^2} \mathrm{Tr}\, \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}} = \frac{\partial K_\mu}{\partial x_\mu} \tag{7.7}$$

where the current

$$K_\mu = \frac{e^2 \varepsilon_{\mu\alpha\beta\gamma}}{8\pi^2} \mathrm{Tr}\left[\underset{\sim}{A}_\alpha \partial_\beta \underset{\sim}{A}_\gamma - \frac{2ie}{3} \underset{\sim}{A}_\alpha \underset{\sim}{A}_\beta \underset{\sim}{A}_\gamma\right] \tag{7.8}$$

is gauge-dependent. The operator $N = \int K_0 d^3x$ measured the topological quantum number used to classify the QCD vacuua. It might appear that the

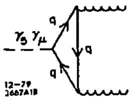


Fig. 19. The triangle anomaly.

U(1) chiral symmetry could be salvaged by considering the summed current $\mathscr{J}_\mu = J_\mu + K_\mu$ which <u>is</u> conserved. The conserved charge would be in this case

$$\int \mathscr{J}_0 d^3x \equiv Q_{tot} = Q_5 + 2n_f N = \sum_{\text{flavors}} (N_{Li} - N_{Ri}) + 2n_f N \qquad (7.9)$$

where N counts the number of gauge-bubbles present in the QCD vacuum. However, even in the absence of the light quarks, we have learned that we cannot characterize the vacuum by the quantum number N. Because of gauge invariance and vacuum tunnelling via instantons, it is its conjugate variable θ that labels the vacuua. So also it will be for the chiral charge Q_5; the variable φ conjugate to Q_5, which is a phase, is used to characterize the chiral structure of the vacuum. This is the same as the situation for spontaneous chiral symmetry breakdown in the more conventional context: the vacuum is <u>not</u> an eigenstate of the chiral charge.

We have not yet motivated why this chiral symmetry-breaking must occur. This has to do with the instanton phenomenon. The different N-vacuua become coupled because of the existence of a non-vanishing quantum tunnelling amplitude (the instanton). With fermions present, we must reanalyze the tunnelling process and keep track of how the quark-states are affected by, say, the creation of an instanton-induced gauge-bubble.[72] Consider a femtouniverse where the quark coupling to transverse gluons may be considered perturbatively (i.e., neglected). If we are in an N = 0 sector the quark states are simple plane waves of definite chirality (with the approximation $m_{u,d,s} = 0$) and quantized momenta.

Is there anything different about the solutions in the presence of a gauge-bubble? The answer is no: only a color-dependent gauge-phase need be multiplied to the quark wave-functions. The energy-eigenvalues (which are gauge invariant) remain unchanged at $\pm \frac{\pi}{V^{1/3}}$, $\pm \frac{\pi\sqrt{2}}{V^{1/3}}$, etc. [Note: for convenience we choose to quantize with antiperiodic boundary conditions $\Psi(\frac{L}{2}) = -\Psi(-\frac{L}{2})$ to avoid eigenstates of zero energy.] However, we may ask what goes on as the tunnelling occurs. We shall analyze this problem in first-quantization. That is, we treat each fermion state individually and only introduce the filled Dirac-sea of negative-energy eigenstates at a later point. Because the tunnelling is semi-classical the coordinates of a given quark must be deformed along with the gauge-field coordinates. In the intermediate configurations between $N = 0$ and $N = 1$ there are no pure gauges; the quark finds itself in real color-electric and color-magnetic fields. Thus the levels will shift--and if the tunnelling rate is slow, they will shift adiabatically:[73,74] only degenerate fermion states mix. However, as the tunnelling becomes complete, one returns again to a pure gauge configuration--and therefore the same set of energy levels as in the beginning. But the important and crucial feature of this process is that <u>the matching of final levels to initial levels is nontrivial</u>. Some of the initial levels move upward; others downward and end up, at $N = 1$, in different states (c.f. Figure 20). What does this mean? It means that, when one second-quantizes and--according to Dirac hole theory--fills all negative-energy states, the tunnelling phenomenon takes one from the $N = 0$ vacuum of filled fermion states to an $N = 1$ state <u>which is not necessarily the $N = 1$ vacuum</u>. In fact, if any of the fermion energy levels does cross zero, then the $N = 1$ state

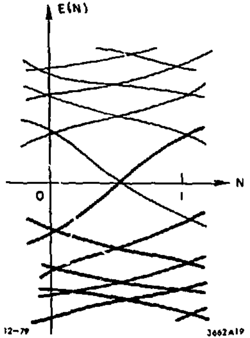


Fig. 20. Schematic picture of shifts in fermion energies as a function of winding-number N.

which is reached will not be the naive N = 1 vacuum state. This is actually what happens. In tunnelling from N = 0 to N = 1, one negative-helicity state (for a given flavor) of originally positive energy dives into the negative energy-sea, while one positive-helicity state emerges from the negative-energy sea and joins with the positive energy states.[75] Thus, the net effect is that the instanton-induced transition from the N = 0 vacuum to the N = 1 vacuum is suppressed; instead the transition simultaneously creates a pair of each flavor of "massless" fermion, since the levels of all flavors are shifted together. In our case, that means the transition is accompanied by creation of three quarks and three antiquarks

$$(\mathrm{vac})_{N=0} \rightleftarrows (\mathrm{vac})_{N=1} + u\bar{u}\; d\bar{d}\; s\bar{s} \qquad (7.10)$$

Remarkably enough,* this process respects the conservation law implied by Equations (7.6) and (7.7):

$$\Delta\left[\sum_i (N_{Ri} - N_{Li})\right] = 2n_f\ \Delta N \qquad (7.11)$$

---

*When applied to the weak-interaction gauge theory,[72] this phenomenon is even more spectacular. In a femtouniverse with dimension small compared to m_W^{-1}, one expects to have an essentially unbroken $SU(2)_L$ non-Abelian gauge theory. Thus the instanton creates one left-handed fermion from each weak doublet, consistent with charge-conservation. For three generations this means:

$$\begin{aligned}(\mathrm{vac})_N \rightleftarrows (\mathrm{vac})_{N+1} &+ \begin{pmatrix}\nu_e\\ e\end{pmatrix} + \begin{pmatrix}u\\ d\end{pmatrix}_1 + \begin{pmatrix}u\\ d\end{pmatrix}_2 + \begin{pmatrix}u\\ d\end{pmatrix}_3\\ &+ \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix} + \begin{pmatrix}c\\ s\end{pmatrix}_1 + \begin{pmatrix}c\\ s\end{pmatrix}_2 + \begin{pmatrix}c\\ s\end{pmatrix}_3\\ &+ \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix} + \begin{pmatrix}t\\ b\end{pmatrix}_1 + \begin{pmatrix}t\\ b\end{pmatrix}_2 + \begin{pmatrix}t\\ b\end{pmatrix}_3\end{aligned}$$

This cannot be an accident; the triangle anomaly must sense at the high momenta the imbalance of fermion levels induced by gauge-bubbles and nonvanishing values of $\int d^4x\, \vec{E}\cdot\vec{B}$. One may recall that even in QED the definition of the electrical-current operator requires a careful symmetrization between positive and negative energies. The monotonic shift in levels of given chirality induced by an $\vec{E}\cdot\vec{B}$ term will be felt not only in terms of levels which cross zero energy but also (in the presence of a high-momentum cutoff) of an induced asymmetry at the highest momenta and consequent multilation of the structure of the current operator.[75]

What now are the stationary states? The tunnelling effective Hamiltonian analogous to Equation (4.6) now has an extra factor to account for the quark pair-creation we have found. Schematically

$$H' \sim \sum_N C\,\delta_{N,N+1} \int \frac{d\lambda}{\lambda^5}\,(\bar{u}u)(\bar{d}d)(\bar{s}s)\,e^{-\frac{2\pi}{\alpha_s(\lambda)}}\left(\frac{2\pi}{\alpha_s}\right)^6 \cdots \qquad (7.12)$$

When H' is evaluated in a θ-vacuum, the N-dependent factor evidently becomes a phase

---

For example, this means there exist $|\Delta B| = |\Delta L| = 3$ virtual transitions

$$(\mathrm{vac})_N \longleftrightarrow (\mathrm{vac})_{N+1} + \nu_e + n + \mu^- + (ccs) + \tau^- + (ttb)$$

Thus a pure 3rd generation baryon could decay into a pure second-generation antibaryon, an antineutron, and 3 antileptons. Regrettably (or perhaps fortunately) the amplitude for this process is extremely small, of order

$$e^{-\frac{2\pi}{\alpha}\sin^2\theta_W} \sim 10^{-60}$$

$$<\theta|H'|\theta> \sim (\text{const})e^{i\theta}\int\frac{d\lambda}{\lambda^5}(\bar{u}u)(\bar{d}d)(\bar{s}s)e^{-\frac{2\pi}{\alpha_s}}\left(\frac{2\pi}{\alpha_s}\right)^6 \ldots + \text{h.c.} \quad (7.13)$$

For diagonal matrix-elements it is necessary to contract the fermion-fields $\bar{u}u \to \langle\bar{u}u\rangle$, etc.; and such contractions vanish if the masses of the quarks are zero. Hence all instanton-related effects will now be proportional to the product $m_u m_d m_s$, where these are the current-algebra masses (a reasonable estimate is $m_u \sim 4$ MeV, $m_d \sim 7$ MeV, $m_s \sim 150$ MeV). Furthermore the phase transformations on the quark fields necessary to put the mass-matrix $\mathcal{M}$ into standard form $\sum_i \bar{q}_i|m_i|q_i$ will also leave a phase ϕ in the expression (Equation (7.13)) for the tunnelling energy in a θ-vacuum. This phase is easily to be found to be

$$\phi = \arg\det\mathcal{M} \quad (7.14)$$

and hence the effective angle relevant to observable effects, such as in Equation (7.13) will be

$$\bar{\theta} = \theta - \phi = \theta - \arg\det\mathcal{M} \quad (7.15)$$

Thus even if strong-interactions have for some reason a boundary condition $\theta = 0$, after inclusion of mass-generation (via Higgs-electroweak mechanisms) the effective angle $\bar{\theta}$ cannot be expected to remain zero.

Thus far we have found that the $U_{3L} \otimes U_{3R}$ symmetry has been broken by instanton effects to an $SU(3)_L \otimes SU(3)_R \otimes U(1)_V$ symmetry. What about spontaneous chiral symmetry breaking? The flavor-octet of axial currents $J_\mu^{\,i} = \bar{q}\gamma_\mu\gamma_5\lambda_i q$ is not plagued with instanton couplings or triangle anomalies, and one may attack the problem in more conventional ways. Even in QED, it has been argued that chiral symmetry can be spontaneously broken,[76] i.e., lepton mass can be generated self-consistently by the mechanism illustrated in Figure 21. The same mechanism with gluon replacing photon might apply in QCD. A different line of argument has been advanced by Callan, Dashen, and Gross,[77] who exploit the instanton tunnelling and propose a mechanism as shown in Figure 22. The effective instanton-induced six-fermion interaction in Equation (7.13) is used to trigger the spontaneous symmetry breakdown. But this is difficult to make quantitative. In addition, recent explicit calculations in strong-coupling lattice gauge theories[78] are also supportive that spontaneous symmetry breaking of SU(3) chiral symmetry will occur in QCD. In all cases much more needs to be done before one can be fully convinced that QCD does imply the spontaneously broken chiral SU(3) symmetry which underlies the successful current-algebra and PCAG phenomenology of ordinary hadrons.

Finally, it is necessary to inquire into the implications of a nonvanishing value of $\overline{\theta}$ for the questions of CP violation. Evidently a term in the Lagrangian

$$L' = \frac{\theta}{24\pi^2} \mathrm{Tr}\, \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}} \tag{7.16}$$

which serves to fix the value of vacuum-θ is P-odd and C-even; hence T-odd. While formally a total divergence, we have had ample evidence that such a term produces nontrivial effects. The greatest threat lies in the experimental limit of 10^{-24} e-cm on the neutron electron dipole moment, implying a value of $|\overline{\theta}|$ of $\lesssim 10^{-8}$. It _is_ necessary that $\overline{\theta}$ be very small.[79] There are various points of view possible:

1. θ is an arbitrary parameter in the most general renormalizable QCD Lagrangian and, since it is subject to divergent renormalizations from higher order loops, must be considered not implausibly small.

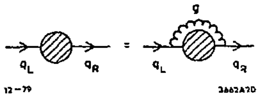


Fig. 21. Possible mechanism for spontaneous chiral symmetry breaking.

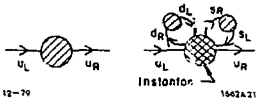


Fig. 22. Possible instanton mechanism for spontaneous chiral symmetry breaking.

2. For strong interactions one simply imposes the condition $\theta = 0$ by hand as a symmetry criterion: T-invariance of strong interactions is demanded from the start. This does not solve the problem completely; weak interactions and Higgs-couplings can reintroduce a non-vanishing θ by radiative effects. Furthermore, these effects can again be divergent, so one can argue they are not small. On the other hand it is not necessarily so that the infinity gives a large value of θ. For example, in the popular SU(5) grand-unified model, the "infinite" CP violating effects only leak into the strong interactions in 14th order. The nominal order of magnitude is[80]

$$\theta << \left(\frac{\alpha}{\pi}\right)^7 \log \infty \tag{7.17}$$

This size is safely "small" ($\frac{\alpha}{\pi} \log \infty \lesssim 1$ for any reasonable value of ∞).

3. We saw that in the limit of vanishing quark mass, the vacuum tunnelling (and therefore CP violating effect) was suppressed, and that in the presence of quark mass the tunnelling amplitude is multiplied by a factor proportional to $m_u m_d m_s$ or better

$$\left(\frac{1}{m_u} + \frac{1}{m_d} + \frac{1}{m_s}\right)^{-1} \approx \frac{m_u m_d}{m_u + m_d} \tag{7.18}$$

Thus, were the "bare" current-algebra mass of at least one quark to be zero, we could avoid the problem.[81] The best candidate is the up-quark, whose mass is estimated, from current algebra considerations, to be ~ 4 MeV. However, zero mass seems to go against successful current-algebra and SU(3) calculations.

4. The mechanism of quark-mass generation can be modified. Peccei and Quinn[82] found that by allowing an additional U(1) symmetry in the Higgs-sector, which was then spontaneously broken, that the CP violating phase could be sloughed off into that U(1) phase of the Higgs sector. However, Weinberg[83] and Wilczek[84] then showed that there should be an almost massless 0^- Goldstone-boson (the axion) with mass $\lesssim$ 10-100 keV. The couplings of this object to matter are sufficiently strong that it should probably have been seen.[85]

Among these options, probably the least unpalatable is option (2): θ is put to zero by hand in the strong QCD theory, with the weak-interaction and Higgs contributions required to be a small perturbation.

We have in this section not provided much of any idea how these effects are calculated quantitatively. The most appropriate and powerful technique utilizes the Feynman path-integral,[86] but even that method is technically quite difficult. Considerable uncertainty remains in the magnitude of these CP violating and other instanton-induced effects.

8. IDEAS ABOUT CONFINEMENT

A great deal of effort has gone into trying to understand the confinement problem, and considerable insight has been attained. Nevertheless, there is no general agreement that the question is understood. Here we shall only mention in the most superficial way some of the approaches:

a. Lattice QCD: One approximates the continuous system with a single cubic lattice, with quarks living on sites and gauge field living on links.[87] The basic element is a line integral

$$\underset{\sim}{U}_{AB} = :e^{i\int_A^B \vec{A}\cdot d\vec{x}}: \tag{8.1}$$

connecting neighboring sites A and B, and the action is a sum of contributions of the 4U's at a time taken around an elementary face (plaquette) of the lattice. This theory is simplest in the strong coupling limit. There Wilson showed[87] that there exists a linear potential between heavy sources—even for lattice QED. The field chooses the shortest path between sources, and any fluctuation costs extra powers of g^2 in energy.

The central problems are to connect the strong-coupling limit to a weak coupling theory at short distances, and to demonstrate a distinction in that limit between the abelian QED and nonabelian QCD theories. Ideally one should be able to compute the string-tension in terms of the perturbation-theory parameter Λ which controls the value of the running coupling constant α_s at short distances.

Important progress has been recently reported. Kogut, Pearson, and Shigemitsu[88] calculate the strong-coupling expansion to several orders and use Pade approximation to extrapolate toward weak coupling. Creutz,[89] using a technique originated by Wilson, reduces the problem to an equivalent statistical-mechanical calculation of a free-energy, which he does via Monte-Carlo techniques on a computer. Both groups, as well as Wilson, find evidence for an abrupt transition from weak to strong-coupling at a critical distance-scale, and at a relatively small value of $\alpha_s \sim 0.1$.

b. MIT Bag: The MIT bag model[90] was originally formulated[91] at a level more phenomenological than QCD. It views the vacuum as a complicated medium, and a hadron a "hole" in the vacuum which is simpler, at least as seen by quarks and gluons. There is an energy cost in creating such a hole; also the vacuum pressure on the hole is compensated by the pressure of the quarks and/or gluons in the interior, leading to a stable hadron.

This picture has enjoyed reasonable phenomenological success. The QCD interest lies in making a connection of this picture to the QCD Hamiltonian. The ideas, not necessarily exclusive, include:

1) Princeton program. Callan, Dashen, and Gross[92] argue that the exterior vacuum (viewed 4-dimensionally) is a dense plasma of instanton events. However they argue that in the presence of color electric fields (i.e., in bag interiors), instanton effects are suppressed and one has a relatively dilute gas. Their calculation of relative properties of these two phases rests heavily upon the use of analogies to statistical mechanics and on taking account of couplings between instantons.

2) Analogy with Meissner Effect: Electric-Magnetic Duality. A monopole in a superconductor undergoes confinement in the way envisaged for QCD: a quantized vortex line connects a monopole-antimonopole pair, providing a linear potential.[93] To exploit this analogy in QCD, one needs to reverse the role of electric and magnetic field; the QCD string contains electric flux, not magnetic. There does exist some $E \leftrightarrow B$ duality in QCD.[94] But this program is evidently a difficult one and at present is not complete.

c. Schwinger-Dyson Equations: Another attack uses the Schwinger-Dyson equations, which sum up the Feynman-diagram expansion.[95] The goal is typically to find a nonperturbative solution consistent with a q^{-4} behavior in the gluon propagator, signalling a confining potential. Problems include how to truncate the infinite set of coupled nonlinear equations for the Green's functions, how to enforce gauge-invariance, and how to interpret the result (the Green's function of the gluon is itself gauge-dependent). On top of this is the straight technical problem, even after brutal truncation, of solving complicated sets of nonlinear integral equations.

d. QCD Strings: This approach treats a piece of bare string, possibly closed, as the basic degree of freedom instead of point quanta. Thus the field degrees of freedom are: $:e^{i\int_A^B \vec{A}\cdot d\vec{s}}:$. The problem is to establish equations of motion and a consistent quantum-mechanical formalism, and then determine properties of dressed strings.[96]

All these, and others not mentioned, are the subject matter of courageous and difficult research. There is optimism that the problem can be understood; perhaps one of these approaches will provide some answers.

9. ALTERNATIVES TO QCD

It is becoming hard nowadays to find serious work on strong interactions which does not start with QCD. This is less a consequence of overwhelming evidence in favor of QCD as it is a consequence of a lack of serious contenders. In my mind the two strongest contenders are the string-model and the Pati-Salam scheme. In the case of the string model, one simply asserts (without any backup from field theory, etc.) that quarks are tied together by strings of unspecified structure and basic origin. It has a close relationship with dual models and the topological expansion of the S-matrix. A major difference with QCD lies in the absence of the gluon degrees of freedom. The recent PETRA data does not encourage this point of view.

The Pati-Salam scheme, based on Han-Nambu integer-charged quarks, has undergone a considerable degree of evolution.[97] The apparent discrepancy with the fractional charge measured in electroproduction can be avoided if the theory is gauged. However, it appears difficult if not impossible to avoid a low mass ($\lesssim 2$ GeV) boson $\tilde{U}$ which mixes with the photon and which has a large leptonic width.[98] This seems to be ruled out on experimental grounds. Also, the τ lepton and b quark do not fit very comfortably into the scheme.

Why the difficulty in finding alternatives? It is simply the short but restrictive list of reasons listed in the introduction, plus the problem of incorporation of the parton-model picture at short distances (solved in QCD by the asymptotic freedom property).

QCD does pass the test on these issues. It possesses as well the distinguished pedigree of being a local gauge theory like QED, which puts the color degree of freedom to work dynamically. It makes it no surprise that it is so widely accepted. Nevertheless QCD does need much better experimental support for it to be truly confirmed. For me, the most reliable tests are those which can be imagined to be carried out in a femtouniverse. These include (1) measurement of the e^+e^- total cross section to an accuracy sufficient to see the radiative correction and (2) measurement of α_s via $e^+e^- \to q\bar{q}g$, with the final partons in a highly non-collinear final state (i.e. $\sim 120°$ away from their neighbors). One will then want to check the gluon spin by measurement of angular correlations in this or other processes.

10. CONCLUSIONS

QCD is a theory of strong interactions which has a starting point as fundamental as QED. While there are major mathematical difficulties in even a successful formulation of QCD as a quantum field theory, these difficulties are ameliorated by quantizing the theory in a tiny box (femto-universe) with a small coupling constant. There do exist processes which we observe and which in principle fit into the femtouniverse. These processes can be calculated perturbatively; in the femtouniverse QCD converges as well as QED.

However, as the box size grows there appear at least six crises:

1. The Gribov gauge-fixing ambiguity impedes the elimination of unphysical degrees of freedom from the canonical (temporal-gauge)formalism.

2. The running coupling constant becomes strong at large distances, and perturbation theory goes out of control.

3. Large instanton effects occur, complicating the question of topological structure of the QCD vacuum.

4. The vacuum rotor-modes mix with transverse-gluon modes and may have significant physical effects.

5. Gluonium and/or color-singlet hadrons should appear as asymptotic states and quarks and gluons should disappear. How (and at what distance scale) does this happen?

6. With light quarks present, spontaneous breakdown of chiral symmetry should occur. Again at what distance scale (and how) does this happen.

But while there remain many unanswered questions about the large-distance aspects of QCD, the small-distance behavior appears to be comprehensible. If the theory is correct, the problem of understanding strong-interaction dynamics can to a great extent be decoupled from the problem of interrelating the strong, weak and electromagnetic force at extremely short distances. That alone would be a great step forward in our understanding of elementary particles.

APPENDIX A

One of the Maxwell Equations

$$\left\{\vec{\underset{\sim}{D}} \cdot \vec{\underset{\sim}{E}}(x,t) - e\underset{\sim}{J}^{o}(x,t)\right\}\Psi = 0 \tag{A.1}$$

is an equation of constraint. The existence of this constraint is connected with the residual gauge-invariance of the theory: the condition $\underset{\sim}{A}_0=0$ does not completely fix the gauge. Time-independent gauge transformations

$$\vec{\underset{\sim}{A}} = \underset{\sim}{S}\vec{\underset{\sim}{A}}'\underset{\sim}{S}^{-1} + \frac{i}{e}(\vec{\underset{\sim}{\nabla}}S)\underset{\sim}{S}^{-1} \tag{A.2}$$

with

$$\frac{\partial \underset{\sim}{S}}{\partial t} = 0 \tag{A.3}$$

can be still carried out. The Gauss-Law operator $(\vec{\underset{\sim}{D}} \cdot \vec{\underset{\sim}{E}} - e\underset{\sim}{J}^{o})$ is actually the infinitesimal generator of these gauge transformations. To see this write

$$\underset{\sim}{S}(x) = 1 - ie\delta\underset{\sim}{\Lambda}(x) \tag{A.4}$$

with $\delta\underset{\sim}{\Lambda}(x)$ a 3x3 matrix. Then by definition

$$\vec{\delta}\underset{\sim}{A} = \vec{\underset{\sim}{A}} - \vec{\underset{\sim}{A}}' = [\delta\underset{\sim}{\Lambda}(x), \vec{\underset{\sim}{A}}(x)] + \vec{\nabla}\delta\underset{\sim}{\Lambda}(x) \tag{A.5}$$

Now in the big Hilbert-space of quantized fields we write the unitary transformation $\mathcal{U}$

$$\mathcal{U} = 1 + i\,\mathrm{Tr}\int d^3y\,[\underset{\sim}{D} \cdot \underset{\sim}{E}(y) - e\underset{\sim}{J}_o(y)]\,\delta\underset{\sim}{\Lambda}(y) \tag{A.6}$$

Then using the canonical commutators, we find

$$\mathcal{U}\vec{\underset{\sim}{A}}(x)\mathcal{U}^{-1} = \vec{\underset{\sim}{A}}(x) + \vec{\delta}\underset{\sim}{A}(x) \tag{A.7}$$

with $\delta A(x)$ given above in Equation (A.5).

Likewise the unitary transformation of the quarks is also generated:

$$\mathcal{U} q(x) \mathcal{U}^{-1} = (1 + ie\delta\underset{\sim}{\Lambda}(x))q(x) \tag{A.8}$$

The Hamiltonian has a very large symmetry under residual gauge transformations. The idea of gauge-fixing is to find all the extra trivial coordinates and remove them from the formalism. In QED this is easy to do, as was illustrated in Chapter 3.

Now we ask what the corresponding procedure is in QCD. Evidently we would like to copy QED as much as possible. We may do this as before by introducing transverse and longitudinal parts of $\vec{\underset{\sim}{E}}$ and $\vec{\underset{\sim}{A}}$ and eliminating $\underset{\sim}{A}_L$ and $\underset{\sim}{E}_L$ from the theory. We begin with the electromagnetic potential and assume there exists a time-independent gauge-transformation which renders $\vec{\underset{\sim}{A}}$ purely transverse

$$\vec{\nabla} \cdot \vec{\underset{\sim}{A}}' = 0 \quad \text{or} \quad \vec{\underset{\sim}{A}}' = \vec{\underset{\sim}{A}}'_T \tag{A.9}$$

This is not self-evident; the equation for the gauge-transformation $\underset{\sim}{S}$ is

$$\vec{\nabla} \cdot (\underset{\sim}{S}^{-1}\vec{\underset{\sim}{A}}\underset{\sim}{S} + \frac{i}{e}\underset{\sim}{S}^{-1}\vec{\nabla}\underset{\sim}{S}) = 0 \tag{A.10}$$

which is not at all transparent. (Note that for QED, with $S=e^{ie\Lambda}$, the above equation is just

$$\nabla^2\Lambda = \vec{\nabla} \cdot \vec{A} \tag{A.11}$$

which allows the solution

$$\Lambda(x) = \int d^3y \, \frac{1}{4\pi|x-y|}\left(\vec{\nabla}_y \cdot \vec{A}(y)\right) \tag{A.12}$$

If this is carried out, then $\vec{\underset{\sim}{A}}_L$ is removed from the Hamiltonian, inasmuch as H is invariant under such gauge transformations. To remove $\vec{\underset{\sim}{E}}_L$, define

$$\vec{\underset{\sim}{E}} = \vec{\nabla}\underset{\sim}{\phi} + \vec{\underset{\sim}{E}}_T \tag{A.13}$$

and eliminate $\underset{\sim}{\phi}$ using the Gauss' law constraint

$$\underset{\sim}{D} \cdot \vec{\underset{\sim}{E}} = \underset{\sim}{D} \cdot \nabla\underset{\sim}{\phi} + \underset{\sim}{D} \cdot \underset{\sim}{E}_T$$

$$= \vec{\nabla} \cdot \underset{\sim}{D}\underset{\sim}{\phi} - ie\left[\vec{\underset{\sim}{A}}, \cdot \vec{\underset{\sim}{E}}_T\right] = e\underset{\sim}{J}_0 \qquad \text{(A.14)}$$

It is easiest at this point to expand all 3×3 matrices in terms of λ-matrices. Then ϕ^A can be obtained formally via

$$\phi^A(x) = \int d^3y\ K_{AB}(x,y;A)\left[ef^{BCD}\vec{A}_T^D(y)\cdot\vec{E}_T^C(y) + eJ_0^B\right] \qquad \text{(A.15)}$$

with the kernel K "defined" by

$$(\vec{\nabla} \cdot \vec{D})_{AB}\ K(x,y;A)_{BC}$$

$$\equiv \nabla^2 K(x,y;A)_{AC} - ef^{ABD}\vec{A}_T^D \cdot \vec{\nabla}K_{BC} = \delta^3(x-y) \qquad \text{(A.16)}$$

In shorthand notation

$$K = \frac{1}{\vec{\nabla}\cdot\vec{D}} \qquad \text{(A.17)}$$

With ϕ defined above, we finally end up with the Hamiltonian

$$H = \int d^3x\left\{\mathrm{Tr}(\underset{\sim}{E}_T^2 + \underset{\sim}{B}^2 + (\vec{\nabla}\underset{\sim}{\phi})^2) + \Psi^\dagger(\vec{\alpha}\cdot(\vec{P} - e\underset{\sim}{A}_T) + \beta m)\Psi\right\} \qquad \text{(A.18)}$$

This somewhat cumbersome form directly generalizes the procedure in QED. It is quite satisfactory for weak-coupling applications, but unsatisfactory for strong coupling. This is because it is known that there exist homogeneous solutions for the Green's function K. Roughly speaking, the $\vec{A}_T\cdot\vec{\nabla}$ term is a potential which supports "bound states" provided A is big enough. If the region where A is large has a size L this means $A \gtrsim 1/eL$ in order to compensate the kinetic-energy term.

Another way of seeing what this means is to look at the gauge-condition

$$\vec{\nabla}\cdot\vec{\underset{\sim}{A}} = 0 \qquad \text{(A.19)}$$

We assumed this could be done in a unique way. However, existence of homogeneous solutions of $\vec{\nabla} \cdot \vec{\underset{\sim}{D}}\underset{\sim}{\phi} = 0$ implies that the above condition, Equation (A.19), is not unique: namely there exist gauge-equivalent fields which each satisfy $\vec{\nabla} \cdot \vec{\underset{\sim}{A}} = 0$. That is, when $\underset{\sim}{A}$ is chosen such that these solutions are close to each other we have

$$\vec{\nabla} \cdot \vec{\underset{\sim}{A}} = \vec{\nabla} \cdot (\vec{\underset{\sim}{A}} + \vec{\delta}\underset{\sim}{A}) = 0 \qquad \text{(A.20)}$$

However, from Equation (A.6), $(D \cdot \underset{\sim}{E} - e\underset{\sim}{J}_0)$ generates gauge transformations

$$\begin{aligned} \delta\vec{A} &= \int d^3y \, [\underset{\sim}{D} \cdot \underset{\sim}{E}(y)\delta\Lambda(y), \vec{\underset{\sim}{A}}(x)] \\ &= \vec{\underset{\sim}{D}}\delta\underset{\sim}{\Lambda} \end{aligned} \qquad \text{(A.21)}$$

Hence

$$\vec{\nabla} \cdot \vec{\underset{\sim}{D}}\delta\underset{\sim}{\Lambda} = 0 \qquad \text{(A.22)}$$

Non-trivial solutions of Equation (A.22) thus imply that fixing $\vec{\underset{\sim}{A}}$ to be transverse does not mean that the gauge has been fixed.

APPENDIX B

Even the vacuum of QCD turns out to have a quite complicated structure for essentially topological reasons. We shall enter this subject here from what is apparently quite an oblique way. It is not the historical path, but we hope it is more physically comprehensible.[99] We start by adding a term to the Lagrangian density of the theory with the form

$$\mathcal{L}' = \frac{e^2\theta}{32\pi^2} \operatorname{Tr} \epsilon_{\mu\nu\alpha\beta} \underset{\sim}{F}^{\mu\nu} \underset{\sim}{F}^{\alpha\beta}$$

$$= \frac{e^2\theta}{16\pi^2} \operatorname{Tr} \underset{\sim}{F}_{\mu\nu} \tilde{\underset{\sim}{F}}^{\mu\nu} = \frac{e^2\theta}{4\pi^2} \operatorname{Tr} \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}} \tag{B.1}$$

where the dual $\tilde{F}_{\mu\nu}$ of $F_{\mu\nu}$ is defined as

$$\tilde{\underset{\sim}{F}}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\alpha\beta} \underset{\sim}{F}^{\alpha\beta}$$

That is, the dual of $\vec{\underset{\sim}{E}}$ is $\vec{\underset{\sim}{B}}$, and vice versa. This interaction term is evidently parity-violating but C-conserving; hence by the TCP theorem also time-reversal non-invariant. We might ask why in the world one would gratuitously throw such a coupling into the strong interactions. One answer is why not: the term is renormalizable and there is a school of thinking that says that one should write down the most general renormalizable theory whenever possible. But even if one restricts the strong interactions by hand to be CP-invariant by setting $\theta = 0$, CP violation elsewhere in the theory can eventually leak back into the strong interaction and induce such a term. So it is reasonable to expect such a coupling at some level, even if it is quite small. However this is not the only reason that such a term is of interest. The $\theta F\tilde{F}$ interaction is quite peculiar, because it turns out that the Lagrangian density $\mathcal{L}'$ is a total divergence. This implies that the Lagrangian L can be written as a time derivative of a function of $\vec{\underset{\sim}{A}}$. To see this easily we go to $\underset{\sim}{A}_0=0$ gauge and write

$$L = \frac{\theta e^2}{4\pi^2} \mathrm{Tr} \int \vec{\underset{\sim}{E}} \cdot \vec{\underset{\sim}{B}}\, d^3x$$

$$= \frac{\theta e^2}{4\pi^2} \mathrm{Tr} \int d^3x\, \varepsilon_{ijk} \frac{\partial \underset{\sim}{A}_i}{\partial t} \left[\frac{\partial \underset{\sim}{A}_i}{\partial x_k} - ie \underset{\sim}{A}_j \underset{\sim}{A}_k \right]$$

$$= \frac{\partial}{\partial t} \left\{ \frac{\theta e^2}{4\pi^2} \mathrm{Tr} \int d^3x\, \varepsilon_{ijk} \left[\frac{1}{2} \underset{\sim}{A}_i \frac{\partial \underset{\sim}{A}_j}{\partial x_k} - \frac{ie}{3} \underset{\sim}{A}_i \underset{\sim}{A}_j \underset{\sim}{A}_k \right] \right\} \tag{B.2}$$

where we have assumed that $\vec{\underset{\sim}{A}}$ vanishes as $|\vec{x}| \to \infty$ so that integration by parts may be carried out. Now when the Lagrangian contains a total time derivative, it is a simple matter to calculate its effect. We review this for a system with a finite number of degrees of freedom. Let

$$L(q_i, \dot{q}_i) = L_0(q_i, \dot{q}_i) + \frac{dF(q_i)}{dt}$$

$$= L_0(q_i, \dot{q}_i) + \sum_i \dot{q}_i \frac{\partial F}{\partial q_i} \tag{B.3}$$

The equations of motion are unchanged:

$$0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = \underbrace{\left[\frac{d}{dt}\frac{\partial L_0}{\partial \dot{q}_i} - \frac{\partial L_0}{\partial q_i} \right]}_{\text{old}} + \underbrace{\left[\frac{d}{dt}\frac{\partial F}{\partial q_i} - \dot{q}_i \frac{\partial^2 F}{\partial {q_i}^2} \right]}_{\text{new}} \tag{B.4}$$

The Hamiltonian formalism changes a little. The new momentum is

$$p_i = \frac{\partial L_0}{\partial \dot{q}_i} + \frac{\partial F}{\partial q_i} = p_i^{(o)} + \frac{\partial F}{\partial q_i} \tag{B.5}$$

However the Hamiltonian remains form-invariant (in order to preserve the equations of motion!):

$$H(p_i,q_i) = \sum_i p_i\dot{q}_i - L$$

$$= \left[\sum_i p_i^{(o)}\dot{q}_i - L_0\right] + \left[\sum_i \frac{\partial F}{\partial q_i}\dot{q}_i - \frac{dF}{dt}\right]$$

old new

$$= H_0\left(p - \frac{\partial F}{\partial q} , q\right)$$

and

$$\left[p_i,q_i\right] = \left[p_i^{(o)},q_j\right] = -i\delta_{ij} \qquad (B.6)$$

Thus the effect of the extra term F can be taken into account by incorporating an extra phase into the wave function. That is, if

$$H_0\Psi_0(q,t) = i\,\frac{\partial\Psi_0}{\partial t} \qquad (B.7)$$

then upon defining

$$\Psi(q,t) = e^{-iF(q)}\Psi_0(q,t) \qquad (B.8)$$

one finds

$$H(p,q)\Psi(q,t) = i\,\frac{\partial\Psi(q,t)}{\partial t} \qquad (B.9)$$

inasmuch as $e^{iF(q)}p\,e^{-iF(q)} = p - \frac{\partial F}{\partial q}$. This looks like it has a trivial effect on the theory. The spectrum of eigenfunctions and eigenvalues of H are simply determined in terms of those of H_0. In particular the energy-eigenvalues are the same. For QCD, the transformation on the great big wavefunction for the field-theory is given by

$$\Psi_\theta(A) = e^{-iN(A)\theta}\Psi_0(A) \qquad (B.10)$$

with

$$N(A) = \frac{e^2}{4\pi^2}\,\mathrm{Tr}\int d^3x\;\epsilon_{ijk}\left[\frac{1}{2}A_i\,\frac{\partial A_j}{\partial x_k} - \frac{ie}{3}A_iA_jA_k\right] \qquad (B.11)$$

Why is this of any interest at all? For pure time-independent gauges ($\vec{\underset{\sim}{E}} = \vec{\underset{\sim}{B}} = 0$) the operator N takes the form

$$\hat{N} \equiv N_{\text{pure gauge}} = \frac{-ie^3}{24\pi^2} \operatorname{Tr} \int d^3x \, \varepsilon_{ijk} \underset{\sim}{A}_i \underset{\sim}{A}_j \underset{\sim}{A}_k \tag{B.12}$$

with

$$\vec{\underset{\sim}{A}} = \underset{\sim}{U}^{-1} \vec{\nabla} \underset{\sim}{U} \tag{B.13}$$

It is here that topological considerations enter. If we make an infinitesimal, time-independent gauge transformation on the above $\vec{A}$ which vanishes at ∞ we find

$$\underset{\sim}{A}_i(x) + [\delta\underset{\sim}{\Lambda}(x), \underset{\sim}{A}_i(x)] + \nabla_i \delta\underset{\sim}{\Lambda}(x) \tag{B.14}$$

and

$$\delta\hat{N} = \frac{-ie^2}{24\pi^2} \operatorname{Tr} \int d^3x \left\{ \varepsilon_{ijk} \left[\delta\underset{\sim}{\Lambda}, \underset{\sim}{A}_i \underset{\sim}{A}_j \underset{\sim}{A}_k \right] + 3\varepsilon_{ijk} \frac{\partial}{\partial x_i} \left(\delta\underset{\sim}{\Lambda} \underset{\sim}{A}_j \underset{\sim}{A}_k \right) \right\}$$

$$= 0 \tag{B.15}$$

Thus <u>N is invariant under infinitesimal gauge transformations which approach the identity transformation as $x \to \infty$</u>. Thus it is also invariant under finite gauge transformations U of the same type which can be reached from the identity continuously. It might seem possible therefore to restrict ourselves to potentials $\vec{\underset{\sim}{A}}$ which can be reached continuously from the trivial potential $\vec{\underset{\sim}{A}} = 0$; in this case (in the absence of gauge fields $\underset{\sim}{E}$ and $\underset{\sim}{B}$), N would vanish and the wave function would not depend upon θ. What follows will demonstrate that things aren't that simple and that (1) there exist other pure-gauge field-configurations $\vec{\underset{\sim}{A}}$ which cannot be reached from the identity by gauge-transformations (at least those which approach unity at $\vec{x} \to \infty$), and which are characterized by integer eigenvalues n of the N-operator, and (2) these n states are dynamically coupled together (by the instanton phenomenon), and that (3) the coupling is important enough that physical observables depend upon θ, despite the fact that its presence in the Lagrangian did not affect the equations of motion of the theory.

To do this, we simply exhibit a prototype example of a gauge function which produces a non-vanishing value of N. We write, as in the text

$$\underset{\sim}{U} = C \exp\left[\frac{i\vec{\tau}\cdot\vec{r}}{r} f(r)\right] \tag{B.16}$$

with C some constant matrix and with the Pauli-matrices $\vec{\tau} = (\lambda_1, \lambda_2, \lambda_3,)$

$$\tau_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \tau_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \tau_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{B.17}$$

defined as the first 3 of Gell-Mann's λ-matrices. The topological games come from the curious coupling of internal-symmetry matrices $\vec{\tau}$ to coordinates. In order that $\underset{\sim}{U} \to \underset{\sim}{1}$ as $r \to \infty$, we must have, as $r \to \infty$

$$f(r) \to \pi n \qquad \text{and} \qquad C = (-1)^n \tag{B.18}$$

with n an integer. In order that $\underset{\sim}{U}$ be nonsingular as $r \to 0$, we must also have $f(r) \to 0$ as $r \to 0$; hence $\underset{\sim}{U} \to (-)^n$ as $r \to 0$. The integer n will in fact provide the characterization of the n-vacuua. Because N is invariant under continuous gauge transformations (provided they are trivial at the boundary) we may choose a simple form for f(r), in particular push it out near to the boundary of the region, which we take as a large sphere of radius R. Notice that the action of $\underset{\sim}{U}$ is to rotate a quark in the internal space by $n\pi$ as one goes out to ∞ from the origin. The axis of rotation in the internal space is dependent on the direction in real space that one travels. To calculate $\hat{N}$ we take a small solid angle $\Delta\Omega$ near the north pole ($x_3 \sim R$; x_1, x_2 small) and calculate the components of $\vec{\underset{\sim}{A}}$ with

$$ie\underset{\sim}{A}_i = \underset{\sim}{U}^{-1}\nabla_i\underset{\sim}{U} \tag{B.19}$$

and $\underset{\sim}{U}$ given by Equation (B.16). Elementary estimates give

$$ie\underset{\sim}{A}_x \cong \underset{\sim}{U}^{-1}C\left[\frac{i\tau_1}{r}\sin f(r)\right] + \text{terms odd in } x$$

$$ie\underset{\sim}{A}_y \cong \underset{\sim}{U}^{-1}C\left[\frac{i\tau_2}{r}\sin f(r)\right] + \text{terms odd in } y$$

$$ie\underset{\sim}{A}_z \cong i\tau_3\frac{\partial f}{\partial r} \tag{B.20}$$

The terms odd in x and y will vanish upon averaging over a region centered symmetrically about $x = y = 0$. Insertion into Equation (B.12) for $\hat{N}$ yields

$$\Delta\hat{N} \sim \frac{-1}{24\pi^2}\cdot 6 \cdot 2i\int_{x,y\ \text{small}} dxdy\int_0^R \frac{dr}{R^2}\sin^2 f(r)f'(r)$$

or

$$\hat{N} = \frac{1}{2\pi^2}\int d\Omega\int_0^{\pi n} df\ \sin^2 f = n \tag{B.21}$$

What have we accomplished? For one thing we have finally uncovered the reason for the factor $32\pi^2$ in Equation (B.1). But we emphasize that it is not yet physics: we have merely found a topological classification of gauge-potentials $\vec{\underset{\sim}{A}}$ in terms of a single integer n. If one restricts oneself to potentials which vanish at spatial infinity, then there is no gauge-transformation obtainable from the identity by a product of infinitesimal transformations which takes a potential characterized by one n-value to a potential characterized by a different n-value.

Enter the instanton. While the n-vacuua are only coupled by gauge transformations nonvanishing at the surface at ∞, there can be dynamical coupling of the n-vacuua. For simplicity consider traveling in $\underset{\sim}{A}$-space (remember the $\underset{\sim}{A}(x)$'s are coordinates in the great big Hilbert-space; the $\underset{\sim}{E}(x)$'s are momenta) from the origin $\underset{\sim}{A} = 0$ in a straight line to an $\underset{\sim}{A}(x)$ in the topological sector $n = 1$. The intermediate $\underset{\sim}{A}$'s cannot be pure gauge; hence by definition there will be $\underset{\sim}{B}$-field in between. Thus the potential energy

$$V(A) = \mathrm{Tr}\int d^3x\, \underset{\sim}{B}^2(x) \tag{B.22}$$

increases as one proceeds away from $n = 0$ and then must decrease again as one reaches the $n = 1$ pure-gauge configuration (c.f. Figure 6). Because the great big Hilbert space is so multidimensional, there are many paths one may travel in going from $\underset{\sim}{A} = 0$ to the $n = 1$ gauge configuration; nevertheless they all share the feature that there is a potential barrier (unless one travels via the coordinates on the boundary). However even if one cannot go around, one can still go through the barrier by quantum mechanical tunnelling. Suppose that e is small. We saw that for a glob of potential with $n > 0$ and spatial extent $\sim L$, we must have dimensionally (c.f. Equation (B.12))

$$e^3A^3L^3 \gtrsim 1 \tag{B.23}$$

or

$$A \gtrsim \frac{1}{eL} \tag{B.24}$$

The intermediate B-field would then be expected to be

$$B \sim \frac{1}{eL^2} \tag{B.25}$$

and

$$V(A) \sim \frac{1}{e^2 L} \qquad (B.26)$$

The height of the barrier is $\sim e^{-2}$; hence for small e we have a large thick barrier and a small tunnelling probability. For such a situation (and only for such a situation) there exists an easy, albeit rather crude, estimate for the tunnelling amplitude: it is the semiclassical approximation. To get the answer, again retreat to a system with a finite number of degrees of freedom.

$$H = \frac{1}{2} \sum_{i=1}^{N} p_i^2 + V(q_1, \ldots q_n) \qquad (B.27)$$

with $V(0) = 0$ and $V(Q_1, \ldots, Q_N) = 0$. For a big thick barrier we write in the classically forbidden region

$$\Psi(q) \sim e^{-S(q)} \qquad (B.28)$$

The tunnelling amplitude is then $\sim e^{-[S(Q) - S(0)]}$. When S is large (coefficient e^{-2}), the unwritten factors which normalize the wave function are relatively inconsequential. There does exist a rather sophisticated technology[100] for their calculation (the old-fashioned way involves what is known as the Van Vleck determinant[101]) but we shall not go into this aspect of the problem here. It will be enough to obtain S(Q). The equation for S in semiclassical approximation ($\partial V/\partial q$ small compared to rate of falloff of ψ) is

$$\frac{1}{2} \sum_i \left(\frac{\partial S}{\partial q_i}\right)^2 + V(q_i) \approx 0 \qquad (B.29)$$

Other than a change in relative sign, this is just the Hamilton-Jacobi equation for the phase point $\vec{q}$ in the potential V(q). The sign-charge (which originates in the e^{-S} instead of the WKB e^{iS}) allows us to interpret Equation (B.29) as the real classical motion of the phase-point $\vec{q}$ starting at rest at the origin and ending at the point Q (again at rest) in an attractive potential $V'(q)=-V(q)$. The quantity S(Q) is, from classical theory, the action J associated with such a classical motion starting at 0 and ending at Q. To see this, simply retreat from the Hamiltonian formalism, writing

$$p_i = \dot{q}_i = \frac{\partial S(q)}{\partial q_i} \qquad H = \frac{1}{2}\sum_i p_i^2 + V'(q) = 0$$

$$L = \sum p_i \dot{q}_i - H = \sum_i p_i q_i = \sum_i \dot{q}_i \frac{\partial S}{\partial q_i} \tag{B.30}$$

Hence the action J is given by

$$J = \int_{-\infty}^{\infty} L\, dt = \int_{-\infty}^{\infty} dt \sum_i \dot{q}_i \frac{\partial S}{\partial q_i} = S(Q) - S(0) \tag{B.31}$$

inasmuch as the motion begins at $q_i=0$ and ends at $q_i=Q_i$.

One last point before reverting to the real continuum QCD problem: instead of changing the sign of the potential, we could equally well change t to it, since the equations of motion

$$\ddot{q} = \frac{-\partial V}{\partial q} \tag{B.32}$$

undergo the desired sign change.

Let us now recapitulate. We get a tunnelling amplitude by the following procedure

1) replace t by it;

2) solve the classical equations for this system, requiring $\vec{E} = \vec{B} = 0$ at $t \to \infty$ (and for us $\vec{E}(x)$ and $\vec{B}(x) \to 0$ as $\vec{x} \to \infty$ as well);

3) calculate the classical action J for this solution. The desired tunnelling amplitude is then $\sim e^{-J}$.

We see that we need a classical solution of the Euclidean QCD equations with finite action in order to couple together all the n-vacuua. This solution was found by Belavin et al.[34] To find it, it is best to abandon $A_0=0$ gauge and exploit the full 4-dimensional symmetry of Euclidean QCD. We search first for the form of the potentials A_μ as $\rho = \sqrt{t^2 + x^2} \to \infty$. They must be pure gauge, fall off as ρ^{-1} and involve the Pauli-matrices $\vec{\tau}$. A choice which satisfies this and looks right is

$$U = \frac{-t + i\vec{\tau} \cdot \vec{r}}{\sqrt{t^2 + r^2}} \tag{B.33}$$

This has a singularity at the origin but, as we shall soon see, otherwise has the right properties to induce a transition between states of differing n. Then a choice for the solution is

$$ieA_\mu = f(\rho)U^{-1}\partial_\mu U \tag{B.34}$$

with

$$\begin{aligned} f(\rho) &\to 1 \qquad \rho \to \infty \\ f(\rho) &\to 0 \qquad \rho \to 0 \end{aligned} \tag{B.35}$$

Direct, tedious calculation shows that

$$f(\rho) = \frac{\rho^2}{\rho^2 + \lambda^2} \tag{B.36}$$

does the job. The potential $\underset{\sim}{A}_\mu$ is indeed no longer singular at $x = 0$, and the QCD equations of motion are satisfied.

We now can put together all the pieces. In doing this let us again recapitulate what we have done. We introduced a term into the Lagrangian density $\mathcal{L}'$ of the form $\frac{e^2\theta}{16\pi^2}\,\mathrm{Tr}\underset{\sim}{F}_{\mu\nu}\underset{\sim}{\tilde{F}}^{\mu\nu}$. This term is a total divergence $\theta\partial_\mu K^\mu$; hence the Lagrangian L was modified by a total time derivative $\theta\frac{\partial}{dt}\int K^0 d^3x$ which simply put a coordinate-dependent phase on the wave function in the great big Hilbert space. The phase-operator $\theta N = \theta\int d^3x K^0$ in a representation in which the potentials are diagonal, measures a topological property of the potentials, characterized by an integer n. Tunnelling between states of different n is possible if there exist classical solutions of the equations of motion with finite action, and which (in $\underset{\sim}{A}_0=0$ gauge) take one from a pure gauge configuration of given n at $t=-\infty$ to a pure gauge configuration of a different n at $t=+\infty$. We apparently found such a solution, albeit in a different gauge. To show that the solution indeed induces a tunnelling transition between states of different n, we introduce in the instanton gauge potential at large distances, Equation (B.34), the polar coordinates shown in Figure 23.

$$t = \rho\cos\psi$$
$$r = \rho\sin\psi \qquad (B.37)$$

We may then write

$$\underset{\sim}{U} = -e^{i\frac{\vec{\tau}\cdot\vec{r}\psi}{r}} \qquad (B.38)$$

Hence at $\psi = 0$ $(t\to+\infty,\ r\to 0)$ $U=-1$,

while at $\psi = \pi$ $(t\to-\infty,\ r\to 0)$ $U=+1$. Now make a <u>continuous</u> gauge transformation to a new $\underset{\sim}{U}$ such that

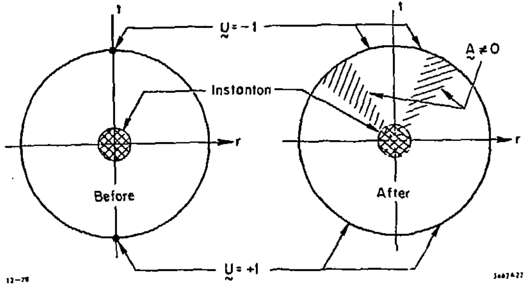


Fig. 23. Effects of a gauge transformation on the asymptotic gauge potentials surrounding an instanton: (a) "Covariant gauge" and (b) a gauge in which $\vec{A} = 0$ in the past and $\vec{A}$ is an expanding gauge-bubble in the future.

$$\underset{\sim}{U} = -e^{\frac{i\vec{\tau}\cdot\vec{r}}{r} f(\psi)} \tag{B.39}$$

with $f(0) = 0$ and $f(\pi) = \pi$ unchanged, but where $f(\psi)$ makes the flip from 0 to π at small angles (i.e., in the future). Thus in the past $\underset{\sim}{U} \approx 1$ and $\vec{\underset{\sim}{A}} \cong 0$, while in the future $\underset{\sim}{U}$ takes the same form as we had for the prototype gauge configuration with $n = 1$. It follows that the instanton induces a tunnelling transition between topological sectors with $|\Delta n| = 1$.

To calculate the tunnelling amplitude associated with a single instanton, we need calculate only the Euclidean action J of the instanton. It is

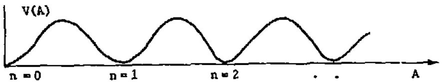
$$\begin{aligned} J &= \mathrm{Tr}\int d^4x\ (\underset{\sim}{E}^2 + \underset{\sim}{B}^2) \\ &\geq 2\mathrm{Tr}\int d^4x\ \vec{\underset{\sim}{E}}\cdot\vec{\underset{\sim}{B}} \end{aligned} \tag{B.40}$$

by the Schwartz inequality. We use this inequality because we recall that $\vec{\underset{\sim}{E}}\cdot\vec{\underset{\sim}{B}}$ is our total divergence, already evaluated from topological considerations. From Equations (B.2) and (B.11)

$$2\mathrm{Tr}\int d^4x\ \vec{\underset{\sim}{E}}\cdot\vec{\underset{\sim}{B}} = \frac{8\pi^2}{e^2}\int dt\ \frac{\partial N}{\partial t} = \frac{8\pi^2}{e^2}\left[N(\infty) - N(-\infty)\right] = \frac{8\pi^2}{e^2} \tag{B.41}$$

Hence the tunnelling amplitude is $\lesssim e^{-8\pi^2/e^2}$. It is in fact the case that the instanton solution is self-dual, i.e., $\vec{\underset{\sim}{E}} = \vec{\underset{\sim}{B}}$, and the inequality is an equality.

We now can crudely estimate the effect on the wave function of the system. The situation is analogous to the problem of a particle in a periodic potential



The eigenfunctions are not packets localized near $\underset{\sim}{A} = 0$ (n = 0), but instead the Bloch-waves

$$\Psi \cong \sum_n e^{in\theta} \Psi_n(A) \qquad (B.42)$$

This is in fact the structure we already constructed from use of the gauge-invariant surface term. This much follows essentially from a requirement of gauge-invariance and not of existence of instantons. What the instanton does is to provide a coupling between n-vacuua which is not a surface-effect, but a volume effect. That is, the effective Hamiltonian for tunnelling between adjacent n-vacuua will, for dimensional reasons, be

$$H' \sim \sum_n V \int \frac{d\lambda}{\lambda^5} e^{-\frac{8\pi^2}{e^2(\lambda)}} \delta n,n-1 \qquad (B.43)$$

The volume factor V occurs because the instanton can be located anywhere in space (i.e., the choice of which set of coordinates $\underset{\sim}{A}_i(x)$ do the tunnelling is open and must be summed over. Likewise the instanton size λ is arbitrary and must be summed over). Since H' has dimension of inverse length, dimensional analysis gives the weight factor. There is an additional factor e^{-12} which requires study of the Van Vleck determinant which normalizes the wave function at the classical turning point.[00,101] The integral over λ is infrared-divergent; hence in a finite and very small volume, the most important instanton is the one which just fits into the box. If the box is small enough that perturbation theory is justified, then in fact the instanton effects remain very small.

The effect of H' on a θ-state is simple to calculate. Recall

$$|\theta\rangle = \sum_n e^{in\theta}\psi_n(A) \tag{B.44}$$

and

$$\frac{H'}{V}|\theta\rangle = \frac{k}{e^{12}}\int\frac{d\lambda}{\lambda^5}\, e^{-\frac{8\pi^2}{e^2(\lambda)}} \sum_n e^{in\theta}\Big[\Psi_{n-1}(A) + \Psi_{n+1}(A)\Big] \tag{B.45}$$

where $\Psi_{n-1}(A)$ and $\Psi_{n+1}(A)$ differ from $\Psi_n(A)$ by an additional lump of gauge configuration somewhere. Thus a θ-vacuum is an eigenfunction of H'/V with eigenvalue

$$H'|\theta\rangle = \frac{-2V\ k\cos\theta}{(e^2)^6}\int\frac{d\lambda}{\lambda^5}\, e^{-\frac{8\pi^2}{e^2(\lambda)}}\,|\tilde{\theta}\rangle \tag{B.46}$$

where in general $|\tilde{\theta}\rangle$ might not be equal to $|\theta\rangle$ but could be deformed into it by a <u>continuous</u> gauge transformation. However physical states are to be gauge-<u>invariant</u>; hence for states $|\theta\rangle$ in that subspace of the great big Hilbert space which are physical, we can replace $|\tilde{\theta}\rangle$ by $|\theta\rangle$. Hence for practical purposes H' is in fact diagonal. We see in particular that the vacuum energy <u>density</u> is θ-dependent, with $\theta = 0$ being the state of lowest energy.

REFERENCES

1. If you don't know, you are invited to consult someone else's lectures.
2. S. Brodsky, these proceedings.
3. An introduction to this method can be found in J. C. Taylor, "Gauge Theories of the Weak Interactions," Cambridge University Press, 1976.
4. Actually an exception to this might be the pion in the limit of massless u and d quarks. In that circumstance pions are believed to be massless Goldstone bosons. However, none of that appears to have much connection with the phenomenon of QCD confinement.
5. A recent review is given by S. Weinberg, I. I. Rabi Festschrift, to be published by the New York Academy of Sciences.
6. See, for example, D. Robson, Nucl. Phys. B130, 328 (1977), also K. Koller and T. Walsh, Nucl. Phys. B140, 449 (1978). A recent detailed study is given by J. Coyne, P. Fishbane and S. Meshkov, NBS/Virginia preprint (1980).
7. S. Weinberg, Phys. Rev. D11, 3583 (1975).
8. See Robson, Ref. 6, for a discussion.
9. Actually an indefinite number of gluons might "bleach" the color field and produce a color singlet state (of fractional charge). That this is a theoretical possibility—although not a very desirable one—was emphasized to me by Marvin Weinstein.
10. For a review, see J. Weis, Acta. Phys. Polon. B9, 1051 (1978).
11. W. Heisenberg and H. Euler, Zeitschrift fur Physik 98, 714 (1936); J. Schwinger, Phys. Rev. 82, 664 (1951).
12. A recent study of this mechanism has been made by A. Casher, H. Neuberger, and S. Nussinov, Phys. Rev. D20, 179 (1979); also

Tel-Aviv preprint TAUP-747/79; see also H. Neuberger, preprint TAUP-721/79.

13. However, the OZI rule does give some motivation for a small coupling.
14. I thank Stanley Mandelstam for pointing this out to me.
15. J. Bjorken, Lecture Notes in Physics 56, "Current Induced Reactions," ed. by J. Koerner, G. Kramer, and D. Schildknecht, Springer-Verlag (N.Y.), 1975.
16. Were $\underline{A}^{\mu}$ not taken to be traceless, the coefficient of the unit matrix would be a U(1) degree of freedom not coupled to color at all. There is no evidence for its existence as a physical degree of freedom.
17. A hint: the quantity Tr $\underline{J}_{\mu}\underline{J}^{\mu}$ is a color invariant.
18. The Lagrangian density is $\mathcal{L} = \mathrm{Tr}\ (\underline{E}^2-\underline{B}^2)$.
19. L. Faddeev and V. Popov, Phys. Lett. 25B, 29 (1967).
20. Our notation follows that of Reference 1.
21. S. Frenkel, Phys. Rev. D13, 2325 (1976), W. Konetschny and W. Kummer, Nucl. Phys. B100, 106 (1975); see also J. Willemsen, Phys. Rev. D17, 574 (1978).
22. This can also be explicitly checked without much difficulty.
23. V. N. Gribov, SLAC preprint SLAC-TRANS-176; S. Mandelstam, invited talk at Washington Meeting of the American Physical Society, 1977.
24. Generalizations of the problems have been made by M. Atiyah and I. Singer. Quantization in axial gauge ($A_3=0$) provides some improvement; yet there remains some residual gauge invariance, as well as possible problems with boundary terms. R. Jackiw and J. Goldstone [Phys. Lett. 74B, 81 (1978)] have presented an SU(2)

example of quantization without subsidiary conditions; however the resultant formalism contains some possibly troublesome singular expressions, as well as appearing to be rather unmanageable. See also V. Baluni and B. Grossman, Phys. Lett. 78B, 226 (1978).

25. We are thinking here of the LSZ formalism as used in Book II.

26. T. D. Lee and M. Nauenberg, Phys. Rev. 133, B1549 (1964); T. Kinoshita, J. Math. Phys. 3, 650 (1962).

27. S. Drell and T. M. Yan, Phys. Rev. Lett. 25, 316 (1970).

28. See the lectures of R. Stroynowski, these proceedings, for an up-to-date review.

29. K. Wilson, Phys. Rev. 179, 1499 (1969).

30. G. Sterman and S. Weinberg, Phys. Rev. Letters 39, 1436 (1977).

31. This problem is under study in collaboration with M. Fischler.

32. R. Jackiw and C. Rebbi, Phys. Rev. Letters 37, 172 (1976); C. Callan, R. Dashen, and D. Gross, Phys. Letters 63B, 334 (1976).

33. Provided, of course, one does not travel via field variables $\vec{A}(x)$ with $\vec{x}$ on the boundary of the box.

34. A. Belavin, A. Polyakov, A. Schwartz, and Y. Tyupkin, Phys. Lett. 59B, 85 (1975).

35. C. Callan, R. Dashen, and D. Gross, Phys. Rev. D17, 2717 (1978).

36. S. Coleman, "The Uses of Instantons," Harvard University preprint to be published in Proceedings of the 1977 International School of Subnuclear Physics "Ettore Majorana."

37. This has been verified in field-theoretical contexts by J. Willemsen, UC-Santa Cruz preprint (1979).

38. L. Baulieu, J. Ellis, M. K. Gaillard, and W. Zakrzewski, Phys. Lett. 77B, 290 (1978); also N. Andrei and D. Gross, Phys. Rev. D18, 468 (1978); R. Carlitz and C. Lee, Phys. Rev. D17, 3238 (1978).

39. G. Veneziano, Nucl. Phys. B117, 519 (1976).

40. D. Robson op. cit.; Reference 6.

41. H. Fritzsch and P. Minkowski, Nuovo Cimento 30A, 393 (1975).

42. R. Jaffe and K. Johnson, Phys. Lett. 60B, 201 (1976).

43. K. Johnson (private communication) argues that the pressure on the bag from the gluon field implies $E^2 > B^2$. The modes considered in Ref. 42 do not satisfy that constraint.

44. See References 6 and 39; also, P. Freund and Y. Nambu, Phys. Rev. Letters 34, 1646 (1975).

45. J. Bolzan, K. Geer, W. Palmer, and S. Pinsky, Phys. Rev. Letters 35, 419 (1975); also J. Bolzan, W. Palmer, a d S. Pinsky, Phys. Rev. D14, 3202 (1976).

46. See for example E. Witten, Nucl. Phys. B156, 269 (1979) for a nice discussion.

47. I am grateful to R. Brower for emphasizing this to me and for very useful discussions.

48. M. Chanowitz, Phys. Rev. D12, 918 (1975); S. Brodsky, D. Coyne, T. DeGrand, and R. Horgan, Phys. Lett. 73B, 203 (1978).

49. R. Barbieri, E. d'Emilio, G. Curchi, and E. Remiddi, Nucl. Phys. B154, 535 (1979).

50. M. Krammer, Phys. Letters 74B, 361 (1978).

51. A. Billoire, R. Lacaze, A. Morel, and H. Navelet, Phys. Lett. 80B, 381 (1979).

52. R. Brandelik et al., Phys. Lett. 74B, 292 (1978).

53. J. Schwinger, Phys. Rev. 125, 1043 (1962); 127, 324 (1962). See also Ref. 59.

54. The approach discussed here has been independently worked out by S. Drell and also D. Stump (private communications). They have studied the nature of this term in greater depth.

55. This is a consequence of the renormalizability of the theory, a fact which is not self-evident in this formalism. A direct check that the next term in the expansion behaves properly is at present being carried out by S. Drell and R. Hughes (private communication).

56. See for example Chapter 19 of Book II.

57. The history of renormalizability is traced quite completely by M. Veltman, Proceedings of the 6th International Symposium on Electron and Photo Interactions at High Energies, ed. H. Rollnik and W. Pfeil, North-Holland (Amsterdam), 1974; p. 429.

58. D.R.T. Jones, Nucl. Phys. B75, 531 (1974 W. Caswell, Phys. Rev. Lett. 33, 244 (1974).

59. T. Appelquist, M. Dine, and I. Muzinich, Phys. Lett. 69B, 231 (1977), also Phys. Rev. D8, 2074 (1978).

60. H. Thacker, C. Quigg, and J. Rosner, Phys. Rev. D18, 287 (1978).

61. J. Richardson, Phys. Lett. 82B, 272 (1979). See also B. Margolis, R. Roskies, and N. de Takacsy, McGill University preprint (1978); R. Levine and Y. Tomozawa, Phys. Rev. D19, 1572 (1979).

62. M. Krammer and H. Krasemann, preprint DESY 79/20.

63. For a review, see J. Jackson, Proceedings of the 1977 European Conference on Particle Physics, Budapest, ed. by L. Jenik and I. Montvay (CRIP, Budapest, 1978), p. 601; also K. Gottfried, Proceedings of the 1977 International Symposium on Leptons and Photons at High Energies, Hamburg, ed. P. Gutbrod (DESY, Hamburg, 1978), p. 667.

64. N. Isgur and G. Karl, Phys. Rev. D19, 2653 (1979).

65. A recent review is given by C. Quigg, Fermilab preprint FERMILAB-CONF 79/74 THY, to be published in the Proceedings of the 9th International Conference on Leptons and Photons at High Energy, Fermilab, Batavia, IL (1979).

66. K. Gottfried, Phys. Rev. Lett. 40, 598 (1978); M. Peskin, Nucl. Phys. B156, 365 (1979); G. Bhanot and M. Peskin, Nucl. Phys. B156, 391 (1979).

67. A. Ore and J. Powell, Phys. Rev. 75, 1696 (1949).

68. E. Witten, Harvard preprint HUTP-79/A007 (1979).

69. M. Shifman, A. Vainstein, and V. Zakharov, Nucl. Phys. B147, 385, 448 (1978).

70. V. Novikov, M. Shifman, A. Vainstein, and V. Zakharov, ITEP preprint (1979).

71. J. Bell and R. Jackiw, Nuovo Cimento 60A, 47 (1969); S. Adler, Phys. Rev. 177, 2426 (1969).

72. G. 't Hooft, Phys. Rev. Lett. 37, 8 (1976); Phys. Rev. D14, 432 (1976).

73. This idea seems to have originated with Callan, Dashen, and Gross, Ref. 35.

74. The adiabatic hypothesis is in fact unnecessary. A comprehensive study has been made by N. Christ (Columbia University Preprint CU-TP-160 (1979)), who relates all this to the work of M. Atiyah and I. Singer on spectral flow and the index theorem [M. Atiyah, N. Hitchen, and I. Singer, Proc. Nat. Acad. Sci. USA 74, 2662 (1977); M. Atiyah, V. Patodi, and I. Singer, Math. Proc. Camb. Phil. Soc. 77, 43 (1975); 79, 71 (1976)].

75. J. Kiskis [Phys. Rev. D18, 3690 (1978)] has explicitly identified the levels as they cross zero energy. See also L. Brown, R. Carlitz, and C. K. Lee, Phys. Rev. D16, 417 (1977), A. Schwartz, Phys. Lett. 67B, 172 (1977), R. Jackiw and C. Rebbi, Phys. Rev. D16, 1052 (1977); N. K. Nielsen and B. Schroer, Nucl. Phys. B127, 493 (1977).

76. M. Baker, K. Johnson, and R. Willey, Phys. Rev. 163, 1699 (1967); J. Cornwall and R. Norton, Phys. Rev. D8, 3338 (1973); R. Jackiw and K. Johnson, Phys. Rev. D8, 2386 (1973).

77. C. Callan, R. Dashen, and D. Gross, Phys. Rev. D12, 2717 (1978); see also D. Caldi, Phys. Rev. Lett. 39, 121 (1977); R. Carlitz and D. Creamer, Annals of Physics 118, 429 (1979).

78. H. Quinn and M. Weinstein (private communication).

79. There is a solid current-algebra calculation connection $\bar{\theta}$ with the neutron electric dipole moment; R. Crewther, P. diVecchia, G. Veneziano, and E. Witten, CERN preprint CERN-TH-2735 (1979); see in this connection V. Baluni, Phys. Rev. D19, 2227 (1979).

80. J. Ellis and M. K. Gaillard, Nucl. Phys. B150, 141 (1979); B. Morel, Harvard preprint HUTP-79/A009 (1979).

81. F. Wilczek, Phys. Rev. Lett. 40, 279 (1978).

82. R. Peccei and H. Quinn, Phys. Rev. Lett. 38, 1440 (1977).

83. S. Weinberg, Phys. Rev. Lett. 40, 223 (1978).

84. F. Wilczek, Phys. Rev. Lett. 40, 279 (1978).

85. T. Donnelly, S. Freedman, R. Lytel, R. Peccei, and M. Schwartz, Phy. Rev. D18, 1607 (1978).

86. The reader is strongly urged to consult the lectures of Sidney Coleman, Ref. 36, for a complete and lucid explanation of the methodology.

87. K. Wilson, Phys. Rev. D10, 2445 (1975).

88. J. Kogut, R. arson, and J. Shigemitsu, Phys. Rev. Lett. 43, 484 (1979).

89. M. Creutz, Brookhaven preprint, September 1979.

90. P. Hasenfratz and J. Kuti, Phys. Reports C, 75 (1978). For a newer review, see R. Jaffe, MIT preprint MIT-CTP-814 (1979).

91. A. Chodos, R. Jaffe, K. Johnson, C. Thorne, and V. Weisskopf, Phys. Rev. D9, 3471 (1974); D10, 2599 (1974).

92. C. Callan, R. Dashen, and D. Gross, op. cit., Ref. 35; also Lectures at La Jolla Institute Workshop on Particle Theory, August 1978 (Princeton preprint), and Phys. Rev. D19, 1826 (1979).

93. H. Nielsen and P. Olesen, Nucl. Phys. B61, 45 (1972).

94. For a review, see S. Mandelstam, UC-Berkeley preprint UCB-PTH-79/9 (1979), to be published in the Proceedings of the 9th International Symposium on Lepton and Photon Interactions at High Energy, Fermilab, Batavia, IL (1979).

95. R. Anishetty, M. Baker, S. Kim, J. Ball, and F. Zachariasen, Phys. Lett. 86B, 52 (1979); M. Baker, J. Lucht, P. Lucht, and F. Zachariasen, Caltech preprint CALT-68-741 (1979); U. Bar-Gadda, SLAC preprint SLAC-PUB-2347; S. Mandelstam, Ref. 94 and UC-Berkeley preprint UCB-PTH 79/8.

96. Y. Nambu, Proceedings of the 19th International Conference on High Energy Physics, Tokyo, Japan (1978), p. 971, and references therein.

97. A recent review is given by J. Pati, University of Maryland preprint MDDP-TR-79-066 (1978), to be published in Seoul Symposium on Elementary Particle Physics, Seoul, Korea (1978).

98. L. B. Okun, M. Voloshin, and V. Zakharov, Moscow preprint ITEP-79 (1979).

99. Our approach roughly follows that of Jackiw and Rebbi (Ref. 32) and of K. Bitar and S. J. Chang [Phys. Rev. D17, 486 (1978); D18 435 (1978)].

100. Again, we recommend the Erice Lectures of S. Coleman (Ref. 36) as an excellent introduction.

101. J. H. Van Vleck, Proc. Nat. Acad. Sci. 14, 178 (1928).