PHYSICS REPORTS (Review Section of Physics Letters) 75, No. 4 (1981) 205–286. North-Holland Publishing Company

THEORIES OF QUARK CONFINEMENT*

Myron BANDER

Department of Physics, University of California, Irvine, California 92717, U.S.A.

Received 12 November 1980

"And since these things are so,
it is necessary to think that in
all the objects that are compound
there existed many things of all
sorts, and germs of all objects,
having all sorts of forms and
colors and flavors."

– Anaxagoras

Contents:

* Supported in part by the National Science Foundation U.C.I. Technical Report 80-12.

0 370-1573/81/0000–0000/$16.40

THEORIES OF QUARK CONFINEMENT

Myron BANDER

Department of Physics, University of California, Irvine, California 92717, U.S.A.

NORTH-HOLLAND PUBLISHING COMPANY - AMSTERDAM

Abstract:

It is conjectured that a non-Abelian gauge theory based on the color SU(3) group will confine quarks. Various techniques that have been applied to this question are reviewed. These include approximate methods based on strong coupling expansions of Hamiltonian and Euclidian lattice theories, instanton improvements on perturbation theory, and solutions of truncated Dyson–Schwinger equations for the gauge field propagator. Formal results based on electric-magnetic duality arguments and on the study of loop field theories are presented. Deconfinement at high temperatures, the inclusion of light quarks, and a possible reconciliation with a hypothetical discovery of free quarks are discussed.

1. Introduction

The concept that elementary particles are composed of fundamental objects with unusual charges called quarks dates back to 1964 [1.1, 1.2]. From the outset it was assumed that the quark was a mathematical concept and that free quarks did not have to exist. The observed hadrons were assigned to multiplets of flavor SU(3) transforming as singlets, octets, or decimets. Quarks were assigned to transform as the fundamental 3 and $\bar{3}$. This was subsequently extended to flavor SU(4) [1.3] and more recently to flavor SU(6).

However, more and more, the structure of elementary particles was being understood on the basis of this simple, even naive, model [1.4]. The subjects to which the quark model gave reasonable answers included hadron magnetic moments, cross sections, weak interactions, etc. It was reasonable to assume that hadrons were composites of quarks that had their own physical reality. It was not unreasonable to look for these quarks as on mass-shell real particles.

1.1. Free quark searches

In a review of recent free quark searches [1.5], we learn that of ten new experiments, eight give negative results. The latter indicate that quarks are not present in ordinary matter to parts in 10^{21}. For masses between 5 and 10 GeV they are not produced in high energy collisions, up to the highest accelerator energies available, to parts in 10^{11} [1.6]. The two experiments yielding possible positive signals [1.7] are either in contradiction with other similar experiments or have certain internal peculiarities. Even accepting some of these results as evidence for free quarks, these are still extremely rare. In the last section we will, briefly, discuss how the theories for permanent confinement could be modified to accommodate such rare quarks. For the present we will assume that quarks are permanently confined in hadrons.

1.2. Need for color

Several serious discrepancies and inconsistencies were evident in the naive quark model.

(i) Aside from the nonexistence of free quarks, only singlets, octets and decimets were observed. No triplets or sextets existed as physical particles.

(ii) The phenomenology of the low lying baryons favored placing the three quarks making up such a baryon into a totally symmetric state. This was a violation of the spin-statistics connection.

(iii) The simple quark model prediction of the value of the cross section for e^+–e^- annihilation into hadrons was off by a factor of about three.

(iv) There was no obvious way to introduce the strong interactions themselves into the model.

Historically, it was the second point that led to the concept of color [1.8]; it did help solve all the above problems. Each quark flavor comes in three "colors" transforming into each other under the color SU(3) group. The phenomenological postulate that encompasses confinement and item (i) above, is that *only color singlets exist as physical states.* Three quarks in a color singlet are totally antisymmetric, solving problem (ii), and the factor of three in the e^+–e^- annihilation into hadrons is taken care of by the three colors. The color quantum number provides a current which can couple to strong interactions. The flavor quantum numbers are more natural for the weak and electromagnetic interactions.

With the introduction of color, the concept of quark confinement takes on a precise form. It states that only color singlets belong to the physical Hilbert space. All nonsinglets have infinite energy.

1.3. Quantum chrodynamics

The idea that the strong interactions may be described by a non-Abelian gauge theory [1.9, 1.10] was given a sharp impetus by the observation that strong interactions are asymptotically free [1.11]. Bjorken scaling [1.12] and more generally the parton model [1.13], had to assume that for high energy inclusive reactions the constituents of the hadron behaved essentially as noninteracting particles and the discovery of asymptotic freedom provided a theoretical justification for these assumptions. The weakness of quark interactions at high energies has been successfully exploited in perturbative calculations [1.14]. After noting the decrease of the coupling constant at high energies, the idea that the interaction strength will increase at low energies or large distances, was put forward. It is hoped that interactions will then become sufficiently strong to confine quarks permanently.

1.4. Outline

This article will introduce various approaches and attempts to show that a non-Abelian, Yang–Mills gauge theory based on color SU(3) confines static quarks. The latter restricts the discussion to a pure gauge theory with no dynamical quarks, i.e. no quark pair production out of the vacuum. Quarks appear only as external sources. In the last section we will briefly discuss how such dynamical quarks modify the arguments for confinement. We limit our discussion to four dimensions. This review will concentrate primarily on work done after 1977. For earlier attempts, see the review by Marciano and Pagels [1.15]. One further approach we will not discuss is the $1/N_c$ expansion. This approximation has given insight into some of the phenomenology of hadron physics, as well as illuminating some difficulties in connection with apparently conserved currents; it has not, as yet, shed light on the question of confinement in four dimensions.

The approaches that will be discussed are:

(i) Lattice gauge theory, renormalization group applications to it, and numerical results (sections 3–5).

(ii) Improvements on perturbation theory based on the inclusion of instantons and on the application of Schwinger–Dyson equations (section 6).

(iii) Electric-magnetic duality (section 7).

(iv) Loop field theories (section 8).

Due to its own interest, and as it focuses on the important features of quark confinement, the deconfinement of quarks at high temperatures is discussed in section 9. The inclusion of dynamical light quarks is briefly presented in section 10, as is some speculation on how many of the confinement ideas can survive a discovery of real quarks. Section 2 is devoted to a short review of gauge theories. Several technical points and calculations are given in the appendices. These are mostly "well known" and are given for purposes of completeness.

2. Quantum Chromodynamics (QCD)

We assume that the strong interactions of quarks are mediated by a non-Abelian gauge theory based on a local SU(3) group. We will refer to this theory as QCD and to the vector fields on which this theory is based as gluons. This section serves as a review of the continuum, both quantum and classical, theory. As confinement is very strongly linked to gauge invariance, we discuss the interesting gauge invariant operators, and in section 2.4 present some physical and mathematical criteria for confinement.

2.1. Classical continuum QCD

We shall study a theory of a *static* triplet of quarks, described by fields $q^\alpha(x)$, $\alpha = 1, 2, 3$ and an octet of gauge vector gluons described by the field A^j_μ, $j = 1, \dots, 8$. It is convenient to combine this octet into a 3×3 matrix

$$A_\mu(x) = \tfrac{1}{2}\lambda^j A^j_\mu(x),$$

where the matrices $\lambda/2$ generate the algebra of SU(3) [2.1],

$$[\tfrac{1}{2}\lambda^i, \tfrac{1}{2}\lambda^j] = \mathrm{i} f^{ijk}\tfrac{1}{2}\lambda^k.$$

We define the covariant derivative of the quark fields and the field strength tensor by:

$$\begin{aligned} \mathrm{D}_\mu q^\alpha &= \partial_\mu q^\alpha - \mathrm{i}gA^{\alpha\beta}_\mu q^\beta \\ \mathrm{D}_\mu \bar{q}^\alpha &= \partial_\mu \bar{q}^\alpha + \mathrm{i}g\bar{q}^\beta A^{\beta\alpha}_\mu \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu - \mathrm{i}g[A_\mu, A_\nu]. \end{aligned} \tag{2.1}$$

The classical theory, whose quantum version we hope will result in confinement, is described by the Lagrangian

$$L[A, \bar{q}, q] = -\tfrac{1}{2}\mathrm{Tr}\, F_{\mu\nu}F^{\mu\nu} + \mathrm{i}\bar{q}\gamma^0 \mathrm{D}_0 q - m_0\bar{q}q. \tag{2.2}$$

The static nature of the quarks is insured by the absence of spacial derivatives acting on the quark fields. This Lagrangian is invariant under the local SU(3) gauge transformation described by the 3×3 matrix

$$\begin{aligned} &S(x)=\exp[\mathrm{i}\omega^j(x)\lambda^j/2] \\ &q(x)\to S^{-1}(x)q(x) \\ &A_\mu(x)\to S^{-1}A_\mu S+\frac{\mathrm{i}}{g}S^{-1}\partial_\mu S. \end{aligned}\tag{2.3}$$

In much of what we shall do we will need the Hamiltonian corresponding to the above Lagrangian. Due to the gauge invariance, not all the variables A_μ are independent, and we may avoid this redundancy by a proper choice of gauge. The most convenient will be the $A_0=0$ gauge. With

$$\begin{aligned} &E_i^\alpha=\partial_0 A_i^\alpha, \qquad E_i=\tfrac{1}{2}\lambda^\alpha E_i^\alpha \\ &B_i=\tfrac{1}{2}\varepsilon_{ijk}F_{jk}, \end{aligned}\tag{2.4}$$

the Hamiltonian is

$$H=\int \mathrm{d}^3x\,\{\mathrm{Tr}[E^2+B^2]+m_0\bar{q}q\}.\tag{2.5}$$

As the $A_0=0$ gauge is not a completely proper gauge, it must be supplemented with the auxiliary condition corresponding to Gauss' law

$$\boldsymbol{D}\cdot\boldsymbol{E}=\nabla\cdot\boldsymbol{E}-\mathrm{i}g[A_i,E_i]=-gq^+\frac{\lambda^\alpha}{2}q\frac{\lambda^\alpha}{2}.\tag{2.6}$$

2.2. Quantization of continuum QCD

The most straightforward method of quantizing a classical theory. is to identify the independent variables and their conjugate momenta and impose canonical commutation (or anticommutation) relations. One then searches for eigenstates of the relevant Hamiltonian. In the present case, the operators and their conjugate momenta are $A_i^\alpha(\boldsymbol{x})$ and $E_j^\beta(\boldsymbol{y})$ for the gauge fields and $q^\alpha(\boldsymbol{x})$, $q^{+\beta}(\boldsymbol{y})$ for the spinors:

$$\begin{aligned} &[E_j^\beta(\boldsymbol{y}),A_i^\alpha(\boldsymbol{x})]=-\mathrm{i}\delta_{ij}\,\delta^{\alpha\beta}\,\delta(\boldsymbol{x}-\boldsymbol{y}) \\ &\{q^{+\alpha}(\boldsymbol{x}),q^\beta(\boldsymbol{y})\}=\delta^{\alpha\beta}\,\delta(\boldsymbol{x}-\boldsymbol{y}). \end{aligned}\tag{2.7}$$

Gauss' law, eq. (2.6) is satisfied by requiring the physical states to form a sector of Hilbert space obeying

$$(\boldsymbol{D}\cdot\boldsymbol{E}+gq^+q)|\mathrm{physical}\rangle=0.\tag{2.8}$$

The operator $(\boldsymbol{D}\cdot\boldsymbol{E}+gq^+q)$ is the generator of infinitesimal gauge transformations. Equation (2.8) guarantees that the physical states are gauge invariant.

An equivalent way of quantizing a theory is through the Feynman functional integral method [2.2, 2.3]. This method yields directly the vacuum expectation values of time ordered products of operators:

$$\begin{aligned}\langle T(O_1(x_1)\cdots O_n(x_n))\rangle &= \frac{1}{Z}\int [\mathrm{d}A\,\mathrm{d}\bar q\,\mathrm{d}q]\,\delta[A_0^\alpha]\,O_1(x_1)\cdots O_n(x_n)\exp\Big\{\mathrm{i}\int \mathrm{d}^4x\,L[A,\bar q,q]\Big\}\\ Z &= \int [\mathrm{d}A\,\mathrm{d}\bar q\,\mathrm{d}q]\,\delta[A_0^\alpha]\exp\Big\{\mathrm{i}\int \mathrm{d}^4x\,L[A,\bar q,q]\Big\}.\end{aligned} \tag{2.9}$$

$L[A, \bar q, q]$ is the Lagrangian (2.2) treated as a classical functional of the commuting variables A_μ, and the anticommuting variables $\bar q$ and q. $[\mathrm{d}A\,\mathrm{d}\bar q\,\mathrm{d}q]$ denotes the functional integration.

In view of further developments, it is useful to transform the functional integrals from Minkowski to Euclidian space; we continue from real time t to $\mathrm{i}t$. In this language we evaluate the analytic continuation of the vacuum expectation discussed above.

$$\begin{aligned}\langle O_1(x_1)\cdots O_n(x_n)\rangle &= \frac{1}{Z}\int [\mathrm{d}A\,\mathrm{d}\bar q\,\mathrm{d}q]\,\delta[A_0^\alpha]\,O_1(x_1)\cdots O_n(x_n)\exp\Big\{-\int \mathrm{d}^4x\,L_\mathrm{E}[A,\bar q,q]\Big\}\\ Z &= \int [\mathrm{d}A\,\mathrm{d}q\,\mathrm{d}q]\,\delta[A_0^\alpha]\exp\Big\{-\int \mathrm{d}^4x\,L_\mathrm{E}[A,\bar q,q]\Big\}.\end{aligned} \tag{2.10}$$

In the above, all the geometry is to be treated as Euclidian, and the Euclidian Lagrangian is given by

$$L_\mathrm{E} = \tfrac{1}{2}\mathrm{Tr}\,F_{\mu\nu}F_{\mu\nu} + \mathrm{i}\bar q\gamma_0 \mathrm{D}_0 q - \mathrm{i}m_0\bar q q.$$

2.3. *Gauge invariant operators*

In the following we will attempt to build all our observables out of gauge invariant, not merely covariant, operators. Neither $q(x)$, $A_\mu(x)$ or even $F_{\mu\nu}$, satisfies this criterion. q^+q and $\mathrm{Tr}\,F^2$ do, but they form too restrictive a class. A much larger class can be built up with the help of the string operator (terminology to become clear further on)

$$U^{\beta\alpha}(y,x;c) = \Big[P_\mathrm{c}\exp\Big[\mathrm{i}g\int_x^y A(z)\cdot\mathrm{d}z\Big]\Big]^{\beta\alpha}. \tag{2.11}$$

The symbol P_c indicates that the exponential is to be path ordered along the curve c. Namely, let us divide the curve c into N segments bounded by z_i and z_{i+1}, with $z_0 = y$ and $z_N = x$. Then

$$\Big[P_\mathrm{c}\exp \mathrm{i}g\int_x^y A\cdot\mathrm{d}z\Big] = \lim_{N\to\infty}\Big[\prod_{i=0}^{N-1}\exp \mathrm{i}g(z_{i+1}-z_i)\cdot A(z_i)\Big].$$

Note that the definition of U is the same in Minkowski and Euclidian space. Under gauge transformations

$$U(y,x;c)\to S^{-1}(y)\,U(y,x;c)\,S(x).$$

Among the most useful *gauge invariant* operators will be the following two,

$$\begin{aligned} &M(y, x; \mathrm{c}) = \bar{q}^{\beta}(y)\, U^{\beta\alpha}(y, x; \mathrm{c})\, q^{\alpha}(x) \\ &W(\mathrm{c}) = \operatorname{tr} U(x, x; \mathrm{c}). \end{aligned} \tag{2.12}$$

$W(\mathrm{c})$ depends only on the curve c and not on the point x.

The operator M has an interpretation as a meson creation operator. A part of a meson wave function should contain a quark at the point y and an antiquark at the point x. Such a state may be created by the operator $q^{+\alpha}(y)\, q^{\alpha}(x)$ acting on the vacuum. This state is, however, not gauge invariant. We may remedy this by introducing a string of gauge fields between x and y. Such a state is obtained by letting the operator $M(y, x; \mathrm{c})$ act on the vacuum. Of course, a true meson wave function will consist of (aside from multi-quark configurations, in the case of light quarks) a superposition of quark and antiquark positions and, always, an even more complicated superposition of strings joining them.

2.4. *Confinement criteria*

The obvious question to ask with respect to confinement is, what is the energy of a state with a quark at $x = 0$ and an antiquark at $x = R$. Such a state will obey Gauss' law, eq. (2.8), or, equivalently, it should be gauge invariant. Based on the discussion above this state will be of the form

$$|q(0), \bar{q}(R)\rangle = \sum_{\mathrm{c}} \psi[\mathrm{c}]\, M(0, R; \mathrm{c})|0\rangle. \tag{2.13}$$

$M(0, R; \mathrm{c})$ is the meson creation operator of eq. (2.12), and the summation over curves c is only a symbolic way of indicating a superposition, with amplitudes $\psi[\mathrm{c}]$ of states with different curves. Let the energy of the lowest lying configuration be $E(R)$.

If there is no confinement, we expect for large R

$$E(R) \to 2m,$$

where m is the renormalized quark mass. Confinement implies that the interquark potential grows without bound. Although any growth consistent with $E(R) \to \infty$ yields confinement, we will illustrate this behavior, again anticipating future developments, with a linear growth.

$$E(R) \xrightarrow[R\to\infty]{} \sigma R. \tag{2.14}$$

σ is a universal tension, which can be determined from heavy quark spectroscopy [2.4],

$$\sigma = 0.2\ (\mathrm{GeV})^2. \tag{2.15}$$

If we could establish the behavior implied by eq. (2.14), we would prove confinement. σ will often be referred to as the string tension. The terminology will become clear subsequently.

Can we translate the previously described behaviors of $E(R)$ into properties of vacuum expectation values of some operators? This would prove useful in a path integral study of confinement. The overlap of two mesonic states at different Euclidian times is

$$M(T,R) = \langle M[(0,\mathbf{0}),(0,\boldsymbol{R});\mathrm{c}]\, M^{+}[(T,0),(T,\boldsymbol{R});\mathrm{c}]\rangle. \tag{2.16}$$

For c we take the straight line from $\mathbf{0}$ to $\boldsymbol{R}$. Placing a complete set of intermediate states between the operators, and remembering that we are working in Euclidian space, we find

$$M(T,R) = \sum_n |\langle 0|M[(0,\mathbf{0}),(0,\boldsymbol{R}),\mathrm{c}]|n\rangle|^2 \exp(-E_n T).$$

The smallest E_n corresponds to the energy of separation of a quark and an antiquark, i.e., the $E(R)$ of the previous paragraph. For large T the behavior of $M(T,R)$ is

$$M(T,R) \sim \mathrm{e}^{-TE(R)}. \tag{2.17}$$

We shall now study $M(T,R)$ directly:

$$M(T,R) = \langle \bar{q}(0,\mathbf{0})\, U((0,\mathbf{0}),(0,\boldsymbol{R});\mathrm{c})\, q(0,\boldsymbol{R})\, \bar{q}(T,\boldsymbol{R})\, U((T,\boldsymbol{R}),(T,\mathbf{0});\mathrm{c})\, q(T,\mathbf{0})\rangle. \tag{2.18}$$

The static quark fields satisfy the equation of motion (Euclidian)

$$\mathrm{i}\gamma^0(\partial_0 - \mathrm{i}gA_0)q = \mathrm{i}m_0 q.$$

Although in the $A_0 = 0$ gauge this is a free field equation, we can solve it in the presence of A_0 and as a result obtain a gauge independent confinement criterion. The quark propagator is

$$\begin{aligned}\langle q^\alpha(t,\boldsymbol{x})\,\bar{q}^\beta(t',\boldsymbol{x})\rangle &= \mathrm{P}\exp\left\{\mathrm{i}\int_{t'}^{t} \mathrm{d}t''\, A_0(t'',\boldsymbol{x})\right\}\langle q^\alpha(t,\boldsymbol{x})\,\bar{q}^\beta(t',\boldsymbol{x})\rangle_{\mathrm{free}} \\ &\sim \exp(-m_0|t-t'|)\, U((t,\boldsymbol{x}),(t',\boldsymbol{x}),\mathrm{c}')\,\delta^{\alpha\beta}\end{aligned} \tag{2.19}$$

where c′ is the line parallel to the t axis and stretching from $\boldsymbol{x}'$ to $\boldsymbol{x}$. Combining eq. (2.18) with eq. (2.19) we obtain

$$M(T,R) \sim \exp(-2m_0 T)\, W[\mathrm{c}].$$

$W[\mathrm{c}]$ is as defined in eq. (2.12), and the curve c is the rectangle depicted in fig. 2.1. Comparing this to eq. (2.18) we obtain

$$W[\mathrm{c}] \sim \exp\{-T[E(R) + 2m_0]\}.$$

If confinement holds, and the form of $E(R)$ discussed in eq. (2.14) is valid, we find

$$W[\mathrm{c}] \sim \mathrm{e}^{-\sigma TR} \sim \mathrm{e}^{-\sigma A[\mathrm{c}]}, \tag{2.20}$$

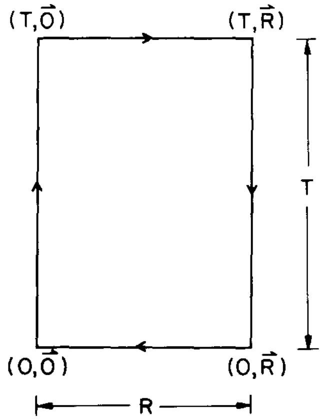


Fig. 2.1. The directed rectangular curve c used to determine the potential of a quark–antiquark pair separated a distance R apart.

where $A[c]$ is the area of the rectangle bounded by c. We note that the reference to quarks has disappeared from the condition on $W[c]$. Therefore we may define a confinement criterion for a pure gauge field. Generalizing eq. (2.20) to an arbitrarily large curve c, not just a rectangle, we say that the theory confines if

$$W[c] \sim e^{-\sigma A[c]}, \tag{2.21}$$

where $A[c]$ is the minimal area of some surface bounded by c [2.5]. If the theory does not confine, then the expected behavior of W is

$$W[c] \sim e^{-\mu P[c]}, \tag{2.22}$$

where $P[c]$ is the perimeter of c. Eq. (2.21) is the Wilson criterion [2.5] for confinement.

A physical picture of confinement is obtained by a study of the ways gauge invariance (Gauss' law) is implemented. For a quark–antiquark pair separated a distance R, Gauss' law requires a color electric flux between them. The question is to what extent does this flux spread in the transverse direction. In Abelian QED this spread is of the order of the separation leading to an electric field, E, of the order of $1/R^2$ and an energy in the field

$$\tfrac{1}{2}\int E^2\, d^3r \sim 1/R$$

yielding the usual Coulomb law. If the transverse spread can be fixed to be independent of the separation, the electric field will be constant and the energy of separation will be proportional to that separation, consistent with the previous pictures of confinement [2.6]. Squeezing color electric flux into tubes of fixed thickness (thick strings) is another way of insuring confinement.

3. Lattice gauge theories

A major impetus for the hope that gauge theories will yield confinement is the study of these theories in the strong coupling limit [2.5]. Results in this regime have been obtained in a lattice formulation of such field theories. In field theories a coupling constant depends on a length scale; thus, in order to distinguish strong from weak coupling we must introduce an intrinsic length into the theory. A lattice provides us with such a scale. We will first present Hamiltonian spatial lattice formulation and then a four dimensional Lagrangian lattice formalism.

3.1. Hamiltonian lattice formulation

As mentioned previously, one of the reasons for introducing a lattice is to enable a strong coupling expansion. With this in mind, it is useful first to rescale the variables so as to remove the coupling constant, g, from eqs. (2.1), (2.3) and (2.6); of course it is bound to reappear elsewhere. Let

$$\tilde{A}_i^a(x) = g\,A_i^a(x), \qquad \tilde{E}_i^a(x) = \frac{1}{g}E_i^a(x).$$

The new variable continues to satisfy the usual commutation relations. Some of the other relations are changed to

$$\begin{aligned} F_{\mu\nu} &= \frac{1}{g}(\partial_\mu\tilde{A}_\nu - \partial_\nu\tilde{A}_\mu - \mathrm{i}[\tilde{A}_\mu, \tilde{A}_\nu]) \\ \tilde{A}_\mu &\to S^{-1}\tilde{A}_\mu S + \mathrm{i}S^{-1}\partial_\mu S \\ \nabla\cdot\tilde{E} &- \mathrm{i}[\tilde{A}_i, \tilde{E}_i] = -q^+q \end{aligned} \tag{3.1}$$

and most important

$$H = \mathrm{Tr}\int \mathrm{d}^3x\,\left\{\left[g^2\tilde{E}^2 + \frac{1}{g^2}\tilde{B}^2\right] + m_0\bar{q}q\right\} \tag{3.2}$$

with

$$\tilde{B}_k = \tfrac{1}{2}g\epsilon_{ijk}F_{ij}. \tag{3.3}$$

Naively one could attempt a strong coupling expansion based on the Hamiltonian (3.2). For large g we might ignore the magnetic term and look for eigenstates of the electric field. However, nothing prevents $\tilde{B}^2$ from becoming very large, or $\tilde{E}^2$ very small. This is the reason we must turn to the lattice.

We discretize space into a simple cubic lattice with lattice spacing a. The lattice sites are vectors of the form

$$\tilde{\boldsymbol{x}} = a(i\hat{\boldsymbol{e}}_x + j\hat{\boldsymbol{e}}_y + k\hat{\boldsymbol{e}}_z).$$

i, j, k are integers and the $\hat{\boldsymbol{e}}_i$ are unit vectors along the lattice directions. The links joining lattice sites will play an important role; the *directed* link from $\boldsymbol{x}$ to $\boldsymbol{x} + a\hat{\boldsymbol{e}}$ will be denoted by $(\boldsymbol{x}, \hat{\boldsymbol{e}})$. We distinguish it from the oppositely directed link $(\boldsymbol{x} + a\hat{\boldsymbol{e}}, -\hat{\boldsymbol{e}})$.

Introducing a lattice forces us to give up various space symmetries. We shall, however, try to maintain gauge invariance. We wish the theory to be invariant under the discrete generalization of the quark field transformation law, eq. (2.3). It is natural to associate the vector gauge fields with the links. In a lattice, geometry scalars are associated with lattice sites, vectors with directed links, antisymmetric tensors with oriented areas, etc. Our basic unit will be the string operator (2.11) [2.5, 3.1]

$$U_{x;\hat{e}} = U(\boldsymbol{x} + a\hat{\boldsymbol{e}}, \boldsymbol{x}; \mathrm{c}) = \mathrm{P}\exp \mathrm{i}\left[\int_{x}^{x+a\hat{e}} \tilde{A} \cdot \mathrm{d}z\right]. \tag{3.4}$$

The integral runs along the link from $\boldsymbol{x}$ to $\boldsymbol{x} + a\hat{\boldsymbol{e}}$. An obvious property, due to the unitarity of U is

$$U^{-1}_{x;\hat{e}} = U_{x+a\hat{e};-\hat{e}}.$$

The continuum limit is achieved by letting a tend to zero, and then, at least on the classical level

$$U_{x;\hat{e}} \approx \exp\{\mathrm{i}a\tilde{\boldsymbol{A}}(\boldsymbol{x}) \cdot \hat{\boldsymbol{e}}\} \approx 1 + \mathrm{i}a\tilde{\boldsymbol{A}}(\boldsymbol{x}) \cdot \hat{\boldsymbol{e}}. \tag{3.5}$$

We shall use this limit to justify the lattice versions of various operators.

U is a unitary 3×3 matrix and thus can be parametrized as any element of SU_3, namely

$$U_{x,\hat{e}} = \exp\left\{\mathrm{i}\frac{\lambda^{\alpha}}{2} \cdot b^{\alpha}_{x,\hat{e}}\right\}. \tag{3.6}$$

Comparing with (3.5) we see that the group parameters b are related to the vector potentials A. It is important to note that the b's vary over a compact manifold in distinction to the continuum A's. Integration over the group denoted in the future by $[\mathrm{d}U]$ will, in reality, be an integration over the b's. We are now in a position to define a gauge invariant operator whose continuum limit will be related to $\tilde{\boldsymbol{B}}$ of (3.3). Consider the trace of a product of four link operators, $U_{x,\hat{e}}$, along a fundamental lattice square, or plaquette [3.2], illustrated in fig. 3.1

$$U_{x;\hat{e}_1,\hat{e}_2} = \{U_{x+a\hat{e}_2;-\hat{e}_2} U_{x+a(\hat{e}_1+\hat{e}_2);-\hat{e}_1} U_{x+a\hat{e}_1;\hat{e}_2} U_{x;\hat{e}_1}\}. \tag{3.7}$$

Using eq. (3.5) we combine the factors of $U_{x;\hat{e}}$ and to order a^2 obtain

$$\mathrm{Tr}\, U_{x;\hat{e}_i,\hat{e}_j} = \epsilon_{ijk}\, \mathrm{Tr}\exp[\mathrm{i}a^2\tilde{\boldsymbol{B}} \cdot \hat{\boldsymbol{e}}_k] \approx \epsilon_{ijk}\, \mathrm{Tr}[1 - \tfrac{1}{2}a^4(\tilde{\boldsymbol{B}} \cdot \hat{\boldsymbol{e}}_k)^2]. \tag{3.8}$$

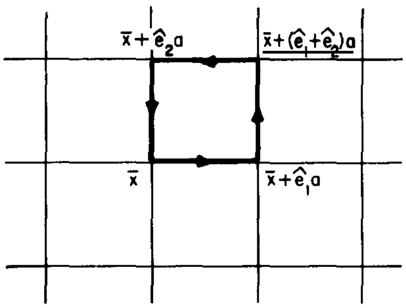


Fig. 3.1. A fundamental plaquette.

When no ambiguity occurs we replace the combination $x; \hat{e}_i, \hat{e}_j$ by p and let

$$U_p = U_{x;\hat{e}_i,\hat{e}_j}.$$

In view of eq. (3.8) a lattice version of

$$H_{\mathrm{M}} = \mathrm{Tr} \int \mathrm{d}^3x \, \frac{1}{g^2} B^2(x)$$

is

$$H_{\mathrm{M}} = \sum_p \frac{1}{g^2 a} \mathrm{Tr}[2 - U_p - U_p^+]. \tag{3.9}$$

We now need a lattice generalization of the electric field. The commutation relations of eq. (2.7) will serve as a guide. With $U_{x,\hat{e}}$ defined in eq. (3.4) we find

$$[E^\alpha(y) \cdot \hat{e}_i, U_{x;\hat{e}_j}] = \delta_{ij} \, \delta^2(y_\perp) \, U(x + a\hat{e}, y; \mathrm{c}) \frac{\lambda^\alpha}{a} \, U(y, x; \mathrm{c}),$$

if y coincides with any of the points in the interval x, $x + a\hat{e}$ and zero otherwise. In the lattice formation $U_{x,\hat{e}}$ has to be considered in total and cannot be broken up along the link. We can define an electric field at the start or end of a link. The two fields associated with the link $(x, \hat{e})$ are $E^\alpha_{x,\hat{e}}$ and $E^\alpha_{x+a\hat{e};-\hat{e}}$, and are determined from their commutation relations with link variables $U_{x,\hat{e}}$. For $\hat{e}$ positive, we postulate [3.1]

$$\begin{aligned}
&[E^\alpha_{x;\hat{e}_i}, U_{y,\hat{e}_j}] = \delta_{ij} \, \delta_{x,y} \, U_{x;\hat{e}_i} \frac{\lambda^\alpha}{2} \\
&[E^\alpha_{x+a\hat{e}_i;-\hat{e}_i}, U_{y;\hat{e}_j}] = -\delta_{ij} \, \delta_{x,y} \frac{\lambda^\alpha}{2} \, U_{x,\hat{e}_i}.
\end{aligned} \tag{3.10}$$

The fact that $U^+_{x;\hat{e}} = U_{x+a\hat{e};-\hat{e}}$ yields the commutation relations for negative $\hat{e}$'s. Use of the Jacobi identities yields the electric field commutation relations

$$\begin{aligned}
&[E^\alpha_{x;\hat{e}_i}, E^\beta_{y;\hat{e}_j}] = \mathrm{i}\delta_{ij} \, \delta_{x,y} f^{\alpha\beta\gamma} \, E^\gamma_{x;\hat{e}_i} \\
&[E^\alpha_{x;-\hat{e}_i}, E^\beta_{y;-\hat{e}_j}] = -\mathrm{i}\delta_{ij} \, \delta_{x,y} f^{\alpha\beta\gamma} \, E^\gamma_{x;-\hat{e}_i} \\
&[E^\alpha_{x;\hat{e}_i}, E^\beta_{y;-\hat{e}_j}] = 0.
\end{aligned} \tag{3.11}$$

The two electric fields are not independent of each other; the defining commutation relations provide the connection

$$\lambda^\alpha E^\alpha_{x+a\hat{e};-\hat{e}} = -U_{x;\hat{e}} \lambda^\beta E^\beta_{x;\hat{e}} U^+_{x;\hat{e}}$$

which implies that

$$E^2_{x;\hat{e}} = E^2_{x+a\hat{e};-\hat{e}}. \tag{3.12}$$

The electric energy

$$H_{\mathrm{E}} = g^2 \,\mathrm{Tr} \int \mathrm{d}^3x\, \tilde{E}^2 \tag{3.13}$$

goes over to

$$H_{\mathrm{E}} = \frac{g^2}{2a} \sum_{\text{links}} E^{\alpha}_{x;\hat{e}} E^{\alpha}_{x;\hat{e}}. \tag{3.14}$$

The total lattice QCD Hamiltonian is

$$H = \frac{g^2}{2a} \sum_{\text{links}} E^2_{x;\hat{e}} + \frac{\mathrm{Tr}}{ag^2} \sum_{p} (2 - U_p - U_p^+) + \sum_{x} m_0 \bar{q}_x q_x. \tag{3.15}$$

The generator of infinitesimal time independent gauge transformations is

$$\sum_{x} \Big\{ \sum_{\hat{e}} (E^{\alpha}_{x;\hat{e}} - E^{\alpha}_{x;-\hat{e}}) + q^+_x \frac{\lambda^{\alpha}}{2} q_x \Big\}. \tag{3.16}$$

We require this operator to annihilate all physical states.

Before proceeding, it is useful to study the spectrum of states of a single link Hamiltonian[1]

$$H_{x;\hat{e}} = \frac{g^2}{2a} E^{\alpha}_{x;\hat{e}} E^{\alpha}_{x;\hat{e}}.$$

For notational simplicity we write

$$E^{\alpha}_{x;\hat{e}} E^{\alpha}_{x;\hat{e}} = E^2$$
$$E^{\alpha}_{x;\hat{e}} = E^{\alpha}$$
$$E^{\alpha}_{x+a\hat{e};-\hat{e}} = E'^{\alpha}$$

E^2 is the quadratic Casimir operator, $C^{(r)}$ [3.3], for a representation r of SU(3). The states corresponding to a given r transform as an (r, r) representation of SU(3) × SU(3) generated by the operators E^{α} and E'^{α}. We will use the collective labels p and p'. If $d(r)$ is the dimensionality of r, we are dealing with $d^2(r)$ states. Aside from E^2 we may choose E^3, E^8, E'^3 and E'^8 to be diagonal

$$E^2|r; p, p'\rangle = C^{(2)}(r)|r; p, p'\rangle$$
$$E^{3,8}|r; p, p'\rangle = \lambda^{3,8}(p)|r; p, p'\rangle$$
$$E'^{3,8}|r; p, p'\rangle = \lambda^{3,8}(p')|r; p, p'\rangle.$$

[1] Such a single link theory has an interesting mechanical interpretation. Had we been dealing with an SO(3) rather than an SU(3) theory, the single link Hamiltonian would be the same as the Hamiltonian for a symmetric top with moment of inertia a/g^2. The operators $E^{\alpha}_{x;\hat{e}}$ generate rotations around space fixed axes and the $E^{\alpha}_{x+a\hat{e};-\hat{e}}$ around body fixed axes. The matrix $U_{x;\hat{e}}$ is the rotation necessary to go from one set to the other. (See, for example, L.D. Landau and E.M. Lifshitz, Quantum Mechanics (Addison-Wesley, Reading, MA, 1958).) Thus, we may view the lattice gauge theory as a collection of interacting rigid SU(3) tops.

Another useful representation is one in which the operator $U_{x;\hat{e}}(b)$ (cf. eq. (3.6)) is diagonal. The overlap between this and the previous set is

$$\langle U(b)|r; p, p'\rangle = D^{(r)}_{p,p'}(U). \tag{3.17}$$

$D^{(r)}_{p',p}$ is the $d(r)$ dimensional representation of SU(3) corresponding to the transformation generated by the b^α (cf. eq. (3.6)). The defining link variables U correspond to $D^{(3)}$.

As an example, and for future use, we will now calculate, in strong coupling perturbation, the low lying spectrum of states.

3.2. Strong coupling Hamiltonian perturbation

The Hamiltonian of eq. (3.15), in a form convenient for strong coupling expansion is [3.4, 3.5]

$$H = \frac{g^2}{2a}\Big\{\sum_{\text{links}} E^2_{x;\hat{e}} + x\,\mathrm{Tr}\sum_p (2 - U_p - U_p^+)\Big\} + \sum_x m_0\,\bar{q}(x)\,q(x) \tag{3.18}$$
$$x = 2/g^4.$$

In this section we will calculate, to zeroth, first and second order in x, the quark–antiquark potential. We will do this in a standard Hamiltonian perturbation theory. However, as we will be interested in the difference in the energies of this configuration and in the vacuum energy, let us calculate the quarkless spectrum first.

Due to Gauss' law, electric flux lines must close. To zeroth order in x the two lowest lying states are the vacuum, with no electric flux anywhere, and the states where the electric flux is in a 3 or $\bar{3}$ representation along the links of some fundamental plaquette. The first state we identify by $|0\rangle$ with energy $E_0^{(0)} = 0$. The second class of states consists of

$$|\mathrm{P}\rangle = U_p|0\rangle$$
$$|\bar{\mathrm{P}}\rangle = U_p^+|0\rangle. \tag{3.19}$$

Using the definition of U, eq. (3.7), and the commutation relations, eq. (3.10), we obtain

$$H_\mathrm{E}|\mathrm{P}\rangle = E_p^{(0)}|\mathrm{P}\rangle$$
$$H_\mathrm{E}|\bar{\mathrm{P}}\rangle = E_p^{(0)}|\bar{\mathrm{P}}\rangle$$

with

$$E_p^{(0)} = \frac{g^2}{2a}\{4C^{(3)}\} = \frac{16}{3}\frac{g^2}{2a}. \tag{3.20}$$

Normalizing the group invariant measure to [3.3]

$$\int \mathrm{d}U\, D^{*\,(r)}_{pq}(b)\, D^{(r')}_{p'q'}(b) = \frac{1}{d(r)}\,\delta_{rr'}\,\delta_{pp'}\,\delta_{qq'}$$

results in having the states $|\mathrm{P}\rangle$ properly normalized.

To first order in x all energies shift by

$$E^{(1)}_{0,p} = 3x \frac{g^2}{2a} N(\mathrm{P}). \tag{3.21}$$

$N(\mathrm{P})$ is the number of plaquettes in our system; we assume for the moment, a finite world. The order x^2 correction is obtained from second order perturbation theory,

$$E^{(2)}_0 = -\frac{g^2}{2a} x^2 \sum_p \frac{|\langle 0|U^+_p|\mathrm{P}\rangle|^2 + |\langle 0|U_p|\bar{\mathrm{P}}\rangle|^2}{(16/3)} = -\frac{g^2}{2a}\left(\frac{3}{8} N(\mathrm{P})\right) x^2. \tag{3.22}$$

Let us now turn to states with a heavy quark at the origin and a heavy antiquark at $R = na\hat{e}_1$. Again, due to Gauss' law, we must have an electric flux joining the two particles. We expect the lowest energy configuration to consist of the shortest flux path possible. This is illustrated in fig. 3.2(a). The lowest order energy of this configuration is

$$E^{0\mathrm{a}}(x) = \frac{g^2}{2a}\frac{4}{3} n + 2m_0 = \frac{g^2}{2a^2}\frac{4}{3} R + 2m_0. \tag{3.23}$$

Comparing the above equation with (2.14) we obtain the result alluded to earlier: *strong coupling lattice gauge theories confine quarks.*

To first order in x the perturbing Hamiltonian connects the state described by fig. 3.2(a) to those represented by figs. 3.2(b)–(f). The relevant Casimir operators are [3.3]

$$C^{(6)} = 10/3, \qquad C^{(8)} = 3$$

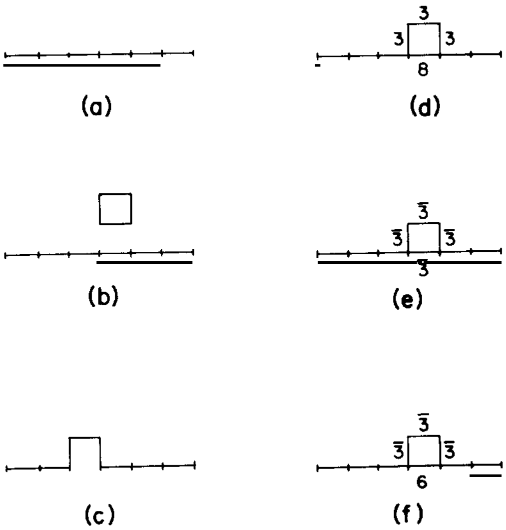


Fig. 3.2. Hamiltonian strong coupling contributions to the interquark potential. Numbers indicate the representation to which a particular link has been excited. If no number appears the 3 representation is implied. Zeroth order (a); first order (b); second order (c–f).

and

$$E^{0b} = E^{0a} + \frac{g^2}{2a}\left(\frac{16}{3}\right)$$

$$E^{0c} = E^{0a} + \frac{g^2}{2a}\left(\frac{8}{3}\right)$$

$$E^{0d} = E^{0a} + \frac{g^2}{2a}\left(\frac{17}{3}\right)$$

$$E^{0e} = E^{0a} + \frac{g^2}{2a}(4)$$

$$E^{0f} = E^{0a} + \frac{g^2}{2a}(6).$$

The first order shifts in the energies of all the configurations are again the common factor of eq. (3.21). The Clebsch–Gordon coefficients yield the matrix elements between configuration (a) and the other ones. The contribution of these configurations to the second order energy shifts are

$$\begin{aligned}
-E^{2b} &= \frac{g^2}{2a}x^2\frac{3}{8}[N(\mathrm{P}) - 4n] \\
-E^{2c} &= \frac{g^2}{2a}x^2\frac{1}{24}[4n] \\
-E^{2d} &= \frac{g^2}{2a}x^2\frac{8}{51}[4n] \\
-E^{2e} &= \frac{g^2}{2a}x^2\frac{1}{9}[4n] \\
-E^{2f} &= \frac{g^2}{2a}x^2\frac{1}{12}[4n].
\end{aligned} \tag{3.24}$$

The terms in the brackets indicate the number of configurations of each type. Combining these results, and subtracting the second order correction to the vacuum, eq. (3.22), we obtain [3.4, 3.5]

$$E(R) = \frac{g^2}{2a^2}R\left[\frac{4}{3} - x^2\frac{11}{153} + \cdots\right] + 2m_0. \tag{3.25}$$

Again we obtain a linearly rising potential leading to a confinement of quarks. Higher order terms will be presented later.

3.3. Euclidian lattice gauge theories

Lagrangian theories in a path integral formalism can also be transcribed to a lattice, this time a four-dimensional one [2.5, 3.2]. In appendix A we show that the lattice analogue of eq. (2.10) is

$$\left\langle \prod O_i(x_i;\hat{e}_i)\right\rangle = \frac{1}{Z}\int \prod_{x;\hat{e}} \mathrm{d}U_{x;\hat{e}} \prod O_i(x_i;\hat{e}_i) \exp\left\{\frac{1}{g^2}\operatorname{Tr}\sum_p (U_p + U_p^+)\right\}$$
$$Z = \int \prod_{x;\hat{e}} \mathrm{d}U_{x;\hat{e}} \exp\left\{\frac{1}{g^2}\operatorname{Tr}\sum_p (U_p + U_p^+)\right\}. \tag{3.26}$$

The plaquette summation now runs over all four dimensions. The continuum limit of the above is just eq. (2.10); however, as discussed in appendix A, this formulation is equivalent to the Hamiltonian formulation *for small g only*. We expect the quantum mechanics based on the different formulations to be the same in the continuum limit. However, various approximate results will be formalism dependent and it is constructive to compare them.

The operator characterizing the confinement properties of the theory is the Wilson loop integral, eq. (2.12), whose lattice analogue is

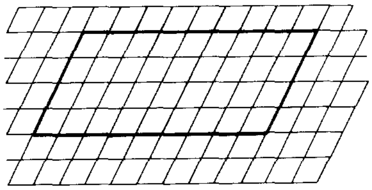
$$W[\mathrm{c}] = \operatorname{Tr} \prod_{(x,\hat{e})\in \mathrm{c}} U_{x,\hat{e}}. \tag{3.27}$$

The product is ordered along the closed curve c. For simplicity we will deal with planar curves only.

3.4. Euclidian lattice strong coupling expansion

The strong coupling expansion of the Euclidian theory [3.2, 3.4, 3.6] is obtained most directly by expanding the exponent in eq. (3.26) as a power series in $1/g^2$. The orthogonality properties of the representation matrices U show that the lowest order nonvanishing contribution to $\langle W[\mathrm{c}]\rangle$ is of the order $(1/g^2)^{N(\mathrm{c})}$. $N(\mathrm{c})$ is the number of plaquettes in the planar area surrounded by the curve c. The situation is illustrated in fig. 3.3(a). Relating the number of plaquettes, $N(\mathrm{c})$, to the area by $A = N(\mathrm{c})a^2$, we obtain

$$W[\mathrm{c}] = \exp(-\ln g^2 A/a^2). \tag{3.28}$$



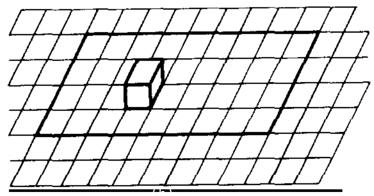


Fig. 3.3. Plaquettes contributing to the Wilson loop in a strong coupling Lagrangian theory. Lowest order (a); first order correction (b).

Comparing this to (2.21) we see that again the strong coupling limit results in confinement with the energy of separation

$$E(R)=\frac{1}{a^2}\ln g^2R. \tag{3.29}$$

The difference between this result and that of eq. (3.23) is due to the aforementioned difference (discussed in greater detail in appendix A) between the two formulations.

One could proceed and pick out the higher order terms contributing to $W[c]$. It is more convenient to modify the expansion somewhat [3.4]. This modification is inspired by the treatment of similar problems in statistical mechanics. Noting that the representations of a group provide a complete set of functions [3.3], we expand

$$\exp\frac{\mathrm{Tr}}{g^2}[U_p+U_p^+]=I(g^2)\sum_r \omega_r(g^2)\,\mathrm{Tr}\,D_p^{(r)} \tag{3.30}$$

with $\omega_1(g^2)=1$. Inverting this expansion yields

$$\begin{aligned}I(g^2)&=\int \mathrm{d}U_p\exp\frac{\mathrm{Tr}}{g^2}[U_p+U_p^+]=1+\mathrm{O}(1/g^2)\\ \omega_3(g^2)&=\int \mathrm{d}U_p\,\mathrm{Tr}\,D_p^{+(3)}\exp\frac{\mathrm{Tr}}{g^2}[U_p+U_p^+]=\frac{1}{g^2}+\mathrm{O}(1/g^2).\end{aligned} \tag{3.31}$$

$\omega_r(g^2)$ start with a power $(1/g^2)^\nu$, where ν is the smallest number of 3's and $\bar{3}$'s necessary to obtain the representation r. Thus an expansion in powers of $1/g^2$ may be recast into an expansion in powers of $\omega=\omega_3/3$. The first contribution comes from the $r=3$ term in eq. (3.30) for each plaquette outlined in fig. 3.3(a)

$$W[\mathrm{c}]=(\omega)^{A/a^2}. \tag{3.32}$$

The factors proportional to $I(g^2)$ are cancelled by similar terms in the expansion of Z. The first correction comes from surfaces depicted in fig. 3.3(b). There are four more plaquettes and the contribution is $(\omega)^{A/a^2+4}$. There are A/a^2 places where this extra box may be attached and four ways of arranging it in the two extra dimensions.

$$W[\mathrm{c}]=\omega^{A/a^2}\left(1+\frac{4A}{a^2}\omega^4+\cdots\right)=\exp\left\{-\frac{A}{a^2}(-\ln\omega-4\omega^4+\cdots)\right\}. \tag{3.33}$$

Again, to this order we have confinement with the energy of separation

$$E(R)=\frac{1}{a^2}(-\ln\omega-4\omega^4+\cdots)R. \tag{3.34}$$

Higher order terms will likewise be presented later.

Lattice theories, both in the Hamiltonian and in the Euclidian–Lagrangian formulation lead to quark confinement at strong couplings. The crucial question to which we turn in section 4 is how these results extrapolate to the weak coupling continuum region.

3.5. *Off axis calculations and string fluctuations*

The calculations in sections 3.3 and 3.4 were done for the case of the quark–antiquark pair along a lattice direction. If the lattice theory possesses a Lorentz invariant continuum limit, then the results should not depend on this choice. There is interest in doing these calculations with the quarks not along a lattice direction. We shall discuss the detailed reasons in section 5 and present the first order result for the quark–antiquark pair off axis, but still in a plane. Results for the three dimensional case are not as yet available; for the ideas we wish to illustrate, the planar case is sufficient.

The situation is illustrated in fig. 3.4. X and Y are in units of the lattice spacing. In this figure we also show several of the possible shortest flux lines. There are, in total, $(X+Y)!/X!Y!$ such flux configurations. The electric energy associated with any of these is

$$E_0 = \frac{g^2}{2a}\frac{4}{3}(X+Y). \tag{3.35}$$

To calculate the corrections due to the magnetic term we must use degenerate perturbation theory. The perturbing Hamiltonian connects paths which differ by one plaquette, such as those marked (a) and (b) in fig. 3.4. The matrix element of this perturbation is

$$-\alpha = -\frac{g^2}{2a}\frac{x}{81}. \tag{3.36}$$

The paths have $N = X + Y$ links. Each link is either horizontal to the right or vertical up. In order to diagonalize the perturbing Hamiltonian, we make a correspondence between the paths and a gas of fermions on N sites [3.7]. If link j is horizontal, we say site j is unoccupied; if link j is vertical then site j contains a fermion. Introducing fermion creation and annihilation operators a_j, a_j^+, we can write the perturbing Hamiltonian as

$$H_1 = -\alpha \sum_{j=1}^{N-1} [a_j^+ a_{j+1} + a_{j+1}^+ a_j]. \tag{3.37}$$

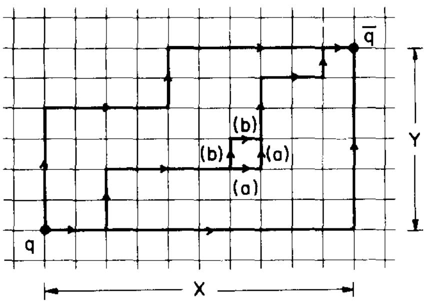


Fig. 3.4. Minimum flux paths connecting an off axis quark pair. The magnetic term in the Hamiltonian connects paths (a) and (b).

The total number of fermions is Y. This Hamiltonian is easily diagonalized. With the momentum modes defined by

$$a(k) = \sqrt{\frac{2}{N}} \sum_{j=1}^{N-1} a_j \sin(kj)$$
$$k = n\pi/N; \qquad n = 1, \ldots, N-1 \tag{3.38}$$

the Hamiltonian becomes

$$H_1 = -2\alpha \sum_k (\cos k)\, a^+(k)\, a(k). \tag{3.39}$$

The ground state consists of a Fermi sea with Y levels filled

$$E_1 = -2\alpha \sum_{n=1}^{Y} \cos(\pi m/N). \tag{3.40}$$

In the limit of large separation the energy of a quark–antiquark pair is

$$E = \frac{g^2}{2a}(X+Y)\left\{1 - \frac{2x}{81\pi}\sin[Y\pi/(X+Y)]\right\}. \tag{3.41}$$

We still have a confining potential, albeit angle dependent. As mentioned earlier, we hope that the angular dependence will disappear when the calculation is taken to all orders and the continuum limit approached.

Our primary interest is not in the energy, but in the fluctuations of the flux string. For the on axis calculations this string did not deviate from the shortest path by more than a few lattice steps. For the off axis case we find that already to first order the fluctuations can be as large as the string itself. For subsequent discussions, we will be interested in the mean square deviation of the flux string from the diagonal path. This will be a measure of how much the string flip-flops.

The distance a point on the string deviates from this diagonal is equal to the number of fermions to the right of it minus the number to the left of it:

$$\delta(l) = \sum_{j \geq l} a_j^+ a_j - \sum_{j<l} a_j^+ a_j. \tag{3.42}$$

In the ground state $\langle\delta(l)\rangle = 0$ but $\langle\delta^2(l)\rangle \neq 0$. The evaluation of this quantity can be accomplished by introducing particle-hole creation and annihilation operators:

$$b^+(k) = a(k); \qquad n \leq Y$$
$$a(k) = a(k); \qquad n > Y. \tag{3.43}$$

The annihilation operator becomes

$$a_j = \sqrt{\frac{2}{N}}\left\{\sum_{n=Y+1}^{N-1} \sin(2\pi nj/N)\, a(k) + \sum_{n=1}^{Y} \sin(2\pi nj/N)\, b^+(k)\right\}. \tag{3.44}$$

Both the a and b operators annihilate the ground state. The evaluation of $\langle\delta^2\rangle$ is now straightforward. We find that in the large N limit, with Y/N fixed,

$$\langle\delta^2\rangle \sim \ln N. \tag{3.45}$$

The importance of this result will be presented in section 5. We find that the fluctuations diverge as the separation increases. If this were true for the on axis situation, it would not become evident till at least the $\ln N$ order of perturbation theory.

What is the reason for these large fluctuations? Consider a quantized string under tension. It is an elementary exercise to show that the transverse fluctuations near the center of the string behave as (3.45) [3.8, 3.9]. This is true for continuous strings. One of the lessons we have learned is that low order perturbation calculations of an off axis quark configuration may be a better approximation to the continuum than on axis ones. Translated to the language of Wilson loops, these results indicate that in the continuum limit, the important surfaces bounding the Wilson loop will fluctuate more and more from the minimal one as the size of the loop increases. This phenomenon has an analogue in the statistical mechanics of spin systems, where it is referred to as surface roughening [3.10].

4. Renormalization group

It is tempting to be satisfied with the previous arguments and conclude that QCD confines quarks, perhaps only for sufficiently large coupling constants. Two related points indicate that we must do more. The strong coupling expansion on a lattice is very far from being Lorentz invariant. We must find a way to extend the previous discussions to such an invariant $a = 0$ continuum limit. Second, in this limit, the theory should go over to a perturbatively asymptotically free SU(3) gauge theory [1.11]. High energy phenomenology favors this version of QCD [1.14]. The standard perturbative limit of QCD *does not* confine quarks, so we cannot compare the strong coupling lattice σ (cf. eq. (2.14)) with a similar perturbative expression. However, there are quantities which exist in both limits and whose properties determine the answers to some of the points raised above. The renormalization group is the most convenient framework for studying the connection between strong and weak coupling.

The basis of renormalization group arguments is the observation that, ultimately, the lattice parameter a is an artifact. No physical quantities may depend on a, especially in the continuum limit of small a. This limit cannot be taken naively as we would lose any scale setting parameter; all dimensional quantities would be either zero or infinite. This is a common difficulty of all field theories with no important intrinsic mass parameters. The problem is cured by various renormalization prescriptions. These amount to taking some quantity that we hope has physical significance, and keeping it fixed as we let a approach zero. All other physical results are expressed in terms of this quantity. One of the crucial questions is whether the quantity that we keep fixed exists for all values of other parameters, as for example, the coupling constant g. If not, then the theory exists in different phases with vastly differing properties. For QCD to be both asymptotically free and confining it must exist in a single phase.

4.1. *Confinement phase*

In the strong coupling limit of lattice theories (cf. sections 3.2 and 3.3) the energy of separation of a quark–antiquark pair is proportional to the separation distance with a universal tension σ. Dimensional considerations lead us to the form

$$\sigma = \frac{1}{a^2} f(g). \tag{4.1}$$

In the sections 3.2 and 3.3 the first two terms in the strong coupling expansion of $f(g)$ were derived. The dependence of σ on a must be artificial and it should be possible to rescale the theory in such a way that the explicit dependence on a in (4.1) is compensated by an implicit dependence of g on a. Differentiating both sides with respect to a yields

$$0 = a\, \mathrm{d}\sigma/\mathrm{d}a = -2\sigma - \beta(g)\, \mathrm{d}\sigma/\mathrm{d}g$$

$$\beta(g) = -a\, \mathrm{d}g/\mathrm{d}a. \tag{4.2}$$

$\beta(g)$ is the Gell-Mann–Low–Callan–Symanzik [4.1, 4.2] function. It is this function that is expected to have both a strong coupling and a weak coupling limit. Rewriting (4.2), we obtain

$$\beta(g) = -2\sigma \Big/ \frac{\mathrm{d}\sigma}{\mathrm{d}g}. \tag{4.3}$$

A zero of β implies a zero of σ (barring the situation where $\mathrm{d}\sigma/\mathrm{d}g = 0$) and at such a point, for finite g, σ cannot be held fixed to define the continuum limit of the theory.

The first two orders in the strong coupling Hamiltonian and Lagrangian theories yield [3.2, 3.4, 3.5]

$$-\beta(g)/g = \left\{1 - \frac{44}{51}\frac{1}{g^8}\right\}; \qquad \text{(Hamiltonian)}$$

$$= \frac{2\mathrm{d}(\ln g)}{\mathrm{d}\ln\omega}\{\ln\omega(1-16\omega^4)+4\omega^4\}; \qquad \text{(Lagrangian)}. \tag{4.4}$$

4.2. Perturbative gluon phase

In standard Feynman perturbation theory there is no linear term in the energy of separation of two quarks. We cannot compare the σ's directly. We must find some other quantity to keep fixed as a and g are varied. This yields a new $\beta(g)$ that we may try to compare to the one obtained in the strong coupling limit. The usual renormalization prescription fixes the three point gluon function at some value M of the external gluon masses, to $g_\mathrm{R}(M)$. One can then vary g and a keeping $g_\mathrm{R}(M)$ fixed and obtain a $\beta(g)$. This procedure is valid if $g_\mathrm{R}(M)$ is, even if only in principle, a measurable quantity. If confinement holds, then our experience with lattice theories indicates that the vacuum is a unique, gauge invariant state and gauge variant expectation values, as the three point function will not exist. We will choose a variant of this prescription by calculating, in perturbation theory, the energy of a static quark–antiquark pair at a distance R. The energy at some fixed distance R is measurable, whether confinement holds or not, and can be the quantity held fixed. The first two terms in a perturbative expansion of β are evaluated in appendix B and it is also shown that these terms are independent of renormalization prescription:

$$\beta(g)/g = -\beta_0 g^2 - \beta_1 g^4$$
$$\beta_0 = 11/16\pi^2; \qquad \beta_1 = 102/(16\pi^2)^2. \tag{4.5}$$

We can invert (4.2) combined with (4.5) to obtain

$$g^2 = \left\{-\beta_0 \ln(\Lambda^2 a^2) + \frac{\beta_1}{\beta_0} \ln[\ln(\Lambda^2 a^2)^{-1}] + \cdots\right\}^{-1}. \tag{4.6}$$

Λ is a normalizing mass. Once we pick a process and a renormalization prescription then such a parametrization can be compared to experiment. Likewise, a knowledge of the finite parts of the coupling constant renormalizations permits a comparison of calculations with different prescriptions. As shown in appendix B, different prescriptions amount to a different choice of Λ.

The phenomenological value of Λ to which we shall compare our calculations is the one defined by the modified minimal subtraction ($\overline{\mathrm{MS}}$) procedure [1.14, 4.3]. Several analyses [4.4] determine

$$\Lambda_{\overline{\mathrm{MS}}} = (300\text{–}500)\ \mathrm{MeV}. \tag{4.7}$$

There is a possibility that higher twist operators [4.5], implying power law corrections to scaling laws, could lower this value by up to a factor of two.

Can we relate these numbers to any other ones? If the theory continues to confine down to small g then we may evaluate σ. Integrating (4.3) with (4.5) as input we find

$$\sigma = \frac{A^2}{a^2} (\beta_0 g^2)^{-\beta_1/\beta_0^2} \exp\left[-\frac{1}{\beta_0 g^2}\right]. \tag{4.8}$$

A^2 is an, as yet undetermined, constant. Between (4.6) and (4.8) we can eliminate g and a and find

$$\sqrt{\sigma} = A\Lambda. \tag{4.9}$$

Both A and Λ are renormalization scheme dependent. We hope that the product is independent. Numerical determinations of A will be discussed in the next section.

4.3. Two scenarios

As was discussed earlier, the continuum limit is reached by letting a approach zero and varying g, while keeping some physical quantity fixed. From the definition of $\beta(g)$, eq. (4.2), g will decrease until it reaches a zero with negative slope of $\beta(g)$. Presumably the theory will be Lorentz invariant at this point. We note that $g = 0$ is a zero of $\beta(g)$. The question is whether, as we decrease g from large values, we encounter other zeroes. Two such scenarios are shown in fig. 4.1. The solid lines indicate the known weak coupling and strong coupling (both Lagrangian and Hamiltonian) results. The dashed-dotted line shows a possible confinement situation. The strong and weak coupling regions are in the same phase. The second scenario corresponds to the dotted line in fig. 4.1. The strong couplings continuum region is

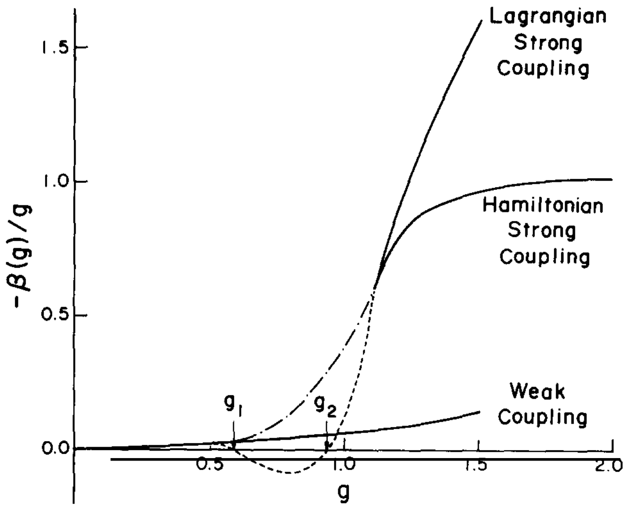


Fig. 4.1. Two scenarios for the QCD renormalization group β-function. The dashed curve refers to a nonconfining situation, while the dashed-dotted curve to a confining one.

given by some finite value of $g = g_2$, which has nothing to do with the asymptotic freedom region of $g < g_1$. Likewise the large distance limit of the asymptotically free theory corresponds to $g = g_1$ while in the strong coupling region g tends towards large values for increasing a. If the first interpretation is valid we obtain confinement.

It is useful to restate these two scenarios directly in terms of the tension $\sigma(g)$. If confinement holds we expect σ to exist down to small g where it is given by eq. (4.8). If confinement does not hold, σ will go to zero at $g = g_2$ and cease to exist for smaller g's. At large g it is given by the coefficient of R in eqs. (3.25) or (3.34).

If we can extend the calculation of σ to small values of g and it matches (4.8) we will be able to determine A and using (4.7) and (4.9) relate two known quantities.

4.4. Real space renormalization program

The renormalization group provides a scheme for directly computing the coupling constant as a function of length scale. The advantages of this procedure are that it is not tied to a specific renormalization prescription and in its approximate forms is less subject to numerical uncertainties encountered in other schemes. The idea for this is based on the block-spin renormalization program [4.6] used in statistical mechanics.

A group of lattice sites is blocked into a single site and new link variables connecting these enlarged sites are introduced. The procedure is outlined in fig. 4.2. The new link variables are functions of the original ones, the U's. Formally this prescription may be expressed by a rewriting of eq. (3.26)

$$Z = \int \prod \mathrm{d}V_l \prod \mathrm{d}U_l\, \delta[V_l - f_l[U]]\, \mathrm{e}^{-S[U]}. \tag{4.10}$$

$f_l[U]$ is a function defining the new link variables. Referring to fig. 4.2 this function depends on the old

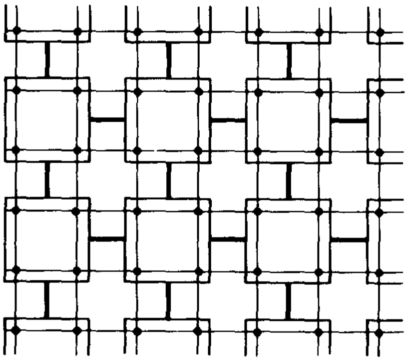


Fig. 4.2. A blocking of lattice sites and plaquettes into a new lattice. Dark lines indicate new links joining new sites made up of four old ones.

link variables in its vicinity. After integrating over the U's we are left with a lattice with twice the lattice spacing and 2^{-4} the number of link variables,

$$Z=\int\prod \mathrm{d}V_l \exp\{-S_{\mathrm{eff}}[V_l]\}. \tag{4.11}$$

S_{eff} is a new action in terms of the new variables. It will be more complicated in as much as it will involve, in addition to the single plaquette terms, traces of loops surrounding two plaquettes, in a plane and out of the plane, etc. As far as the physics of large distances is concerned, $S_{\mathrm{eff}}[V]$ and $S[U]$ should have the same form with only a different coupling constant g as the lattice spacing changes from a to $2a$. In this manner we can map out $g(a)$ and establish whether the strong and weak coupling regions go smoothly into each other.

In practice [4.7], a choice of f is difficult due to the requirement that V be a unitary matrix as U was. We will mention explicit calculations in the next section.

5. Numerical results

Two approximation schemes have been used to study lattice gauge theories. One is an extension to higher orders of the strong coupling expansion presented in section 3, both for the Hamiltonian and Lagrangian theories. In the second scheme, the world is restricted to a rather small lattice (10^4 lattice points being the largest considered till now) and a Monte Carlo lattice path integration is performed. A Monte Carlo integration can also be used to perform an approximate real space renormalization group evolution as described in section 4.4.

Unfortunately, due to computer limitations, many of these calculations have been for an SU(2) rather than an SU(3) theory. Qualitatively, we expect the same results for both theories. The major analytic differences for these two theories are the analogs of (3.25) and (4.5). For SU(2)

$$\begin{aligned} &E(R)=\frac{g^2}{2a^2}R\left[\frac{3}{4}-x^2+\cdots\right] \\ &\beta_0=\frac{22}{3}\frac{1}{16\pi^2}; \qquad \beta_1=\frac{136}{6}\frac{1}{(16\pi^2)^2}\,. \end{aligned} \tag{5.1}$$

5.1. Strong coupling expansion

The strong coupling expansion with low order terms given in eqs. (3.25) and (3.34) has been carried out to higher orders. We shall first discuss the Hamiltonian formulation. To sixth order [3.5] in x ($x = 2/g^2$) the string tension is

$$\sigma = \frac{g}{2a}\frac{4}{3}[1 - 0.054x^2 - 0.028x^3 - 0.0095x^4 + 0.0032x^5 - 0.0018x^6]. \tag{5.2}$$

Using (4.3) we find

$$\frac{-\beta(g)}{g} = 1 - 0.22x^2 - 0.17x^3 - 0.041x^4 + 0.025x^5 + 0.023x^6. \tag{5.3}$$

This series is plotted in fig. 5.1. We note that this series very quickly takes the strong coupling limit into the perturbative region, and to this order almost matches it. Above this weak coupling–strong coupling cross over point there is no hint of a zero in the β function and the confining scenario of fig. 4.1 is reproduced. Two observations are in order. First, the strong coupling and perturbative approximations come together in the small g region, where even the second order correction to β is small. In this region the perturbative β function is independent of the way the coupling constant is defined (see appendix B). The second point is of greater physical interest. The transition region between weak and strong coupling extends over a very small range of g, say for $1 < g < 1.5$. Using (4.6), this translates into a narrow range of momenta for a running coupling constant $g(q^2)$. QCD changes very rapidly from a low momentum, strong coupling, confining theory to a high momentum, perturbative regime. Such a trend is observed in nature. Physics below 2 GeV is dominated by resonances which, presumably, are due to the strong interquark potential. Above 2 GeV the scaling [1.12], parton [1.13], perturbative QCD region sets in; there is no large energy band where neither or both descriptions are necessary.

The string tension is plotted in fig. 5.2. The strong coupling expansion does break off from the strong limit and does show a tendency to approach the form suggested by (4.8). To each order, however, the curve rather than matching it, is at best tangent to this prediction. Note, there is no indication of a zero of σ. To each order we can determine an A; from the third to the sixth, A ranges between 300 and 170.

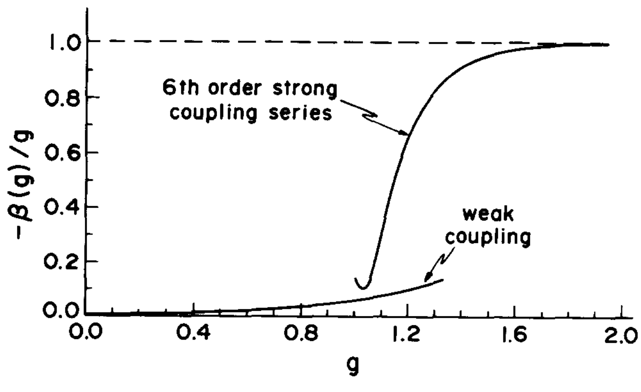


Fig. 5.1. Hamiltonian strong coupling β-function for an SU(3) gauge theory. (From ref. [3.5].)

Extrapolations to infinite order yield

$$A = 70 \pm 15. \tag{5.4}$$

From (4.9) and (2.15) we can determine Λ_{SL} [cf. appendix B]

$$\Lambda_{\mathrm{SL}} = (6.4 \pm 1.3)\,\mathrm{MeV}. \tag{5.5}$$

With the aid of eqs. (B.21), (B.22) and (B.23) we find the corresponding [cf. Note added in proof]

$$\Lambda_{\overline{\mathrm{MS}}} \approx (67 \pm 15)\,\mathrm{MeV}. \tag{5.6}$$

This value is considerably smaller than the one of eq. (4.7). However, we have not included quark fields in our calculations and these are expected to have a profound effect (see section 10.1), as is the inclusion of higher twist operators [cf. discussion following (4.7)]. It is compatible with Monte Carlo results to be discussed.

The strong coupling Euclidian formulation has been carried out to twelfth order in ω [3.6]. The first two terms were given in (3.34),

$$-\sigma a^2 = \ln \omega + 4\omega^4 + 12\omega^5 - 10\omega^6 - 36\omega^7 + 195.5\omega^8 + 113.1\omega^9 + 498.2\omega^{10} - 2548\omega^{11} + 4184\omega^{12}. \tag{5.7}$$

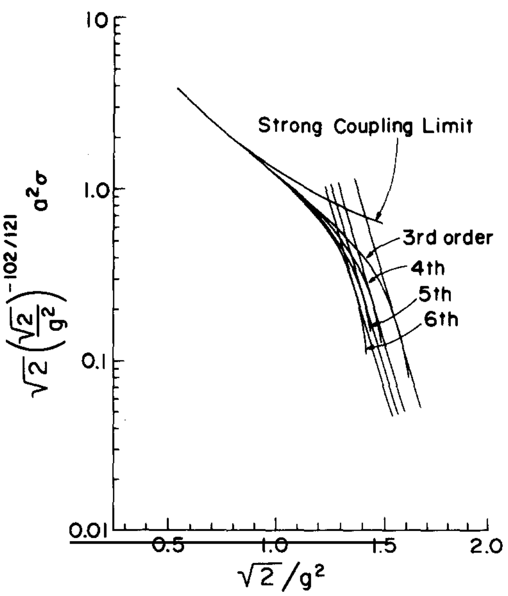


Fig. 5.2. Hamiltonian strong coupling string tension for an SU(3) theory. (From ref. [3.5].) For each order, the straight lines are best fits to eq. (4.8).

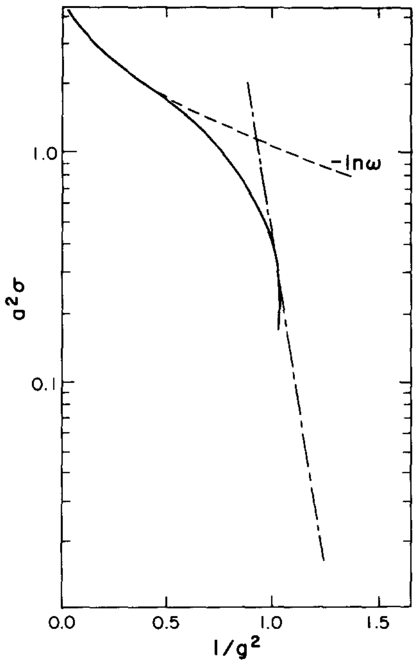


Fig. 5.3. Lagrangian strong coupling string tension for an SU(3) theory. (From ref. [3.6].) The short-long dashed line is a fit to eq. (4.8).

The plot of this result is shown in fig. 5.3 and the inferred value of Λ is

$$\Lambda_{\overline{\text{MS}}} \approx 55\,\text{MeV}, \tag{5.8}$$

in agreement with (5.6).[2]

The β function is similar to that of fig. 5.1.

Taken at face value these results indicate strongly that QCD confines; several caveats are in order.

(i) The obvious question is whether the order of the expansions are sufficiently large to infer such a result?

(ii) Is the discrepancy between $\Lambda_{\overline{\text{MS}}}$ calculated in this scheme, and the phenomenological value, curable by the inclusion of light quarks?

(iii) The curves for the string tension do not really show a straight line perturbative region.

(iv) There are questions regarding the fluctuations of the string presented in section 3.5 and discussed further in section 5.4.

5.2. Monte Carlo evaluations of the string tension

Recently, approximate evaluations of the lattice functional integrals, eq. (3.26), have been performed using Monte Carlo integration techniques [5.2, 5.3, 5.4]. A finite N^4 lattice system with periodic boundary conditions is set up. Practical limitations of even the fastest computers restrict N to be rather small. Most reliable results were obtained for a theory based on SU(2) rather than SU(3). Qualitatively, we expect theories based on either of these groups to behave in a similar way. If one confines, so should the other. For SU(2), results are available for up to a 10^4 lattice [5.3], while for SU(3) [5.4] 6^4 was the largest size, used with most results obtained on a 4^4 lattice[2']. We will not go into the details of this method; it is suited for evaluating averages, as for example the Wilson loop. The string tension, σ, may be determined in one of two ways. Let $W[I, J]$ be the value of a Wilson loop for a rectangle with sides of I and J lattice separations. Evaluating this for several I, J's, we may fit $W[I, J]$ to

$$W[I, J] = \exp[-A + 2B(I + J) + CIJ]. \tag{5.9}$$

A fit to C yields

$$\sigma = C/a^2. \tag{5.10}$$

In the second method, again based on (2.21), one defines

$$X[I, J] = -\ln\left\{\frac{W[I, J]\; W[I-1, J-1]}{W[I, J-1]\; W[I-1, J]}\right\}. \tag{5.11}$$

For large loops $X \to \sigma a^2$. Results will be presented using both techniques. The limitations of the lattice size limit the loops to 5×5 for SU(2) and 3×3 for SU(3).

Results for SU(2) are shown in fig. 5.4. The strong coupling predictions are followed out to about

[2] The two loop corrections to the β function have not been properly taken into account.

[2'] Calculations on 16^4 lattices exist for large, discrete, non-Abelian subgroups of SU(2). (G. Bhanot and C. Rebbi, CERN report TH2979 (1980).)

$g^2 = 2$ where the tension σ breaks and matches on to the form eq. (4.8). The slope provides the value of A

$$A_{SU(2)} = 77 \pm 12. \tag{5.12}$$

As this result is for an SU(2) theory we have no physical quantity to compare it to. We further note that the break between string coupling and weak coupling is fairly sharp. The transition region in g^2 is very limited. This is reminiscent of results found for SU(3) in the strong coupling expansion case, and will serve as a guide for SU(3) Monte Carlo calculations.

The results for SU(3) are far less reliable than those for SU(2). However, assuming that the theories based on these groups parallel each other, we may infer interesting results from the SU(3) calculations depicted in fig. 5.5:

$$A_{SU(3)} = 200 \pm 60. \tag{5.13}$$

The results in appendix B permit us to evaluate

$$\Lambda_{\overline{MS}} \approx (70 \pm 20)\,\mathrm{MeV}. \tag{5.14}$$

This value is compatible with (5.6) and (5.8). Its significance, as well as the significance of the sharp break (more inferred from the SU(2) calculations, than from the less accurate SU(3) one), was discussed in section 5.1.

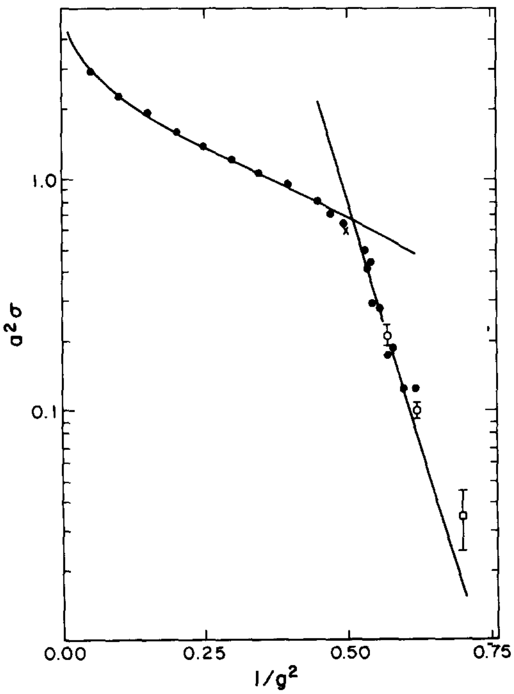


Fig. 5.4. Monte Carlo evaluation of the string tension for an SU(2) theory. (From refs. [5.3, 5.4].) The lines are fits to the Lagrangian strong coupling result and eq. (4.8).

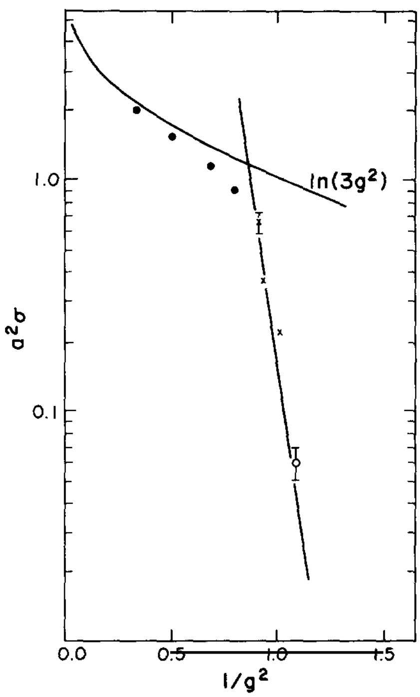


Fig. 5.5. Monte Carlo evaluation of the string tension for an SU(3) theory. (From ref. [5.4].) The lines are fits to the Lagrangian strong coupling result and eq. (4.8).

What are the shortcomings of this approach?

(i) Is even a 10^4 lattice, not to mention a 4^4 one, sufficient to infer the behavior of the theory in the large space-time limit?

(ii) Are the sizes of the Wilson loops large enough to distinguish area from perimeter laws?

(iii) Are these calculations sufficiently accurate, especially when σA is large so that the value of the Wilson loop is small? The statistical errors of the Monte Carlo calculations should be kept smaller than the loop expectation value itself.

(iv) What are the effects of uncontrolled fluctuations of the surfaces spanning the Wilson loop? This question is addressed in the next subsection.

A remark in support of these calculations (with a partial answer to the first two questions) is that there is a consistency in that the relevant distances are small over a large range in coupling constants. By relevant distance, in lattice spacings we understand $(\sigma a^2)^{-1}$. This is the only distance available and plays the role of the correlation length in statistical mechanical problems. Both for SU(2) and SU(3), this distance is of the order of one lattice separation at the transition from strong to weak coupling. Thus, small lattices and small loops may reveal the large structure. It is unlikely that we can believe these calculations when this length becomes of the order of the lattice size. For SU(2) this would correspond to $\sigma a^2 \sim 0.1$ while for SU(3) to much larger values of this parameter.

An attempt to avoid some of the difficulties discussed above is to give up the calculation of the string tension and perform a real space renormalization group calculation as described in section 4.4. The results should yield the coupling constant as a function of length scale directly. At present, the results are very preliminary [4.7]. A blocking algorithm has been devised for an SU(2) theory; it is not readily extendible to SU(3).

5.3. *String fluctuations-roughening*

We noted in section 3.5 that the flux string connecting quarks off a lattice direction fluctuates strongly, even in the lowest nontrivial order. The fluctuations in the middle of the string are of the order of the logarithm of the interquark separation. For the situation where the quarks are located on axis, such phenomena will not be seen in low orders of perturbation. For quark separation R/a, we would have to go to an order $\ln(R/a)$ in x to see this effect. Does this imply some sort of singular behavior for the on axis string tension which would impede a smooth continuation between strong and weak coupling regimes? We shall present evidence that there is a changeover from narrow flux strings to rapidly fluctuating ones at a finite g_R. This is similar to the roughening transitions of the statistical mechanics of spin systems and we borrow the language and say that the string becomes rough [3.10].

The definition of the transverse fluctuations is not unique. For the Hamiltonian theory a convenient procedure is the following [3.5, 3.7]. Consider the quarks separated along the z direction. A measure of the transverse fluctuations is

$$\delta = \langle E_z^2 r_\perp^2\rangle / \langle E_z^2\rangle . \tag{5.15}$$

E_z measures the color electric field on links parallel to the unperturbed string at a transverse distance $r_\perp^2 = x^2 + y^2$. A strong coupling series for the dimensionless quantity $\sigma\delta$ has been obtained [3.5]. Through order $x^{9/2}$ the terms are positive and the tendency for the series to diverge can be tested by looking for poles in Padé approximants. A consistent answer is obtained for the location of a pole at $x_R \approx 1.8$.

A Euclidian version of such calculations can be set up [3.8]. Consider a large planar curve c and a plaquette P parallel to the minimal surface spanning c and a transverse distance $x_\perp$ from it. We may define

$$E(x) = \frac{1}{2}\frac{\langle W[\mathrm{c}]\,U[\mathrm{P}]\rangle - \langle W[\mathrm{c}]\rangle\langle U[\mathrm{P}]\rangle}{\langle W[\mathrm{c}]\rangle} \tag{5.16}$$

as the quantity to insert onto eq. (5.15) for δ. Using this definition, a string coupling series was developed for a theory based on SU(2) and again an indication of a roughening transition was found in the neighborhood of the transition point.

Another calculation [3.9] addresses itself to the question of whether a given plaquette in the minimal area of a Wilson loop is spanned by the important surfaces contributing to W[c]. This is accomplished by reversing the sign of the contribution of the selected plaquette to the action. Let

$$p_{\mathrm{W}} = \frac{\langle W[\mathrm{c}]\exp\{-(2/g^2)\,\mathrm{Tr}(U_p + U_p^+)\}\rangle}{\langle W[\mathrm{c}]\rangle} \tag{5.17}$$

and

$$p_0 = \langle \exp\{-(2/g^2)\,\mathrm{Tr}(U_p + U_p^+)\}\rangle. \tag{5.18}$$

A roughening transition occurs when $p_{\mathrm{W}} = p_0$, i.e., when the surfaces spanning the Wilson loop fluctuate so rapidly that it does not affect a specific plaquette. Again, calculations have been carried out for a theory based on SU(2) and the results agree with the above.

What are the consequences of such a roughening transition? Some were discussed earlier, namely that a strong coupling expansion for the string tension may run into a block as it approaches the weak coupling region. It is unlikely that the string tension itself vanishes at that point. If the interpretation is that this roughening transition predicts an onset of unbounded fluctuations, then the results of section 3.5 indicate that a calculation with the quarks totally off axis (in all dimensions) may bypass this problem. Already in low orders the string is rough. The implications for the Monte Carlo calculations are unclear. Large fluctuations of the surface spanning the Wilson loops may invalidate finite lattice results as these fluctuations may see the walls of the lattice. If these fluctuations are only logarithmic then they may be sufficiently soft that even present results are valid. Work should be forthcoming elucidating both points.

6. Improvements on perturbation theory

It is rather obvious that standard perturbation theory will not result in confinement. The quark–antiquark potential is of the form

$$E(R) = -\frac{4}{3}\frac{1}{4\pi R}\,g^2(R) \tag{6.1}$$

with the zeroth and first order corrections to $g^2(R)$ given by eq. (B.4)

$$g^2(R) = g^2\left[1 + \frac{11}{16\pi^2} g^2 \ln \frac{R^2}{a^2} + \cdots\right]. \tag{6.2}$$

Higher order corrections contribute higher powers of logarithms and no finite order will result in a linear growth of the energy. It would be interesting to find improvements on perturbation theory that would yield confinement. First, it is an approach complementing the strong coupling one discussed previously. Second, we would always work with a continuum, Lorentz invariant theory. Finally, because of the previous point we could more readily develop a phenomenology, especially one involving light quarks.

Two methods have been used to study QCD in the small g region with a view of determining whether this theory does or does not confine. The first method involves perturbing the theory not only about the vacuum built around the $\boldsymbol{A} = 0$ solution, but likewise about nontrivial gauge transforms of this potential. In a Lagrangian language this amounts to performing functional integrals around instantons [6.1, 6.2]. In the second method one sets up an approximate set of Dyson–Schwinger equations and tries to solve them at large distances.

6.1. Instanton contributions to confinement

One way to view perturbation theory on a Hamiltonian, such as that of (2.5), is to search for a classical minimum of the potential, the magnetic term in this case, and expand the Hamiltonian around this minimum. To lowest order one has the familiar collection of harmonic oscillators. In the present case the obvious minimum is for the vector potential $\boldsymbol{A}^\alpha = 0$. The unperturbed Hamiltonian is

$$H = \tfrac{1}{2}\int d^3x\, [:(E):^2 + :(\nabla \times A)^2:]. \tag{6.3}$$

The ground state wave functional consists of Gaussian fluctuations around the minimum. Clearly, any gauge transformation of $\boldsymbol{A}^\alpha = 0$ will likewise be a minimum. The requirement that Gauss' law, eq. (2.8), be obeyed by the states of the theory eliminates most of this redundancy; namely, we do not have to worry about minima which differ from $\boldsymbol{A} = 0$ in finite regions of space. There are, however, classes of vector potentials, even though gauge equivalent to $\boldsymbol{A} = 0$ differ from it at a sufficiently large distance that they have to be taken into account separately. The form of these potentials is discussed in appendix C. These potentials may be labeled by a positive or negative integer ν. The trivial $\boldsymbol{A} = 0$ potential belongs to the $\nu = 0$ class. The Hamiltonian expanded around a potential of the νth class is

$$H^{(\nu)} = \tfrac{1}{2}\int d^3x\, [:(E)^2: + :(\nabla \times A - ig[A^{(\nu)}, A])^2:]. \tag{6.4}$$

Denote by $|\nu\rangle$ the ground state of the harmonic oscillations around the νth minimum. Due to quantum mechanical tunneling which connects states built around the different minima,

$$|\psi\rangle = \sum_{\nu=-\infty}^{\infty} |\nu\rangle \tag{6.5}$$

has a lower expectation value of the Hamiltonian and, therefore, is a better candidate for the ground state of the full theory. To find the energy corresponding to $|\psi\rangle$, we evaluate, for large T

$$e^{-E_0 T} = \frac{\langle\psi|e^{-HT}|\psi\rangle}{\langle\psi|\psi\rangle} = \sum_\nu \langle\nu|e^{-HT}|0\rangle. \tag{6.6}$$

The overlap matrix element $\langle\pm 1|e^{-HT}|0\rangle$ is proportional to the exponential of the action of a solution connecting the $\nu = \pm 1$ class to the $\nu = 0$ class, or looking at (C.8), to $\exp[-8\pi^2/g^2]$. The quantum mechanical zero point energy is very difficult to evaluate. The result is [6.3]

$$\begin{aligned} \langle\pm 1|e^{-HT}|0\rangle &= T\int d^3x \frac{d\rho}{\rho^5} D(\rho) \\ D(\rho) &= C[S_0(g(\rho))]^6 \exp - [S_0(g(\rho))]. \end{aligned} \tag{6.7}$$

The power of ρ in the above may be inferred from dimensional arguments. $S_0(g(\rho))$ is the instanton action, eq. (C.8), corrected for one loop renormalization effects. With the help of (4.6) we may rewrite the action to this order as

$$S_0(g(\rho)) = -11 \ln \Lambda\rho - \frac{51}{21} \ln(\ln \Lambda\rho). \tag{6.8}$$

Λ and the constant C in (6.7) depend on the details of the renormalization scheme.

Equation (6.7) may be viewed as a tunneling amplitude between two minima of a potential. However, tunneling can occur in a more complicated way by going back and forth, using instantons and anti-instantons in such a way that we get between the desired minima in the interval T. We are interested in the action of a configuration of several instantons and anti-instantons, and the Gaussian fluctuations around them. Although the action of a configuration corresponding to only instantons or anti-instantons is additive, the action involving both is more complicated. If the separation of an instanton and anti-instanton is less than their sizes, ρ, then the difference in the action from a sum is significant and difficult to handle. We shall, for the moment, ignore this point and later outline the region where this approximation is expected to be valid. We will also mention an approximation taking into account the above point.

With the above in mind, the overlap on the right hand side of (6.6) is taken to be

$$\langle\nu|e^{-HT}|0\rangle = \sum_{n_+ + n_- = \nu} \frac{T^{n_+ + n_-}}{n_+! n_-!} \left[\int d^3x \frac{d\rho}{\rho^5} D(\rho)\right]^{n_+ + n_-} = \frac{1}{\nu!}\left[2T\int d^3x \frac{d\rho}{\rho^5} D(\rho)\right]^\nu, \tag{6.9}$$

yielding

$$e^{-E_0 T} = \exp\left[2T\int d^3x \frac{d\rho}{\rho^5} D(\rho)\right]. \tag{6.10}$$

We see that the energy density of the ground state is lowered from the naive perturbative value by

$$\Delta\epsilon \doteq -2\int \frac{d\rho}{\rho^5} D(\rho). \tag{6.11}$$

As mentioned earlier, this approximation presupposes that $\rho < \Delta x$, where Δx is some average instanton anti-instanton separation. To take this into consideration we shall cut-off the ρ integration at some critical ρ_c. ρ_c may be estimated by requiring that the product of $\Delta\epsilon$, the volume of an instanton, and the Euclidian time during which it acts, is less than one. The precise assumption made in ref. [6.1] is

$$2\int_0^{\rho_c} \frac{d\rho}{\rho^5} D(\rho) \frac{\pi^2}{2} \rho^4 < 1. \tag{6.12}$$

Irrespective of many of the somewhat uncontrollable assumptions made along the way, we do have a better candidate for the vacuum state than the perturbative one. Within the same spirit can we find the energy of a static quark–antiquark pair separated by a distance R? To answer the question, we first consider the instanton variational calculation for the lowest energy in the presence of a given static external gauge field. Later we shall take the field to be the Coulomb gauge field due to the static quarks. The idea of separating fields into an instanton and an external field makes sense only if the instanton size is less than the distances over which the external field varies. In this case, we may define the instanton to be a solution of the classical equations of motion approaching the external field at large distances. In addition, if the external field is sufficiently weak the total action may be evaluated without too much difficulty. It consists of the action of a pure instanton, eq. (C.8), the action of the external field, $\frac{1}{4}T\int d^3x\,(F^{\text{ext}})^2$, and the interaction action. For an instanton located at x this term will be of the form

$$S_{\text{int}}(x) = \tfrac{1}{2}F^{\text{ext},a}_{\mu\nu}(x)\, F^{\text{inst},a}_{\mu\nu}(x)\, \tfrac{1}{2}c\, \pi^2\rho^4 \tag{6.13}$$

where c is of the order of unity. Detailed calculations [6.1, 6.2] give $c = 2$ and

$$S_{\text{int}}(x) = \frac{2\pi^2}{g(\rho)}\rho^2 R_{a\alpha}\, \eta_{\alpha\mu\nu}\, F^{\text{ext};a}_{\mu\nu}(x). \tag{6.14}$$

An analogous expression, holds for the anti-instanton interaction with the external field. It is easy to see that the analogue of (6.9) is

$$\begin{aligned}\langle\nu|e^{-HT}|0\rangle &= \sum_{n_+ + n_- = \nu} \frac{T^{n_+ + n_-}}{n_+!\,n_-!}\left\{\left[\int d^3x\, dR\, \frac{d\rho}{\rho^5} D(\rho) \exp\{S_{\text{int}}(x)\}\right]^{n_+}\right.\\ &\quad\times\left.\left[\int d^3x\, dR\, \frac{d\rho}{\rho^5} D(\rho)\exp\{\bar{S}_{\text{int}}(x)\}\right]^{n_-} \exp\left\{-\frac{T}{4}\int d^3x\, (F^{\text{ext},a}_{\mu\nu})^2\right\}\right\}\\ &= \frac{1}{\nu!}\left\{T\int d^3x\, dR\, \frac{d\rho}{\rho^5} D(\rho)[\exp\{S_{\text{int}}(x)\} + \exp\{\bar{S}_{\text{int}}(x)\}]\right\}^{\nu} \exp\left\{-\frac{T}{4}\int d^3x\,(F^{\text{ext}})^2\right\}. \end{aligned} \tag{6.15}$$

Thus, the energy of the lowest lying state consistent with a given external field is

$$E = -2\int^{\rho_c} d^3x \frac{d\rho}{\rho^5} D(\rho) - \int^{\rho_c} d^3x\, dR\, \frac{d\rho}{\rho^5} D(\rho)\,[\exp\{S_{\text{int}}(x)\} + \exp\{\bar{S}_{\text{int}}(x)\} - 2] + \tfrac{1}{2}\int d^3x\,(F^{\text{ext}})^2. \tag{6.16}$$

The first term is the instanton correction to the vacuum energy, eq. (6.12); the last term is the energy of the external field, while the middle term represents the external field-instanton interactions. We will assume that the external field is weak and expand the exponential in the middle term to second order. The first order term vanishes by the group integration. Using

$$\begin{aligned}&\int \mathrm{d}R\, R_{a\alpha} R_{a'\alpha'} = \tfrac{1}{8}\delta_{aa'}\delta_{\alpha\alpha'}\\&\eta_{\alpha\mu\nu}\eta_{\alpha'\mu'\nu'} = \delta_{\mu\mu'}\delta_{\nu\nu'} - \delta_{\mu\nu'}\delta_{\nu\mu'},\end{aligned} \tag{6.17}$$

we obtain

$$\begin{aligned}&E = E_{\mathrm{vac}} + \frac{1}{4\mu}\int \mathrm{d}^3x\, (F^{\mathrm{ext}})^2\\&\mu = 1/(1-\eta)\\&\eta = \frac{\pi^2}{2}\int^{\rho_c} \frac{\mathrm{d}\rho}{\rho} D(\rho)\, S(g(\rho)).\end{aligned} \tag{6.18}$$

An approximate treatment [6.1, 6.2] of the instanton anti-instanton interactions modifies the relation of μ to η

$$\mu = \eta + \sqrt{1+\eta^2}. \tag{6.19}$$

Equation (6.18) has a resemblance to the electrodynamics of dielectric media. This correspondence is even stronger when we note the relation between F^{ext} and external sources. In the case where F^{ext} is purely electric, we find, by varying (6.18) with respect to A_0

$$\nabla\cdot\frac{1}{\mu}\boldsymbol{E}^{\mathrm{ext},\alpha} = \rho^{\mathrm{ext},\alpha}. \tag{6.20}$$

In the case that the external source consists of a quark–antiquark pair, the analogs of (6.1) and (6.2) are

$$E(R) = -\frac{4}{3}\frac{1}{4\pi R}\tilde{g}^2(R)$$

with

$$\tilde{g}^2(R) = g^2(R)\,\mu[g(\rho)] \tag{6.21}$$

and μ given by (6.19); $\tilde{g}(R)$ is the coupling constant appropriate to the distance R. From $\tilde{g}(R)$ we can obtain the renormalization group function $\tilde{\beta}$. Before we can compare $\tilde{\beta}$ with other results, we must discuss one further restriction. The treatment of the interaction of an instanton with the external field presupposed instanton sizes smaller than the typical distance over which the external field varied. Though larger instantons may contribute to the vacuum energy, it is plausible that they do not see the external field at all and do not effect (6.21). In ref. [6.1] instanton sizes were integrated up ρ_c with $\frac{2}{3} < \rho_c/R < 1$. The other limit on ρ_c is (6.12) and these two limits set an upper bound on the separation for which this discussion is valid; this in turn puts an upper limit on g. In practice, the upper limit for which we may believe all those approximations is $g < 3$. In fig. 6.1 we plot $\tilde{\beta}(g)/g$ [6.2] evaluated from (6.21) superimposed on the perturbative and strong coupling approximation. The instanton contributions rapidly take the perturbative result towards the strong coupling region without any hint of additional zeroes, indicating that the confining phase and the perturbation phase are the same. The significance of the sharpness of the rise was discussed previously in section 4.[3]

Following the strong coupling results $\ln g^2(a)/a^2$ should, for large a, approach the string tension. Results (with a correct relation among the various Λ's) for a theory based on the group SU(2) are presented in fig. 6.2 [6.4]. The relation between the lattice cut-off and string tension is in agreement with strong coupling and Monte Carlo results, eq. (5.12), in that $\sqrt{\sigma} \approx 70\,\Lambda_{\mathrm{E.L.}}$.

We close this discussion with an intuitive picture of how, in this approach, instantons are responsible for confinement and include a critique of this technique. Equation (6.20) suggests that we define a color displacement field

$$\begin{aligned} \boldsymbol{D}^{\alpha}(\boldsymbol{x}) &= \frac{1}{\mu}\boldsymbol{E}^{\alpha}(\boldsymbol{x}) \\ \nabla \cdot \boldsymbol{D}^{\alpha} &= \rho^{\mathrm{ext},\alpha}. \end{aligned} \tag{6.22}$$

The energy stored in the color field is

$$E = \tfrac{1}{2}\mu \int D^{\alpha} D^{\alpha}\, \mathrm{d}^3x. \tag{6.23}$$

As μ becomes larger due to instanton effects, energy arguments would like D to decrease. D is, however, constrained by Gauss' law, eq. (6.22), tying it to the external sources; the flux of D through any surface between a quark–antiquark pair is proportional to the quark color electric charge. The configuration of minimum energy picks a tube of finite radius connecting the quarks and forces most of the field into it; likewise it expels instantons from the inside of the tube. Inside the tube $\mu \approx 1$ and outside $D \sim 0$. The change in energy from that of the vacuum occurs in a region whose volume is proportional to the interquark separation resulting in the normal confining potential. We shall return to this picture of flux tubes later on. Even if this approach is correct, the details and approximations are open to many questions.

1. Do other configurations besides instantons make a contribution to confinement?
2. Have instanton anti-instanton interactions been properly accounted for? Perhaps a correct treatment of instanton–instanton interactions caused by Gaussian fluctuations around the multi-instanton solutions [6.5] are as important as configurations with both instantons and anti-instantons.
3. Is the treatment of the size cut-off adequate?

6.2. Dyson–Schwinger equations

Earlier in this section we showed that confinement would not obtain in any finite order of

[3] The results for the instanton corrected β function may have some uncertainties as the relations between lattice and continuum Λ's used in this calculation is not correct (cf. appendix B). It is not expected to have a large effect on β [6.4].

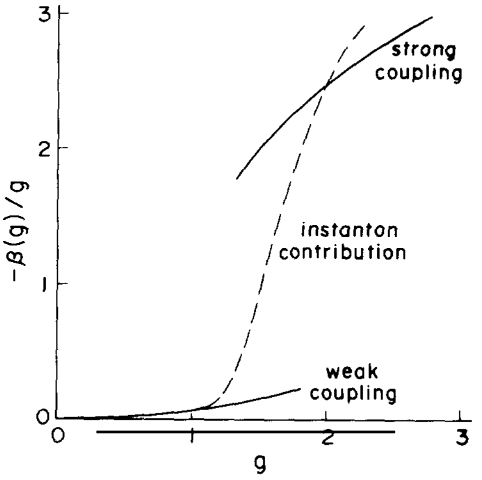


Fig. 6.1. Instanton interpolation between weak and strong coupling. (From ref. [6.2].)

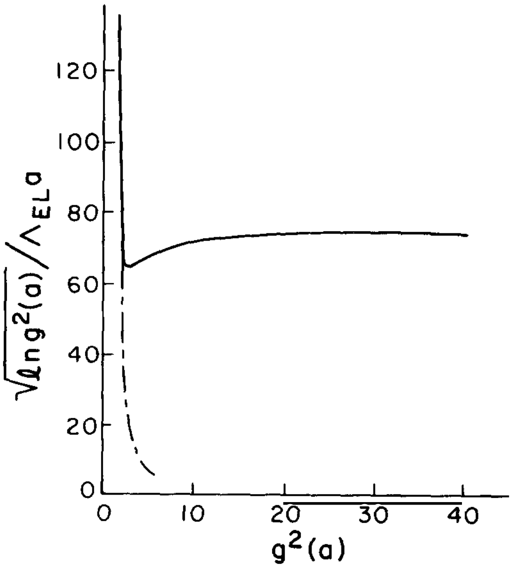


Fig. 6.2. Instanton evaluation of the string tension. The dashed-dotted line is a continuation of the asymptotic freedom result. [D. Gross, private communication.]

perturbation theory. Perhaps an appropriate summation of an infinite set of Feynman graphs could yield confinement? The first question we ask is for what quantity we should set up these diagrams. Ideally, it would be for the energy of two separated quark sources. We first show that this is approximately related to the gluon propagator. It is well known that, in low orders of perturbation theory, an interparticle potential is related to the propagator of an exchanged particle. We will establish this result, and obtain its limitations, for non-Abelian theories. The difference in energies between a state with a static source density $\rho(x)$ coupled to a gauge field and the vacuum energy of a pure gauge field is obtained from

$$Z[\rho, T] = \frac{\langle 0|\exp\left\{-T\left(H + g\int \mathrm{d}^3x\, \rho^\alpha(\bar{x})\, A_0^\alpha(\bar{x})\right)\right\}|0\rangle}{\langle 0|\mathrm{e}^{-TH}|0\rangle}. \tag{6.24}$$

For large T, this yields

$$-\frac{1}{T}\ln Z[\rho, T] = E[\rho] \tag{6.25}$$

with $E[\rho]$ the sought after energy. Another interpretation of (6.24) is that this expression is the generating functional for vacuum expectation values of products of the operator A_0 [2.2]. Specifically

$$\ln Z[\rho, T] = \tfrac{1}{2}\int \mathrm{d}^4x\, \mathrm{d}^4y\, \rho^\alpha(x)\, \Delta_{00}^{\alpha\beta}(x-y)\, \rho^\beta(y) + (\text{terms involving higher powers of } \rho). \tag{6.26}$$

Δ_{00} is the time–time component of the full gauge field propagator. Due to the higher powers of the density ρ appearing in (6.26), the energy of two separated sources cannot be expressed purely in terms of an interparticle potential (plus self energy contribution). It is only if we neglect these terms that we obtain such a potential.[4] For a quark and antiquark separated by a distance R, a combination of (6.25) and (6.26) yields

$$E(R)=\int \Delta_{00}(R,t)\,\mathrm{d}t. \tag{6.27}$$

Going over to momentum space

$$E(R)=\frac{1}{(2\pi)^3}\int \Delta_{00}(\boldsymbol{k}^2)\exp(-\mathrm{i}\boldsymbol{kR})\,\mathrm{d}\boldsymbol{k}. \tag{6.28}$$

The large distance behavior of the potential is determined from the behavior of the propagator near $\boldsymbol{k}=0$. A singularity of the form $k^{-\alpha}$ leads to

$$E(R)\sim R^{\alpha-3}. \tag{6.29}$$

For $\alpha=2$ we obtain the usual Coulomb behavior. A linear growth of the interquark potential would come about from a k^{-4} infrared behavior of the propagator.

We may also understand why such a singular behavior will lead to confinement, as well as the limitations of looking only at the propagator, by studying the behavior of the Wilson loop expectation value. (Recall the discussion in section 2.4.) This matrix element can be expanded in terms of connected n-point functions of the gauge potentials,

$$\left\langle \mathrm{Tr}\,\mathrm{P}\exp \mathrm{i}\oint A\,\mathrm{d}x\right\rangle=\mathrm{Tr}\exp\left\{-\tfrac{1}{4}\oint \mathrm{d}x^{\mu}\,\mathrm{d}x^{\nu}\,\Delta_{\mu\nu}(x-y)+\cdots\right\}, \tag{6.30}$$

where the dots indicate contributions from n-point functions with $n>2$. If we keep only the propagator part, then the small k singular structure of $\Delta(k^2)$ determines the behavior of (6.30). For $\alpha=2$, $\Delta\sim 1/(x-y)^2$ and the double integral in (6.30) is proportional to the circumference of the loop; for $\alpha=4$, $\Delta\sim\ln|x-y|$ and this term is proportional to the area. Within the limitations of keeping the two point function contribution to the energy or Wilson loop, we see that a determination of the infrared behavior of the propagator determines whether the theory confines or not.

Clearly, perturbation theory will not change the $1/k^2$ behavior to a $1/k^4$ one. Higher orders, at best, increase the small k^2 singularity structure by power $\log(k^2)$. We shall address ourselves to the question of whether a selective, but infinite set of perturbation diagrams could yield a $1/k^4$ behavior. The starting point will be the Dyson–Schwinger equations for the propagator. We shall discuss, in turn, two approaches [6.6, 6.7] to solving these equations. They yield similar results. The first one [6.6] will be presented in greater detail as it is more involved. The axial gauge, $n\cdot A=0$, is used in ref. [6.6]; in ref. [6.7] a covariant one is used. The equation for the "inverse" propagator (refer to appendix D for the definition of the inverse propagator and the reason for placing it in quotation marks) is shown diagramatically in fig. 6.3. The algebraic form of this equation is (as ordered in fig. 6.3)

$$\begin{aligned}\pi_{\mu\nu}(q)=\mathrm{i}(q^2g_{\mu\nu}-q_\mu q_\nu)&+\tfrac{1}{2}\int(\mathrm{d}p)\,\mathrm{i}\Gamma^0_{\nu\lambda\lambda'}(q,-r,-p)\Delta_{\lambda\sigma}(r)\,\Delta_{\lambda'\sigma'}(p)\,\mathrm{i}\Gamma_{\mu\sigma\sigma'}(q,r,p)\\&-g^2\int(\mathrm{d}p)\,[g_{\mu\nu}\Delta_{\lambda\lambda}(p)-\Delta_{\mu\nu}(p)]+(\text{last two diagrams}).\end{aligned} \tag{6.31}$$

[4] A similar problem is encountered in appendix B.

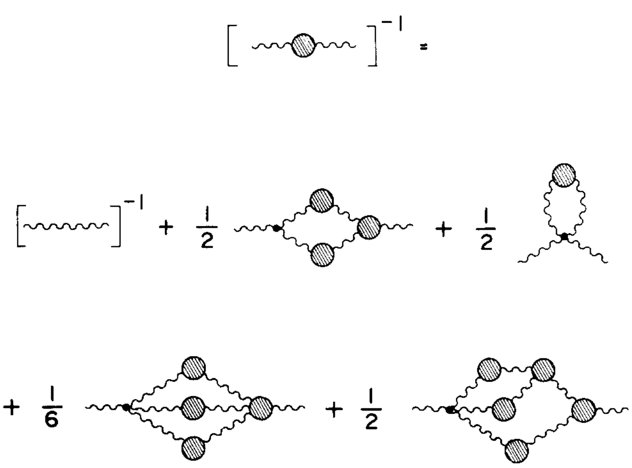


Fig. 6.3. Full Dyson–Schwinger equation for the gluon propagator. Shaded circles indicate full propagators and full vertices.

$(\mathrm{d}p)$ is shorthand for $3\mathrm{d}^4p/(2\pi)^4$. Γ^0 is the bare three point vertex and Γ is the full vertex. The last two terms are not written out in full as, under certain simplifying assumptions, they do not contribute to the resulting equations.

Equation (6.31) is not complete unless we specify the vertex Γ. The Ward identities (appendix D) provide part of this vertex. They do not determine the transverse part, namely the one that vanishes when any index is contracted with its momentum. Is the part of the vertex provided by the Ward identities (the longitudinal part) sufficient for the determination of the infrared structure of the propagator? In similar calculations in Abelian QED it is sufficient [6.8]. This is due to the fact that as far as the infrared on shell singularities are concerned, the longitudinal part carries all of these. For example, the vertex of a spin 1/2 particle, has a part proportional to γ_μ, which is both determined by the Ward identity and has all the infrared singularities; the other part of this vertex, namely the one proportional to $\sigma_{\mu\nu}q^\nu$, vanishes in the infrared limit. In QCD this has been demonstrated only to the one loop level [6.9]. Additionally, in QED it is the singular part of the vertex that drives the singular parts of the propagator. To pursue this treatment for non-Abelian theories the same assumptions must be made. Are these assumptions valid in QCD? QED is a weak coupling theory in the infrared and a strong coupling one in the ultraviolet; QCD, if it confines, is just the opposite and such assumptions may not hold.

To proceed further we would have to write out the most general form of the propagator in terms of known spin terms and unknown scalar functions of q^2. Then we would determine, through the Ward identities, the longitudinal part of the vertex and solve the resulting coupled integral equations. To simplify this discussion we shall make the ansatz that the singular part of the propagator is proportional to the free propagator, with the latter given in (D.3),

$$\Delta_{\mu\nu}(q) = -\frac{\mathrm{i}Z(q^2)}{q^2+\mathrm{i}\epsilon}\left[g_{\mu\nu} - \frac{n_\mu q_\nu + n_\nu q_\mu}{n\cdot q} + \frac{n^2 q_\mu q_\nu}{(n\cdot q)^2}\right]. \tag{6.32}$$

If $Z(q^2) \sim 1/q^2$, then according to the discussions in the beginning of this section, we will obtain confinement. Finite Z would yield a Coulomb like potential. This ansatz yields a simple form of the "inverse" propagator

$$\pi_{\mu\nu}(q) = \mathrm{i}Z^{-1}(q^2)[q^2 g_{\mu\nu} - q_\mu q_\nu] \tag{6.33}$$

and the Ward identity, eq. (D.8), can be solved for the longitudinal vertex [6.6, 6.10]

$$\begin{aligned}\Gamma_{\lambda\mu\nu}(p, q, r) = -\{&g_{\lambda\mu}[Z^{-1}(p^2)p_\nu - Z^{-1}(q^2)q_\nu] \\ &+[Z^{-1}(p^2) - Z^{-1}(q^2)][p\cdot q\, g_{\lambda\mu} - q_\lambda p_\mu](q-p)_\nu/(p^2-q^2) \\ &+\text{cyclic permutations of } \lambda,\ \mu,\ \nu\}.\end{aligned} \tag{6.34}$$

Both the propagator, (6.32), and the vertex, (6.34), depend on one unknown scalar function $Z(q^2)$. We insert these results into the integral equation, (6.31), to obtain an equation for this function. The spin structure is simplified by contracting both sides with the vector n_μ. We now see why this procedure eliminates the last two diagrams of fig. 6.3. Both of these diagrams involve a bare four gluon vertex coupling an external line to a full propagator. This vertex contains no derivatives; it consists of a sum of terms as $g_{\mu\nu}g_{\lambda\eta}$, where the indices refer to the polarizations of the lines entering the vertex. Thus, upon contracting the diagram with n_μ, we always obtain expressions involving $n_\lambda \Delta_{\lambda\eta}$; this expression vanishes. This argument does not hold for the second diagram because the n_μ can be contracted with some of the momenta in the vertex; it does not hold for the tadpole diagram because the indices corresponding to the outside lines can be contracted yielding a factor of n^2.

Proceeding as outlined above, we obtain

$$\begin{aligned}q^2(1-1/\gamma)\, Z^{-1}(q) = q^2(1-1/\gamma) + g^2 \int (\mathrm{d}k)\, K_{\sigma\sigma'}(k, k', n) \\ \times\left\{-\frac{Z(k)}{Z(q)}\frac{Z(k')-Z(q)}{k'^2-q^2}(q+k')_\sigma q_{\sigma'} + Z(k)\delta_{\sigma\sigma'}\right. \\ \left. + \frac{Z(k)-Z(k')}{k^2-k'^2}(k\cdot k'\,\delta_{\sigma\sigma'} - k_\sigma\cdot k_{\sigma'})\right\} + c(\gamma).\end{aligned} \tag{6.35}$$

In the above, $k' = q - k$ and

$$\begin{aligned}&\gamma = n^2q^2/(n\cdot q)^2 \\ &K_{\sigma\sigma'} = \frac{n\cdot(k-k')\, n\cdot k'}{n^2}\Delta^{(0)}_{\lambda\sigma}(k)\,\Delta^{(0)}_{\lambda\sigma'}(k') \\ &c(\gamma) = -\int (\mathrm{d}k)\, Z(k)\,(2+\gamma)/k^2.\end{aligned} \tag{6.36}$$

Gauge invariance requires that the q^2 independent, quadratically divergent $c(\gamma)$ cancels a similar term from the integral (this is the requirement that no mass terms appear in the gauge potential propagator). Before proceeding with the solution we must renormalize this equation. Let us assume that we have regularized this equation in such a way as to eliminate all quadratic divergences. We still must worry about the logarithmic ones. We can rewrite eq. (6.35)

$$Z^{-1}(q) = 1 + g^2 \int K(k, q, n)\, Z(k)(\mathrm{d}k) + g^2 Z^{-1}(q) \int L(k, q, n)\, Z(k)\, Z(q-k)\, (\mathrm{d}k). \tag{6.37}$$

This equation still contains logarithmic divergences. Renormalizability of the underlying theory guarantees that these can be eliminated by defining a renormalized $Z_{\mathrm{R}}(q)$ and a renormalized coupling constant g. $Z_{\mathrm{R}}(q)$ is defined to be one at $q_\mu = M_\mu$ for some fixed M_μ

$$Z(q) = Z_{\mathrm{R}}(q)\, Z(M). \tag{6.38}$$

Defining a renormalized coupling constant

$$g_{\mathrm{R}}^2(M) = \frac{g^2 Z(M)}{1 + g^2 Z(M) \int (\mathrm{d}k)\, K(k, M, n)\, Z_{\mathrm{R}}(k)}, \tag{6.39}$$

which in view of (6.37) is also given by

$$g_{\mathrm{R}}^2(M) = \frac{g^2 Z^2(M)}{1 - g^2 Z^2(M) \int (\mathrm{d}k)\, L(k, M, n)\, Z_{\mathrm{R}}(k)\, Z_{\mathrm{R}}(M-k)}, \tag{6.40}$$

we find an integral equation involving only finite quantities

$$\begin{aligned} \frac{1}{Z_{\mathrm{R}}(q)} = {} & 1 + g_{\mathrm{R}}^2(M) \int [K(k, q, n) - K(k, M, n)]\, Z_{\mathrm{R}}(k)\, (\mathrm{d}k) \\ & + \frac{g_{\mathrm{R}}^2(M)}{Z_{\mathrm{R}}(q)} \int [L(k, q, n)\, Z_{\mathrm{R}}(q-k) - L(k, M, n)\, Z_{\mathrm{R}}(M-k)]\, Z_{\mathrm{R}}(k)\, (\mathrm{d}k). \end{aligned} \tag{6.41}$$

Let us first look at perturbation theory; it should be valid in the asymptotically free ultraviolet region. We will not go into details, but the results are (combined with a renormalization group argument)

$$Z_{\mathrm{R}}(q) \to (1 + g_{\mathrm{R}}^2(M) \ln(q^2/M^2)/\pi^2)^{11/16}. \tag{6.42}$$

The exact result is (eq. (6.2))

$$Z_{\mathrm{R}}(q) \to (1 + 11 g^2(M) \ln(q^2/M^2)/16\pi^2). \tag{6.43}$$

To first order in g^2 the results agree. We do not expect complete agreement as the neglect of the transverse part of the vertices as well as our ansatz on the propagator was geared to the infrared. An exact solution of eq. (6.41) is beyond our scope. An approximate solution based on a trial function with adjustable parameters was obtained [6.7]. With confinement and eq. (6.42) in mind, the form chosen was

$$Z_{\mathrm{R}}(M) = \frac{AM^2}{q^2} + (1-A) \frac{q^2(M^2 + \mu_2^2)}{M^2 + \mu_2^2} \left(1 + \frac{g^2(M)}{\pi^2} \ln \frac{q^2 + \mu_3^2}{M^2 + \mu_3^2}\right)^{11/16}. \tag{6.44}$$

A, μ_2/M and μ_3/M are variational parameters. Solutions with input and output $Z_R(q)$ matching closely for $0<q^2<2.5$ and independent of γ to 5% are shown in fig. 6.4 [6.11], $A\neq 0$ indicates an infrared singularity implying confinement.

A comparison with previous results can be made through the β-function. We use eq. (6.39) to re-express the finite coupling constant $g^2(q)$ at some arbitrary point as a function of the equally finite constant at a fixed value $q^2=M^2$,

$$\frac{1}{g_R^2(q)}=\frac{1}{g_R^2(M)\,Z_R(q)}+\int (\mathrm{d}k)\,[K(k,q,n)-K(k,M,n)]\,Z_R(k). \tag{6.45}$$

From this we obtain $\beta(q)=-q\,\partial g/\partial q$. Results based on the approximate solution, eq. (6.44), are shown in fig. 6.5. Qualitatively, they are the same as comparable functions based on the strong coupling expansion, fig. 5.1. The transition from strong to weak coupling is not as rapid and it occurs for somewhat higher values of g. This may be due in part to the inadequacy of this treatment at small g, as shown in (6.42) and (6.43). Differences in definitions of the coupling constant make it difficult to give precise comparisons between the various approaches.

The second study [6.7] of the Dyson–Schwinger equations makes further simplifying assumptions. Comparing (6.32) and (6.34) we note that corrections to the three gluon vertex are partially cancelled by corrections to the propagator. This is true for the supposedly singular parts of the corrections. Assuming complete cancellation, the first two terms of fig. 6.3 reduce to those depicted in fig. 6.6. Neglecting four gluon couplings (the tadpole diagram is removed by mass renormalization) and Fadeev–Popov ghosts [2.2] (we work in a covariant Landau gauge) the new version of eq. (6.31) is

$$\pi_{\mu\nu}(q)=\mathrm{i}(q^2g_{\mu\nu}-q_\mu q_\nu)+\tfrac{1}{2}\int (\mathrm{d}p)\,\mathrm{i}\Gamma^0(q,-r,-p)\,\Delta_{\lambda\sigma}(r)\,\Delta^0_{\lambda'\sigma'}(p)\,\mathrm{i}\Gamma^0_{\mu\sigma\sigma'}(q,r,p). \tag{6.46}$$

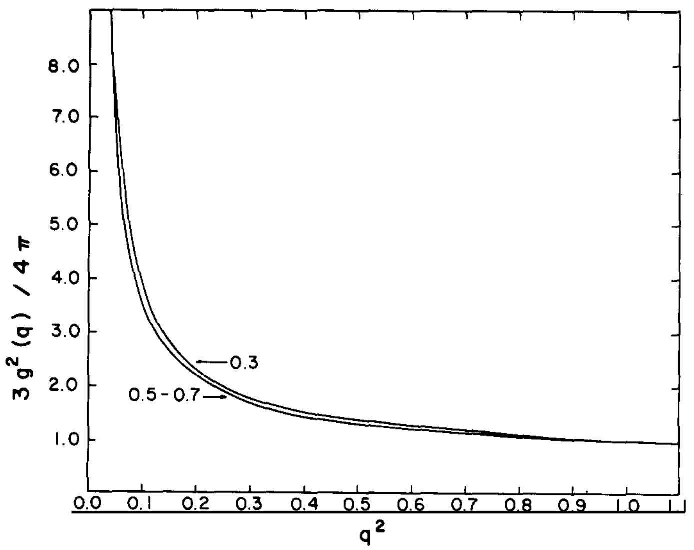


Fig. 6.4. Approximate solution of the Dyson–Schwinger equations. The two curves are for different values of the gauge parameter $\gamma=n^2q^2/(n\cdot q)^2$. [J. Ball, private communication.]

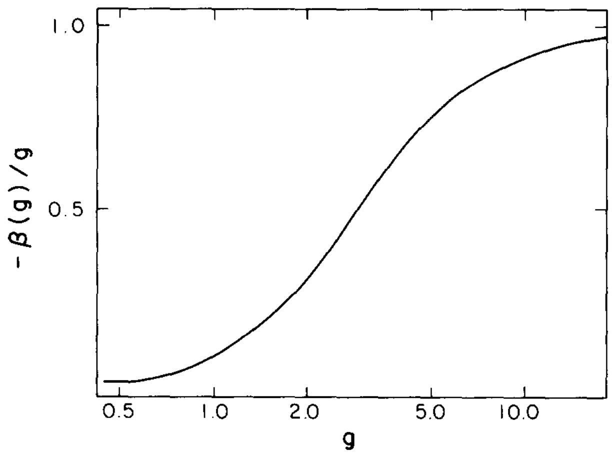


Fig. 6.5. Renormalization group β-function from the solution of the approximate Dyson–Schwinger equation. (From ref. [6.6].)

In the Landau gauge the propagator is

$$\Delta_{\mu\nu}(q) = -\frac{\mathrm{i}Z(q^2)}{q^2 + \mathrm{i}\epsilon}[g_{\mu\nu} - q_\mu q_\nu/q^2] \tag{6.47}$$

with $\pi_{\mu\nu}$ as in (6.33). Eq. (6.46) provides the desired equation for Z:

$$Z(q^2) = \left[1 - \frac{1}{q^2}\int K(q^2, k^2)\, Z(k^2)\, \mathrm{d}k^2\right]^{-1}. \tag{6.48}$$

$K(q^2, k^2)$ is a known kinematic function. This equation has been solved numerically with an ansatz similar to (6.44) and again yields a non-zero value for A, implying confinement.

What are the shortcomings of these approaches?

1. Is a calculation of the gluon propagator meaningful? In a confining theory it does not exist as it is not gauge invariant.

2. How important in determining the interquark potential are the other terms of eq. (6.26) or (6.30)? This question is addressed in ref. [6.7].

3. How valid are the various truncations of the Dyson–Schwinger equations? Does that part of the vertex determined by the Ward identities drive the infrared singularities of the propagator?

In spite of points 1 and 2 above, it is possible that the β function calculated in this approach may be a

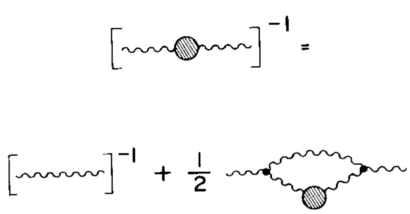


Fig. 6.6. Dyson–Schwinger equation assuming a full cancellation of the singular parts of the vertex and propagator.

correct one based on some renormalization scheme. The absence of zeroes between strong and weak couplings would indicate confinement.

7. Electric-magnetic duality and the phases of QCD

The approximate results discussed in the earlier sections indicate that the perturbative, small coupling constant regime, and the large coupling one are in the same phase. As the strong coupling phase does confine and if we take the previous results seriously, then QCD is a theory with permanent quark confinement. However, all the previous results are approximate and for several of the approximations are hard, if not impossible, to improve upon. Certainly, it would be desirable to have nonapproximate results. To date such arguments have not been able to determine whether QCD does or does not confine. They have been able to illuminate the possible phases of QCD. These techniques also isolate which features of non-Abelian theories may be responsible for confinement. Some limited numerical work has been done based on arguments to be presented.

The inspiration for these studies comes from statistical mechanics. The technique we wish to borrow from that area is the Kramers–Wannier [7.1] duality transformation. In statistical mechanics this transformation relates high and low temperature phases of a theory. In our case, such a transformation, if it exists, would relate strong and weak coupling regimes. In the strong coupling region confinement comes about easily, while in the perturbative one Lorentz invariance is manifest. The advantage of such transformations is that we could transform a theory with one set of known properties into a domain where the other ones could be studied with relative ease.

It is straightforward to implement such transformations in the case of Abelian theories. For non-Abelian theories, even non-gauge ones, there is no simple generalization of such duality transformations. An inspection of the QCD Hamiltonian, eq. (3.2), shows that a transformation of $g \leftrightarrow g^{-1}$ involves a transformation of $E \leftrightarrow B$. The ideal situation would have been had it been possible to find a potential A' dual to the usual vector potential A such that the relations of the electric and magnetic fields to A' were the opposite of those to A. An attempt along this line has been made [7.2]. We shall present a different approach due to 't Hooft [7.3]. Not only do we find the possible phases of QCD but also the features necessary for confinement.

As mentioned earlier, a duality transformation involves an interchange of electric and magnetic gauge fields. We do not expect an eigenstate of the QCD Hamiltonian, or more specifically the vacuum eigenstate to have definite values for such field variables. In non-Abelian theories operators for magnetic and electric fields are not even gauge invariant. In Abelian theories we might hope to classify states by the amount of electric or magnetic flux they possess. For non-Abelian theories, even this will turn out to be subtle.

If space is unconstrained no directions are preferred and we do not expect flux through some arbitrary surface to be relevant. For this discussion it is convenient to constrain the world to a box with dimensions a_1, a_2, a_3 and at the end study the thermodynamic limit $a_i \to \infty$. We may then discuss the fluxes along these preferred directions. A flux tube can be prevented from wandering from one direction to another by giving the system the geometry of a torus, namely identifying opposite sides. What this implies for the gauge potentials will be discussed further on.

What relevance does this have for confinement? Let us consider a quark and antiquark separated by a large distance along the three direction. Gauss' law forces us to have a definite amount of electric flux along this direction. We then ask what is the energy of a configuration with a certain amount of flux in this direction. In a perturbative or nonconfining regime the flux spreads out in the transverse directions and the energy behaves as a_3/a_1a_2. In the $a_i \to \infty$ limit it goes to zero. If confinement holds we expect the electric flux to be concentrated into tubes of fixed transverse size and the energy to behave as a_3. Thus the behavior of a configuration with a definite amount of electric flux determines whether a theory does or does not confine.

7.1. Definition of magnetic flux

It is a standard exercise of classical electrodynamics to find the vector potential outside a magnetized infinite cylinder. We assume the magnetization is uniform and the total flux through any surface in Φ. For a cylinder stretching along the 3 direction the answer is $A_3 = 0$ and the transverse components, in cylindrical coordinates are

$$A_i = \frac{\Phi}{2\pi}\frac{\hat{\varphi}_i}{\rho}. \tag{7.1}$$

As expected A is a pure gauge

$$\boldsymbol{A} = \frac{\Phi}{2\pi}\nabla\varphi. \tag{7.2}$$

The "gauge function" $(e\Phi\varphi/2\pi)$ is not single valued. The jump in this function in going around a magnetized cylinder is just $e\Phi$. If charged fields are present and we want this potential to have no effect on them (no Bohm–Aharanov effect) we require that the gauge transformation be single valued, or that $\Phi = 2\pi n/e$. Of course this potential satisfies the usual definition of magnetic flux

$$\Phi = \int \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{s} = \oint \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{l}. \tag{7.3}$$

The form of A in eq. (7.2) is not convenient for the periodic geometry we have in mind.

For a rectangular area $-a_1/2 < x < a_1/2$, $-a_2/2 < y < a_2/2$ the field outside some flux can equally be a pure gauge generated by the nonsingle valued function

$$\lambda(x, y) = \frac{\Phi}{2\pi}\arcsin\left(\frac{2y}{a_2}\right). \tag{7.4}$$

λ is continuous as we go around the boundary of the rectangle and has a jump discontinuity at $x = a_1/2$, $y = 0$. The potential generated by (7.4) is periodic in the y direction and changes sign between $x = -a_1/2$ and $x = a_1/2$. In what sense do we have periodicity?

$$A(a_1/2, y) = A(-a_1/2, y) + \nabla(2\lambda). \tag{7.5}$$

The corresponding transformation on a charged field is $\exp(2\mathrm{i}e\lambda)$. Requiring this gauge transformation to be single valued places the same restriction on Φ as discussed earlier.

Such arguments may be carried over to the non-Abelian case. We will consider not only truly periodic potentials, but also potentials which are periodic up to gauge transformations

$$
\begin{aligned}
A(a_1/2, y) &= \Omega_1^+(y)\, A(-a_1/2, y)\, \Omega_1(y) + (\mathrm{i}/g)\, \Omega_1^+(y)\, \nabla\Omega_1(y) \\
A(x, a_2/2) &= \Omega_2^+(x)\, A(x, -a_2/2)\, \Omega_2(x) + (\mathrm{i}/g)\, \Omega_2^+(x)\, \nabla\Omega_2(x).
\end{aligned}
\tag{7.6}
$$

A further gauge transformation can bring Ω_2 to unity and we then have a situation analogous to (7.5). Single valuedness would require $\Omega_1(-a_2/2) = \Omega_1(a_2/2)$. This is too restrictive [7.5]. As long as there are no nonzero triality fields present (no dynamical quarks) then it is sufficient to demand that

$$\Omega_1(a_2/2) = \Omega_1(-a_2/2)z \tag{7.7}$$

with $z \in Z_3$ the center of the group. The center of SU(3) consists of elements of the form

$$
\begin{aligned}
&z = \exp 2\pi\, \mathrm{i} m Q \\
&Q = \frac{1}{3}\begin{pmatrix} -1 & & \\ & -1 & \\ & & 2 \end{pmatrix} .
\end{aligned}
\tag{7.8}
$$

If we do not insist that Ω_2 is the identity transformation, then (7.7) generalizes to

$$\Omega_1(a_2/2)\, \Omega_2(-a_1/2) = \Omega_2(a_1/2)\, \Omega_1(-a_2/2)z. \tag{7.9}$$

The non-Abelian versions of (7.2) and (7.4) are

$$S(x, y) = \exp[\mathrm{i} m\varphi Q] \tag{7.10a}$$

$$S(x, y) = \exp[\mathrm{i} m \arcsin(2y/a_2) Q]. \tag{7.10b}$$

Applying either of the above transformations increases the "magnetic flux" by m units (modulo N) in the three direction. This discussion may be extended to other planes and "twisted" periodic [7.6] gauge potentials can be classified by three numbers $\boldsymbol{m}$.[5] This definition of flux in the non-Abelian case is quite different from the Abelian one. There is no analogue of (7.3).

A quantum mechanical state having a prescribed amount of magnetic flux $\boldsymbol{m}$ is described by a wave functional $\psi_{\boldsymbol{m}}(A)$ which vanishes for $A(x)$ not carrying $\boldsymbol{m}$ units of flux.

7.2. Definition of electric flux

We again appeal to an analogy with Abelian theories in order to define electric flux. In such a theory the commutation relations analogous to eq. (2.7) are

$$[E_i(\boldsymbol{x}), A_j(\boldsymbol{y})] = -\mathrm{i}\delta_{ij}\, \delta(\boldsymbol{x} - \boldsymbol{y}). \tag{7.11}$$

[5] The classification of gauge potentials into various magnetic flux classes is not unique. Namely, a given A_μ with periodic boundary conditions may belong to classes with different $\boldsymbol{m}$'s. For example, the trivial configuration $A_\mu = 0$ belongs to classes with all $\boldsymbol{m}$'s. There exist nontrivial potentials that also exhibit ambiguity. It is conjectured that such nonclassifiable potentials form a set of functional measure zero.

For the moment we consider a situation where some electric flux is confined to a thin tube of area a following some curve C (not necessarily closed). Let C be parametrized by $x_i(\tau)$. The operator determining the amount of flux along C is

$$F_E(x) = aE \cdot \frac{\mathrm{d}x}{\mathrm{d}\tau} \Big/ \left| \frac{\mathrm{d}x}{\mathrm{d}\tau} \right| . \tag{7.12}$$

A state, described by $\psi(A)$, is said to have a certain flux along C if it is an eigenstate of F_E. It is easy to see that the operator

$$\mathscr{E}_\mathrm{c} = \exp \mathrm{i}g \int_C A \cdot \mathrm{d}l \tag{7.13}$$

creates a unit of flux along C. This follows from the commutation relations (7.11)

$$F_E(x)\, \mathscr{E}_\mathrm{c} \psi(A) = \mathscr{E}_\mathrm{c}\, F_E \psi(A) + g\mathscr{E}_\mathrm{c}\psi(A). \tag{7.14}$$

This concept may be generalized to non-Abelian theories by defining the creation operator for a unit of electric flux (in units of g) along the 3 direction to be

$$\mathscr{E}_3 = \mathrm{P}\, \exp \mathrm{i}g \int_C A \cdot \mathrm{d}l. \tag{7.15}$$

A is now the non-Abelian vector potential. C is any curve connecting $x_3 = -a_3/2$ and $x_3 = a_3/2$ (a_3 is the extent of our box world in the three direction). Although C could have any shape we shall in the future take it to be a straight line going along the 3 direction. Its location in the 1–2 plane is arbitrary. Given a state with no electric flux, $\psi_0(A)$, $\mathscr{E}_3$ acting on it creates a unit of flux in the 3 direction

$$(\mathscr{E}_3)^{e_3} \psi_0(A) = \psi_{e_3}(A). \tag{7.16}$$

We may generalize to any of the other two directions to define $\psi_e(A)$. There is another way of looking at states with electric flux which clarifies how to choose ψ_0 and places restrictions on the integers e_α. Consider a gauge transformation g_k such that it is unity on the $x_3 = -a_3/2$ plane and belongs to the center for $x_3 = a_3/2$. On that plane $g_{k_3} = \exp(2\pi\, \mathrm{i}k_3)/3$. It is easy to see that under such gauge transformations

$$g_{k_3} \mathscr{E}_3(\mathrm{c})\, g_{k_3}^{-1} = \exp(2\pi\, \mathrm{i}k_3/3)\, \mathscr{E}_3(\mathrm{c}) \tag{7.17}$$

and thus

$$g_k \psi_{e_3}(A) = \exp(2\pi\, \mathrm{i}k_3 e_3/3)(\mathscr{E}_3)^{e_3} g_{k_3} \psi_0(A). \tag{7.18}$$

We choose $\psi_0(A)$ to be invariant under the transformations g_k and thus

$$g_k \psi_e(A) = \exp(2\pi\, \mathrm{i}\boldsymbol{k} \cdot \boldsymbol{e}/3)\, \psi_e(A). \tag{7.19}$$

Thus we see that we can distinguish the integers e_α only within module 3 classes. We will also need eigenstates of $\mathscr{E}_\alpha(\mathrm{c})$. These are superpositions of the ψ_e,

$$\psi(\boldsymbol{k}; A) = \frac{1}{\sqrt{27}} \sum_{\boldsymbol{e}} \exp(2\pi \mathrm{i} \boldsymbol{k} \cdot \boldsymbol{e}/3) \psi_{\boldsymbol{e}}(A) \tag{7.20a}$$

$$\mathscr{E}_\alpha(\mathrm{c})\, \psi(\boldsymbol{k}; A) = \exp(2\pi \mathrm{i} k_\alpha/3)\, \psi(\boldsymbol{k}; A) \tag{7.20b}$$

$$g_l\, \psi(\boldsymbol{k}; A) = \psi(\boldsymbol{k} + \boldsymbol{l}; A). \tag{7.20c}$$

7.3. Energy of configurations with specified electric and magnetic flux

We may now combine the ideas of the previous two sections and define states with both electric and magnetic flux. A state with magnetic flux vector $\boldsymbol{m}$ and electric flux vector $\boldsymbol{e}$ is denoted by $\psi_{\boldsymbol{e};\boldsymbol{m}}(A)$. We also define states analogous to those of (7.20)

$$\psi_{\boldsymbol{m}}(\boldsymbol{k}; A) = \frac{1}{\sqrt{27}} \sum_{\boldsymbol{e}} \exp(-2\pi \mathrm{i} \boldsymbol{e} \cdot \boldsymbol{k}/3)\, \psi_{\boldsymbol{e};\boldsymbol{m}}(A). \tag{7.21}$$

The energy of this configuration may be determined by studying

$$\exp[-\beta F(\boldsymbol{e}, \boldsymbol{m}; \boldsymbol{a}, \beta)] = \sum \langle \psi^*_{\boldsymbol{e};\boldsymbol{m}}(A), \mathrm{e}^{-\beta H} \psi_{\boldsymbol{e};\boldsymbol{m}}(A) \rangle. \tag{7.22}$$

H is the QCD Hamiltonian; β is a Euclidian time (we use β to draw out the analogy with statistical mechanics); a_α is the length of our box world in the α direction. The summation on the right hand side is over all states consistent with the specified electric and magnetic flux configurations. In the large β limit F approaches the energy of the ground state subject to the given electric and magnetic fluxes. (In order to prevent double counting a resolution of the questions raised in footnote 4 must be made.) Define

$$W(\boldsymbol{k}, \boldsymbol{k}'; \boldsymbol{m}; \boldsymbol{a}, \beta) = \sum \psi_{\boldsymbol{m}}(k'; A)\, \mathrm{e}^{-\beta H} \psi_{\boldsymbol{m}}(k; A) \tag{7.23}$$

then

$$\exp(-\beta F) = \frac{1}{27} \sum_{\boldsymbol{k},\boldsymbol{k}'} \exp\{2\pi \mathrm{i}(\boldsymbol{k} - \boldsymbol{k}') \cdot \boldsymbol{e}/3\}\, W(\boldsymbol{k}, \boldsymbol{k}'; \ldots). \tag{7.24}$$

We can sum over $\boldsymbol{k}, \boldsymbol{k}'$ keeping $q = \boldsymbol{k} - \boldsymbol{k}'$ fixed and let

$$W(\boldsymbol{q}; \boldsymbol{m}; \boldsymbol{a}, \beta) = \sum_{\boldsymbol{k}} W(\boldsymbol{q} + \boldsymbol{k}, \boldsymbol{k}; \boldsymbol{m}; \boldsymbol{a}, \beta).$$

(It is also true that $\langle W(\boldsymbol{k}, \boldsymbol{k}'; \ldots) \rangle$ depends only on the difference $\boldsymbol{k} - \boldsymbol{k}'$.) Repeating (7.22) we obtain

$$\exp - [\beta F(\boldsymbol{e}, \boldsymbol{m}; \boldsymbol{a}, \beta)] = \frac{1}{27} \sum_{\boldsymbol{q}} \exp(2\pi \mathrm{i} \boldsymbol{q} \cdot \boldsymbol{e}/3)\, \langle W(\boldsymbol{q}, \boldsymbol{m}; \boldsymbol{a}, \beta) \rangle. \tag{7.25}$$

What is the purpose of introducing the functions W? It turns out that they have properties that can be easily carried over from the electric to the magnetic case.

The W's have a functional integral representation similar to the standard vacuum amplitudes. Equations (7.6) and (7.20) provide the necessary modifications

$$\langle W(q, m; a, \beta)\rangle = \int [\mathrm{d}A]_{q,m} \exp\left[-\int L\, \mathrm{d}^4x\right]. \tag{7.26}$$

The x integration is restricted to the previously described box of lengths a_i in the three directions, and to β in the t direction. The notation $[\mathrm{d}A]_{q;m}$ implies that the functional integration is restricted to potential configurations satisfying twisted periodic boundary conditions. The magnetic flux restriction forces the potentials to obey periodicities in the i–j plane as given in (7.9) and those in the t–i plane given by q_i. The latter follows from (7.23) and (7.20c). All these results may be summarized in a four dimensional notation (cf. (7.6))

$$A\left(\cdots \frac{a_i}{2} \cdots\right) = \Omega_i^+ A\left(\cdots - \frac{a_i}{2} \cdots\right)\Omega_i + \frac{\mathrm{i}}{g}\Omega_i^+ \nabla \Omega_i, \tag{7.27}$$

with the dots denoting the other three variables; Ω_i is likewise a function of these three variables. Following (7.9) we define n_{ij} by

$$\Omega_i\left(\cdots \frac{a_j}{2} \cdots\right) \Omega_j\left(\cdots - \frac{a_i}{2} \cdots\right) = \Omega_j\left(\cdots \frac{a_i}{2} \cdots\right) \Omega_i\left(\cdots - \frac{a_j}{2} \cdots\right) \exp[2\pi\, \mathrm{i} n_{ij} Q]. \tag{7.28}$$

This translates into the following relations to the magnetic and electric fluxes

$$\begin{aligned} &n_{ij} = \epsilon_{ijk} m_k \qquad (i, j = 1, 2, 3)\\ &n_{it} = q_i. \end{aligned} \tag{7.29}$$

7.4. Duality

We are familiar with the fact that Lorentz transformations mix up electric and magnetic fields. The only remanent of Lorentz or really Euclidian invariance we have is 90° rotations. Consider the transformation which changes the unit vectors, $\hat{\boldsymbol{x}}_\alpha$, along our four dimensional box as follows

$$\begin{aligned} &\hat{\boldsymbol{x}}_1 \to \hat{\boldsymbol{x}}_2, \qquad \hat{\boldsymbol{x}}_2 \to -\hat{\boldsymbol{x}}_1\\ &\hat{\boldsymbol{x}}_3 \to -\hat{\boldsymbol{x}}_t, \qquad \hat{\boldsymbol{x}}_t \to \hat{\boldsymbol{x}}_3. \end{aligned} \tag{7.30}$$

Closed curves in the ij planes also change

$$\begin{aligned} &C_{23} \leftrightarrow C_{1t}, \qquad C_{13} \leftrightarrow C_{2t}\\ &C_{12} \leftrightarrow C_{12}, \qquad C_{2t} \leftrightarrow C_{13}. \end{aligned} \tag{7.31}$$

If we denote the first two components of a three vector $\boldsymbol{k}$ by $\tilde{k}$ and this two vector with its components interchanged by $\hat{k}$ we find that in $W(\boldsymbol{q}, \boldsymbol{m}; \boldsymbol{a}, \beta)$

$$\begin{aligned}(\tilde{a}, a_3, \beta) &\leftrightarrow (\hat{a}, \beta, a_3)\\ \tilde{m} &\leftrightarrow \tilde{q}\\ m_3 &\leftrightarrow m_3\\ q_3 &\leftrightarrow q_3,\end{aligned} \tag{7.32}$$

and thus, using (7.29)

$$W(\tilde{q}, q_3, \tilde{m}, m_3; \tilde{a}, a_3, \beta) = W(\tilde{m}, q_3, \tilde{k}, m_3; \tilde{a}, \beta, a_3), \tag{7.33}$$

and

$$\exp[-\beta F(\tilde{e}, e_3, \tilde{m}, m_3; \tilde{a}, a_3, \beta)] = \frac{1}{9}\sum_{\tilde{q},\tilde{l}} \exp\left\{\frac{2\pi \mathrm{i}}{3}[\tilde{q}\cdot\tilde{e} - \tilde{m}\cdot\tilde{l}]\right\} \exp[-a_3 F(\tilde{l}, e_3, \tilde{q}, m_3; \hat{a}, \beta, a_3)]. \tag{7.34}$$

Equation (7.34) is the basis for the electric-magnetic duality. It relates configurations with electric and magnetic fluxes, in some directions, interchanged. It is exact.

7.5. Flux energies

As the third components do not play a role in this discussion let us set $e_3 = m_3 = 0$. If no massless particles are present large volume and β limits are approached exponentially. Let $a_i, \beta \to \infty$. Not all F can approach zero. If they did, then the right hand side of (7.34) would approach $\delta(\tilde{e})\,\delta(\tilde{m})$, implying that only $F(0,0; \tilde{a}, a_3\beta)$ is zero. Consistency with (7.34) demands that 9 of the 81 possible flux combinations yield vanishing F's; these are referred to as light fluxes. Cancellation of the imaginary terms requires that if $(\tilde{e}_1, \tilde{m}_1)$ and $(\tilde{e}_2, \tilde{m}_2)$ are light fluxes then

$$\tilde{e}_1 \cdot \tilde{m}_2 = \tilde{e}_2 \cdot \tilde{m}_1 \qquad (\mathrm{mod}\ 3). \tag{7.35}$$

As $W(\tilde{q}, \tilde{m}; \tilde{a}, \beta)$ is a sum of diagonal matrix elements of positive operators, $\exp - \beta H$, it itself is positive. Equation (7.24) implies that if some $(\tilde{e}, \tilde{m})$ is a light flux, then some of the $W(q, \tilde{m}; \tilde{a}, \beta)$ are nonzero and in turn $\exp[-\beta F(0, \tilde{m}; \tilde{a}, \beta)]$ is nonzero implying that $(0, \tilde{m})$ is a light flux. If $(\tilde{e}, \tilde{m})$ is a light flux, (7.35) implies that $\tilde{e}\cdot\tilde{m} = 0$. Let us consider $\tilde{e}$ and $\tilde{m}$ both in the 1 direction. We then have $\tilde{e}_1\tilde{m}_1 = 0$ which implies that either e_1 or $m_1 = 0$.[6] This, plus invariance under the interchange of 1 and 2 and the requirement that there are 9 light fluxes leads us to the solution that *either* all electric or all magnetic fluxes are light. All other flux combinations lead to infinite F's. These are referred to as heavy fluxes.

How do we expect the F's for these situations to behave? In Abelian theories a flux in the 1 direction is created by fields that spread out in the 2 and 3 directions and the energy of such a configuration is

$$E \cong F \sim C a_1 / a_2 a_3. \tag{7.36}$$

[6] Had we been dealing with SU(4) then for $\tilde{e}_1 = \tilde{m}_1 = 2$, $\tilde{e}_1\tilde{m}_1 = 0$ (mod 4) and we would obtain more solutions of (7.35).

As the box size increases F tends to zero. For the heavy fluxes we do not have a rigorous argument for how their energy behaves, but our intuition tells us that the flux will be confined to a long tube with a finite, fixed thickness. For a heavy flux in the one direction

$$E = F \sim \sigma a_1. \tag{7.37}$$

σ will turn out to be the string tension discussed in previous sections. With assumption (7.37) it will be possible to obtain the energies of configurations with light fluxes. We need, however one more assumption, namely that whichever fluxes, electric or magnetic are light or heavy we have

$$F(\boldsymbol{e}, \boldsymbol{m}; \boldsymbol{a}, \beta) \xrightarrow[a,\beta\to\infty]{} F_{\mathrm{e}}(\boldsymbol{e}; \boldsymbol{a}, \beta) + F_{\mathrm{m}}(\boldsymbol{m}; \boldsymbol{a}, \beta). \tag{7.38}$$

This is a reasonable assumption; it is true for Abelian theories and if the heavy fluxes occupy a negligible portion of space while the light ones fill up all of it, the interference between them is negligible even for non-Abelian situations.

With the arguments presented thus far, we have no way of telling whether it is the electric or magnetic fluxes that are heavy or light. For the present discussion let us assume that it is the electric fluxes that are heavy – this corresponds to a confining situation. The role of electric and magnetic fluxes is interchangeable. We first wish to evaluate $F_{\mathrm{e}}(\boldsymbol{e}; \boldsymbol{a}, \beta)$. We have made the assumption, eq. (7.37) that a single flux tube in direction i has energy σa_i. The summation over states carrying a fixed electric flux includes a summation over the number of such fundamental tubes as well as their positions. A flux tube, in say the one direction, will contribute

$$\gamma_1 = \lambda a_2 a_3 \exp(-\beta\sigma a_1) \tag{7.39}$$

to $\exp(-\beta F_{\mathrm{e}})$. The exponent is, due to (7.37), just the factor $\exp(-\beta H)$; $a_2 a_3$ comes from the summation over positions of this tube; λ is an undetermined coefficient reflecting the normalization of these wavefunctions. Summing over a number of such tubes in both the 1 and 2 direction, we find

$$\begin{aligned} &\exp[-\beta F_{\mathrm{e}}(\tilde{e}; \boldsymbol{a}, \beta) + c(\boldsymbol{a}, \beta)] \\ &\quad = \sum_{n_1^\pm, n_2^\pm} \frac{1}{(n_1^+)!(n_1^-)!(n_2^+)!(n_2^-)!}\, \gamma^{n_1^+ + n_1^-} \gamma_2^{n_2^+ + n_2^-} \delta_3(n_1^+ - n_1^- - \tilde{e}_1)\, \delta_3(n_2^+ - n_2^- - \tilde{e}_2). \end{aligned} \tag{7.40}$$

The summation is over fluxes carrying one unit up or down the αth direction. Flux tubes carrying $n > 1$ units of flux contribute $\exp(-\beta n^2 \sigma a)$, and can be neglected compared to $n = 1$; flux tubes along diagonals, i.e., $\tilde{e} = (1, 1)$ contribute $\exp -\beta\sigma\sqrt{a_1^2 + a_2^2}$ and are less important than the ones retained. Using the definition

$$\delta_3(n) = \frac{1}{3}\sum_{k=0}^{2} \exp(-2\pi\, \mathrm{i} nk/3), \tag{7.41}$$

we obtain

$$\exp\{-\beta(F_{\mathrm{e}} + c)\} = \frac{1}{9}\sum_{\tilde{k}} \exp\left\{2\pi\, \mathrm{i}\tilde{k}\cdot\tilde{e}/3 + \sum_{\alpha=1}^{2} 2\gamma_\alpha \cos\frac{2\pi k_\alpha}{3}\right\}. \tag{7.42}$$

c will be determined from the requirement that $F_e(0; \boldsymbol{a}, \beta) = 0$. Note that for $e \neq 0$ $F_e(\tilde{e}, \boldsymbol{a}) \to \infty$ as expected. (7.12) is in a form where we can easily apply the duality relation (7.36) to find

$$\exp\{-a_3 F_m(\tilde{m}, 0; \hat{a}, \beta a_3)\} = \exp\left[\sum_\alpha 2\gamma_\alpha \left(\cos \frac{2\pi m_\alpha}{3} - 1\right)\right]. \tag{7.43}$$

We have incorporated the requirement that $F_m(0; \boldsymbol{a}) = 0$, and obtain

$$E_m(\boldsymbol{m}; \boldsymbol{a}) = \sum_{\alpha=1}^{3} 2\lambda \left(1 - \cos \frac{2\pi m_\alpha}{3}\right) a_\alpha \exp[-\sigma \epsilon_{\alpha\beta\gamma} a_\beta a_\gamma]. \tag{7.44}$$

The energy of a configuration with one unit of flux in the one direction is proportional to $a_1 \exp(-\sigma a_2 a_3)$ and thus vanishes as the box is extended in all directions. We find the result that if the electric fluxes are heavy, then in the thermodynamic limit there are no restrictions on the amount of magnetic flux the ground state of the system may contain. Had we chosen the magnetic fluxes to be heavy, then this statement could be made about the electric properties of the ground state.

Why do we refer to the situation with heavy electric fluxes as the confining phase? A state with one unit of flux can be created by placing a quark and antiquark at opposite walls of our box. The energy would be as given in (7.37) which is the familiar linear potential discussed previously. What are the characteristics of a phase where the magnetic fluxes are heavy? In analogy with the above, magnetic charges would be confined; electric sources could move apart as much as one wished, as their energy of separation would rapidly go to zero as indicated by (7.44), with electric and magnetic indices interchanged. Only short range interactions remain indicating that the gauge potentials become massive. This is a signal that a Higgs mechanism [2.2] took over, and this phase is referred to as a Higgs phase.

Are other phases possible? Yes, if massless particles are present and some of the previous interchanges of limits are not valid. This new phase partly resembles QED in that massless real gluons would exist as physical particles. This phase could be the perturbative one or one with a partial Higgs mechanism operating, and one or several U(1) subgroups remaining unbroken and their gauge fields turning into real states. In both these cases the ground state of the theory would be a superposition of states with various amount of electric and magnetic flux.

7.6. Applications

Can these considerations be put to use to help us resolve the central question as to whether QCD is a confining theory? A strong coupling constant expansion [7.7] of the vortex energies very similar to the expansion for the string tension discussed in sections 3 and 4 has been carried out. The results are consistent with the confining nature of QCD. An interesting numerical evaluation of F_m has been performed by a Monte Carlo integration similar to that presented in section 4. In this work the periodic world was discretized into a lattice. The calculations were limited to lattices of size 5^4. Again, due to numerical complications, the calculations were performed for an SU(2) rather than an SU(3) theory. For a d^4 lattice the quantity evaluated was

$$\mu(g, d) = \tfrac{1}{2} d[F_m(0; \boldsymbol{d}, d) - F_m(1_3; \boldsymbol{d}, d)], \tag{7.45}$$

1_3 denotes a unit of magnetic flux in the 3 direction. From previous arguments, $\mu(g, d)$ should, for a confining theory, approach zero as d increases. For large g the theory does confine and thus we expect μ to approach zero. This is true even for finite d. For $g=\infty$, $dF_{\mathrm{m}}(\boldsymbol{m};\boldsymbol{d},\boldsymbol{d})$ is the entropy of the system with a restriction on the boundary conditions of the toroidal world. The total number of configurations is independent of these boundary conditions, thus

$$\mu(g=\infty, d)=0, \tag{7.46}$$

and we can restrict ourselves to an evaluation of $\partial\mu/\partial g$. It turns out that it is simpler to evaluate the latter by Monte Carlo techniques. The results are presented in fig. 7.1. Figure 7.1(a) shows the results of the Monte Carlo calculations on 3^4, 4^4 and 5^4 lattices together with smooth curves through the data. These curves are based on theoretical prejudice [7.9] and are replotted in fig. 7.1(b). As the number of

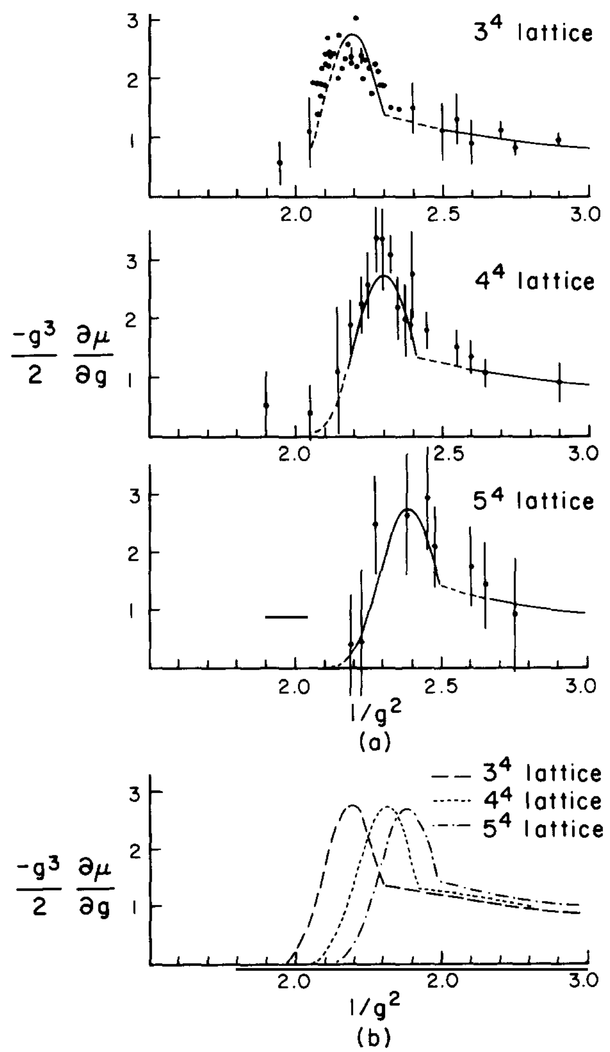


Fig. 7.1. Monte Carlo calculation of the derivative of the free energy differences between configurations with one unit and no units of magnetic flux (a); theoretical fits to the same quantity for various lattice sizes (b).

lattice points increases the region where $\partial\mu/\partial g$ is small (possibly zero) extends towards smaller values of g. Should this pattern persist then $\partial\mu/\partial g$ would vanish in the thermodynamic limit, and as a consequence of (7.46) so would μ itself.

A variational approach consistent with the previous discussions, even though predating it, was Mandelstam's trial vacuum [7.10] consisting of a gas of magnetic monopoles. As we saw in a confining phase electric flux is squeezed into thin tubes. In an ordinary superconductor the opposite phenomenon takes place – magnetic fields are squeezed into vortices (the Meisner effect). In a superconductor it is the free electric charges that form a solenoid around any magnetic flux and keep it from spreading in the transverse direction. Thus we expect that if electric flux is trapped into vortices then there must be a solenoidal current of magnetic monopoles doing this confining. Q.C.D. does have such monopoles, namely the Wu–Yang [7.11] solutions of the static Q.C.D. equations. A vacuum build up of a superconducting gas of such monopoles does have a lower energy than the perturbative one. We do not, however know whether it is the "best" vacuum state, and the renormalization difficulties are severe enough to prevent a systematic improvement on this trial state.

Another thing we have learned from these considerations is what topological structures are necessary to obtain confinement. By this we mean the configurations in a functional integral formulation of QCD that will result in an area law for the Wilson loop. We have no guarantee that these configurations actually saturate the functional integral. The configurations we have in mind are gauge equivalent to $A_\mu = 0$ almost everywhere. Due to certain nonsingle valuedness they do contribute to the Wilson loop, C. We assume that these discontinuities act at isolated points along C. By continuity these points belong to a three dimensional volume bounded by some two dimensional closed surface. There are some, as yet unspecified, currents on these surfaces that generate the potentials A_μ. A prototype expression is

$$A_\mu(x) = \frac{Q}{\pi g}\epsilon_{\mu\nu\lambda\sigma}\int_s \frac{(x-y)_\nu\, ds_{\lambda\sigma}(y)}{(x-y)^4}. \tag{7.47}$$

The integration is over a closed two dimensional surface, s, whose infinitesimal area element is denoted by $ds_{\lambda\sigma}$; Q is the generator of the center of the group, i.e. $\exp 2\pi \mathrm{i} Q \in Z_3$. An elementary calculation yields

$$g\oint_C A\cdot dx = 2\pi Q W[C, s] \tag{7.48}$$

where $W[C, s]$ counts the number of times the curve C winds around the surface s. [(7.45) is a generalization to four dimensions of Gauss' formula for the winding number of two curves in three dimensions.] The contribution to the Wilson loop of this configuration is

$$\frac{\mathrm{Tr}\,P}{3}\exp\left\{\mathrm{i}\oint_C A\cdot dx\right\} = \exp\left\{\frac{2\mathrm{i}\pi}{3} W[C, s]\right\}. \tag{7.49}$$

A winding number 1 (mod 3) is obtained if the surface s and any surface with C as a boundary intersect once. As the size of the curve C increases, "small" surfaces s will do this only if they are located near the periphery C itself. Configurations whose A_μ is generated by such "small" surfaces will contribute terms vanishing as an exponential of the circumference of C. In order to obtain the desired area law surfaces whose size is of the same magnitude as C must contribute [7.12]. Do we expect such large surfaces to be important? The field strength tensor corresponding to (7.45) is

$$F_{\mu\nu}(x) = Q\epsilon_{\mu\nu\lambda\sigma} \int \delta^4(x-y)\, \mathrm{d}s_{\lambda\sigma}/g. \tag{7.50}$$

If we thicken the surface, i.e. renormalize the theory, then the action of this configuration is

$$\mathrm{Tr} \int F_{\mu\nu}F_{\mu\nu}\, \mathrm{d}^4x \sim S/g^2. \tag{7.51}$$

S is the area of the surface. A priori it might appear that large surfaces are suppressed; this suppression might be compensated by the number of surfaces of a given area, namely the entropy of the configurations. We expect the entropy to be at best proportional to S itself and therefore for g sufficiently large, large surfaces are favored and an area law will result. Actually the entropy of free surfaces may grow faster than S (see discussion at the end of section 8.2) and restrictions as to self-intersections may have to be imposed. With such restrictions in force we do not possess techniques to evaluate this entropy and cannot, at present, say whether such large surfaces are important for all g.

What connection do these surfaces have to the field configurations in a box? We saw that magnetic flux tubes are important in a confining vacuum. The time evolution of such tubes are surfaces in four dimensions.

8. Loop field theories

In ordinary nongauge theories the quantities of interest are the Wightman functions of Minkowski space or their continuations to the Euclidian region, the Schwinger functions. These are vacuum expectation values of products of field operators. For gauge theories such products of gauge potentials are not gauge invariant and their expectation values in the true vacuum of the theory may not even exist. Although we did concern ourselves with such products in earlier sections, most of the time we studied gauge invariant operators such as the Wilson loop operator, eq. (2.12). More generally we could study expectation values of the form

$$W[\mathrm{C}_1 \dots \mathrm{C}_n] = \left\langle \mathrm{Tr}\exp\left\{\mathrm{i} \int_{\mathrm{C}_1} A\cdot \mathrm{d}l\right\} \dots \mathrm{Tr}\exp\left\{\mathrm{i} \int_{\mathrm{C}_n} A\cdot \mathrm{d}l\right\}\right\rangle \tag{8.1}$$

with the C_i's arbitrary closed curves. For the gauge group SU(3) operators of the form (see eq. (2.11))

$$U(x,y;\mathrm{C}_1,\mathrm{C}_2,\mathrm{C}_3) = d^{ijk}d^{lmn}U^{il}(x,y;\mathrm{C}_1)U^{jm}(x,y;\mathrm{C}_2)U^{kn}(x,y;\mathrm{C}_3) \tag{8.2}$$

are likewise allowed as are the operators $M(x,y;\mathrm{C})$ of eq. (2.12), if quarks are present. For nongauge theories there are reconstruction theorems [8.1] that permit the determination of the properties of the physical Hilbert space from the Wightman functions. No such theorems exist for the present case. It is tempting to speculate that these gauge invariant expectation values imply an underlying gauge invariant Hilbert space. We shall concentrate on how these functions may be used to establish whether these theories confine. Wilson's criterion for $W[\mathrm{C}]$ is the most natural in this context. For this reason, and for ease of presentation, we will limit our discussion primarily to the $W[\mathrm{C}_1 \dots \mathrm{C}_n]$, and more specifically to just $W[C]$. The dynamics of the theory will force us to look at more involved operators. The equations satisfied by these loop expectation values are very reminiscent of the equations appearing in dual resonance models [8.2].

This approach is a new way of looking at gauge theories. Its main advantage is that it can be formulated in a continuum language; thus Lorentz invariance is at least formally maintained, while at the same time only gauge invariant structures are discussed. If developed further it could combine the best features of lattice theories and continuum theories. At present the mathematics is sufficiently new and untested that we have very little experience in solving the resulting equations, and we cannot isolate the features of the equations that could be responsible for confinement. Similar equations for presumably nonconfining QED would not look too different. All we can say, by combining ideas of this and previous sections, is that if confinement holds then the numbers that arise from the various analyses are consistent with each other. It may be hoped that upon deeper analysis these loop field equations will be able to establish confinement by themselves and even provide a basis for a phenomenology.

8.1. Functional equations for loop operators

Before deriving equations for the expectation values of loop operators we have to discuss the parametrization of curves. A curve C starting at $x_\mu^{(i)}$ and ending at $x_\mu^{(f)}$ is parametrized by a set of points $x_\mu(s)$, with $0 \le s \le 2\pi$, such that $x_\mu(0) = x_\mu^{(i)}$ and $x_\mu(2\pi) = x_\mu^{(f)}$. For a closed curve $x_\mu(0) = x_\mu(2\pi)$. Parametrizations are not unique. If $f(s)$ is any function of s such that $f(s)$ is monotonic with $f(0) = 0$, and $f(2\pi) = 2\pi$, then $x_\mu(f(s))$ is as good a parametrization as the original one. We do not expect a physical quantity such as $W(C)$ to depend on the parametrization. Consider an infinitesimal reparametrization

$$x_\mu(s) \to x_\mu(s + \epsilon(s)) = x_\mu + \epsilon(s)\, \partial x_\mu / \partial s. \tag{8.3}$$

Reparametrization invariance implies

$$W[C(x_\mu(s) + \epsilon(s))] = W[C(x_\mu(s))],$$

which in turn yields

$$\frac{\partial x_\mu(s)}{\partial s} \frac{\delta W[C]}{\delta x_\mu(s)} = 0. \tag{8.4}$$

To obtain dynamical equations for these loop operators, we will draw on our experience with ordinary, nongauge theories. Equations for Wightman functions follow from the equations of motion and equal time commutation relations. They are obtained by comparing such functions at infinitesimally differing points in spacetime. In the present situation we will compare the $W[C]$'s for curves varying infinitesimally from each other [8.3–8.7]. Consider a displacement (fig. 8.1(a))

$$x_\mu(s) \to x_\mu(s) + \epsilon_\mu(s)$$

$$W[C(x_\mu + \epsilon_\mu)] - W[C(x_\mu)] = \int \mathrm{d}s\, \epsilon_\mu(s) \frac{\delta W[C(x_\mu)]}{\delta x_\mu(s)}. \tag{8.5}$$

We now wish to obtain an expression for $\delta W/\delta x_\mu$. Expanding the exponential we find that

$$W[\mathrm{C}] = \sum \frac{(\mathrm{i}g)^n}{n!} \int_0^{2\pi} \mathrm{d}s_1 \ldots \mathrm{d}s_n \,\mathrm{Tr}\,\mathrm{P}\left\langle A_\mu[x(s_1)] \frac{\mathrm{d}x_\mu(s_1)}{\mathrm{d}s_1} \cdots A_\nu[x(s_n)] \frac{\mathrm{d}x_\nu(s_n)}{\mathrm{d}s_n} \right\rangle. \tag{8.6}$$

P denotes the path ordering $\theta(s_1 - s_2) \cdots \theta(s_{n-1} - s_n)$. Performing the variation indicated in (8.5) we find that

$$\delta A_\mu[x(s)] = \frac{\partial A_\mu}{\partial x_\nu} \epsilon_\nu, \qquad \delta \frac{\mathrm{d}x_\mu}{\mathrm{d}s} = \frac{\mathrm{d}\epsilon_\mu}{\mathrm{d}s}. \tag{8.7}$$

Combining these results we obtain

$$\delta A_\mu \frac{\mathrm{d}x_\mu}{\mathrm{d}s} = \frac{\partial A_\mu}{\partial x_\nu} \epsilon_\nu \frac{\mathrm{d}x_\mu}{\mathrm{d}s} + A_\mu \frac{\mathrm{d}\epsilon_\mu}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}(A_\mu \epsilon_\mu) + \left(\frac{\partial A_\mu}{\partial x_\nu} - \frac{\partial A_\nu}{\partial x_\mu}\right) \epsilon_\nu \frac{\mathrm{d}x_\mu}{\mathrm{d}s}. \tag{8.8}$$

The variation of (8.6) yields

$$\delta W = \sum \frac{(\mathrm{i}g)^n}{\cancel{(n-1)!}} \int_0^{2\pi} \mathrm{d}s_1 \cdots \mathrm{d}s_n \,\mathrm{Tr}\,\mathrm{P}\left\langle A_\lambda \frac{\mathrm{d}x_\lambda}{\mathrm{d}s_1} \cdots \delta\left(A_\mu \frac{\mathrm{d}x_\mu}{\mathrm{d}s_i}\right) \cdots A_\nu \frac{\mathrm{d}x_\nu}{\mathrm{d}s_n} \right\rangle. \tag{8.9}$$

Inserting (8.8) into (8.9), the first term on the right hand side of (8.8) can be integrated by parts, resulting in δ-functions from the differentiation of the step functions appearing in P. By using the cyclic property of the trace the result is

$$\frac{\delta W}{\delta x_\nu(s)} = \mathrm{i}g \,\mathrm{Tr}\left\langle\left\{F_{\alpha\nu}[x(s)] \frac{\mathrm{d}x_\alpha}{\mathrm{d}s} U(x(s), x(s); \mathrm{C})\right\}\right\rangle. \tag{8.10}$$

$F_{\mu\nu}$ is the field strength tensor, eq. (2.1). Let us vary this expression once more.

$$\begin{aligned} \frac{\delta^2 W}{\delta x_\nu(s')\, \delta x_\mu(s)} = &-g^2 \left\langle \mathrm{Tr}\left\{\theta(s - s')\, F_{\alpha\nu}[x(s)] \frac{\mathrm{d}x_\alpha}{\mathrm{d}s} \right.\right. \\ &\left.\left. \times U(x(s), x(s'); \mathrm{C})\, F_{\beta\mu}[x(s')] \frac{\mathrm{d}x_\beta}{\mathrm{d}s'} U(x(s'), x(s), \mathrm{C}) + (s \leftrightarrow s')\right\}\right\rangle \\ &+ \mathrm{i}g\, \delta'(s - s') \langle \mathrm{Tr}\{F_{\mu\nu}[x(s)]\, U(x(s), x(s); \mathrm{C})\}\rangle \\ &+ \mathrm{i}g\, \delta(s - s') \left\langle \mathrm{Tr}\left\{\mathrm{D}_\mu F_{\alpha\nu} \frac{\mathrm{d}x^\alpha}{\mathrm{d}s} U(x(s), x(s); \mathrm{C})\right\}\right\rangle. \end{aligned} \tag{8.11}$$

The first term on the r.h.s. of (8.11) comes from varying the path inside the U matrix. The second term is due to a variation of the tangential derivative and the last arises from varying $F_{\alpha\nu}$ itself. D_μ is a covariant derivative and its non-Abelian part comes from varying U at the point $x(s)$.

Before studying (8.11) let us introduce another type of variation of the curve C and a resulting equation that $W(\mathrm{C})$ satisfies. Examining fig. 8.1(a) we see that the previous variation of the curve C consists of adding a "tooth" to C which extends a distance $|\epsilon|$ perpendicular to the curve and runs a distance $|dx/ds|$ along C. The second type of variation consists of adding to C, at the point x, an infinitesimal curve enclosing an area $\delta\sigma_{\mu\nu}$ [8.8, 8.9]. This "keyboard" variation is illustrated in fig. 8.1(b) and may be evaluated by computing $U(\Delta) = \exp\{ig \oint_\Delta A \cdot dl\}$ for infinitesimal curves Δ:

$$U(\Delta) = 1 + ig \oint_\Delta A \cdot dl - g^2 P \oint A \cdot dl \oint A \cdot dl'$$

$$\approx 1 + ig(\partial_\mu A_\nu - \partial_\nu A_\mu)\,\delta\sigma_{\mu\nu} - g^2 \int [A_\mu, A_\nu]\,\delta\sigma_{\mu\nu} = 1 + ig\,F_{\mu\nu}\,\delta\sigma_{\mu\nu}. \tag{8.12}$$

Therefore

$$\frac{\delta W}{\delta\sigma_{\mu\nu}(x)} = ig\,\mathrm{Tr}\langle F_{\mu\nu}(x)\,U(x, x; C)\rangle. \tag{8.13}$$

It is easy to show that

$$\partial_\rho\,\delta W/\delta\sigma_{\mu\nu} = ig\langle \mathrm{Tr}\,D_\rho F_{\mu\nu}\,U(x, x; C)\rangle \tag{8.14}$$

which, except for the δ-function, is the same as the last term of (8.11).

Let us consider for the moment the contracted form of (8.14):

$$\partial_\mu\,\delta W/\delta\sigma_{\mu\nu} = ig\langle \mathrm{Tr}\{D_\mu F_{\mu\nu}\,U(x, x; C)\}\rangle. \tag{8.15}$$

The term $D_\mu F_{\mu\nu}$ appears in the equation of motion; we have to be careful in applying these due to the noncommutativity of various operators. One way of evaluating the above expression is through the path

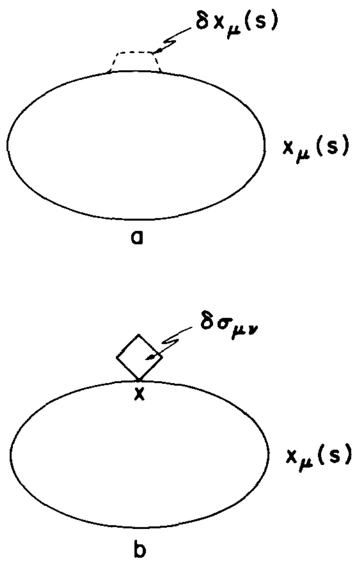


Fig. 8.1. Infinitesimal loop displacement or "tooth" derivative (a); infinitesimal addition to a loop or "keyboard" derivative (b).

integral formalism. We note that

$$\frac{\delta}{\delta A_\nu(x)} \exp\left\{\mathrm{i} \int \mathrm{d}^4x\, L[A]\right\} = \mathrm{i} D_\mu F_{\mu\nu} \exp\left\{\mathrm{i} \int \mathrm{d}^4x\, L[A]\right\} \tag{8.16}$$

and therefore

$$(8.15) = \frac{g}{Z} \int [\mathrm{d}A]\, \mathrm{Tr}\left\{U(x, x; \mathrm{C}) \frac{\delta}{\delta A_\nu(x)}\right\} \exp\left\{\mathrm{i} \int \mathrm{d}^4x\, L[A]\right\}. \tag{8.17}$$

A functional integration by parts gives

$$\begin{aligned}(8.15) &= -g\, \mathrm{Tr}\left\langle \frac{\delta}{\delta A_\nu(x)} U(x, x; \mathrm{C})\right\rangle \\ &= -\mathrm{i}\, g^2 \left\langle \mathrm{Tr}\left\{\int_{\mathrm{C}} \mathrm{d}y_\nu\, \delta(x-y) \frac{\lambda^\alpha}{2} U(x, y; \mathrm{C}) \frac{\lambda^\alpha}{2} U(y, x; \mathrm{C})\right\}\right\rangle. \end{aligned} \tag{8.18}$$

In this form we note that (8.15) is nonzero for self-intersecting curves; we must have a point y on the curve coincident with the reference point x. Do we get a contribution just from this point itself? To answer this question we must be more specific about defining the singular expressions appearing in all the formulas above. For example, the term involving $U(x, x; \mathrm{C})$ should be modified to $U(x+\epsilon, x-\epsilon; \mathrm{C})$ with ϵ a small displacement along C. If C is smooth around x, then this point will not contribute to (8.18), but only genuine self-intersections will. A different regularization scheme [8.6] likewise discards these δ-functions for nonintersecting points. For a curve with one intersection, splitting C into two closed curves C_1 and C_2, one can use the completeness properties of the λ matrices and the unit matrix to simplify

$$\begin{aligned}(8.15) = -\mathrm{i}g \int_{\mathrm{C}} \mathrm{d}y_\nu\, \delta(x-y) \Big\langle \Big[&\frac{8}{9} \mathrm{tr}\, U(x, x; \mathrm{C}_1)\, \mathrm{tr}\, U(x, x; \mathrm{C}_2) \\ &- \frac{1}{12} \mathrm{tr} \frac{\lambda^\alpha}{2} U(x, x; \mathrm{C}_1)\, \mathrm{tr} \frac{\lambda^\alpha}{2} U(x, x; \mathrm{C}_2)\Big]\Big\rangle. \end{aligned} \tag{8.19}$$

With these results we now return to (8.11) and let $s' \to s$ symmetrically from above and below and contract μ and ν. Restricting discussion to nonself-intersecting curves we drop the last term; we also set $\delta'(s-s') = 0$. Thus

$$\frac{\delta^2 W[\mathrm{C}(x(s))]}{\delta x_\mu(s)\, \delta x_\mu(s)} = -g^2 \frac{\mathrm{Tr}}{4} \langle\{F_{\alpha\mu}(x)\{F_{\beta\mu}(x), U(x, x; \mathrm{C}\}_+\}_+\rangle \frac{\mathrm{d}x_\alpha}{\mathrm{d}s} \frac{\mathrm{d}x_\beta}{\mathrm{d}s}. \tag{8.20}$$

Written in this double commutator form the operator ordering is correct and the result holds for arbitrary matrix elements not just vacuum expectation values. Between (8.20) and (8.19) we have the analogs of the field equations and Ward identities for loop theories. The solutions of these equations may establish confinement. A study of the solutions of (8.15) for theories based on the group SU(N), in the limit of large N has been made [8.9].

8.2. Solutions of approximate loop equations

Equation (8.20) relates $W[C]$ to a more complicated gauge invariant quantity. As in ordinary field theory, the complete set of equations is infinite. In order to obtain some information, we must truncate this set by some approximation. The form we hope to maneuver (8.20) into is

$$\frac{\delta^2 W[C]}{\delta x_\mu(s)\,\delta x_\mu(s)} = \sigma^2\left(\frac{dx}{ds}\right)^2 W[C] + \text{(self-intersection terms)}. \tag{8.21}$$

σ will turn out to be the string tension discussed previously.

Equation (8.21) has appeared in the context of dual resonance theories [8.2], where σ is equal to $1/2\pi\alpha'$, with α' the slope of Regge trajectories. With $\alpha' = 0.9\,\text{GeV}^{-2}$ we get a reasonable agreement with the value of σ given in eq. (2.15). We shall not pursue the relation with the dual model further, but give the argument for the relation of σ in (8.21) to the string tension.

General solutions of (8.21) do not exist. Certain special solutions all point to the above interpretation of σ. In two dimensions we have a solution of the form

$$W_2[C] = \text{cnst}\exp\left[-\frac{\sigma}{2}\oint_C \epsilon_{\mu\nu}x_\mu\, dx_\nu\right] = \text{cnst}\exp[-\sigma A], \tag{8.22}$$

with A denoting the area enclosed by C. Upon comparing this with the Wilson criterion, eq. (2.20), we obtain the sought for interpretation of σ. It is plausible to extend this argument to planar curves in four dimensions. In four dimensions special solutions have been found [8.6, 8.10]. Let

$$\sigma_{\mu\nu} = \frac{1}{2}\int x_\mu\, dx_\nu$$

and

$$A_\pm = \sqrt{(\sigma_{\mu\nu} \pm \epsilon_{\mu\nu\alpha\beta}\sigma_{\alpha\beta})^2},$$

then

$$W_4[C] \sim \frac{1}{A_\pm^{1/2}} K_{1/2}(\sigma A^\pm) \tag{8.23}$$

is a solution. For planar curves $A_\pm$ is just the area enclosed by C. Asymptotically in A, W has the correct behavior for a confining theory. Appealing to results from the theory of relativistic strings [8.2], and some mild technical assumptions, the confining behavior of the solutions of (8.21) can be demonstrated [8.11]. Corrections to the leading behavior of $W[C]$ have also been obtained [8.12]. This last study makes progress towards some of the renormalizability questions inherent in (8.21).

Is there any approximation to (8.20) which might bring it to the form of (8.21)? A tempting one is to replace the term $\{F_{\alpha\mu}, F_{\beta\mu}\}_+$ by its vacuum expectation value [8.13]. Comparing with (8.21) we would obtain

$$\sigma^2 = -\frac{g^2}{12}\text{Tr}\langle F_{\mu\nu}F_{\mu\nu}\rangle. \tag{8.24}$$

This result is inconsistent. The left hand side of (8.24) is positive. In a Euclidian formulation the right hand side is intrinsically negative. In a Minkowski formulation $F_{\mu\nu}F^{\mu\nu} = B^2 - E^2$; from the discussion in section 7 we expect that in a confining phase the vacuum is primarily magnetic and again the right hand side of (8.24) is negative.

Section 7 provides a clue to a more reasonable approximation. The Wilson loop operator creates a unit of electric flux along the curve C. The right hand side of (8.20) is approximated in the following way:

$$-g^2 \frac{\mathrm{Tr}}{4} \langle F_{\alpha\mu}F_{\beta\mu}U + 2F_{\alpha\mu}UF_{\beta\mu} + UF_{\alpha\mu}F_{\beta\mu}\rangle \frac{\mathrm{d}x_\alpha}{\mathrm{d}s}\frac{\mathrm{d}x_\beta}{\mathrm{d}s} \approx -\frac{g^2}{6}\mathrm{Tr}\left\langle\!\!\left\langle F_{\alpha\mu}\frac{\mathrm{d}x_\alpha}{\mathrm{d}s}F_{\beta\mu}\frac{\mathrm{d}x_\beta}{\mathrm{d}s}\right\rangle\!\!\right\rangle W[\mathrm{C}]. \tag{8.25}$$

The notation $\langle\langle\ \ \rangle\rangle$ denotes an expectation value in a state containing a unit flux along C. In order to obtain a finite electric field, the flux must be spread out over a tube of finite thickness. Take the cross sectional area of such a tube to be A, then with the help of (7.10) we find that inside the tube

$$(AE^{(3)}\,\mathrm{d}x/\mathrm{d}s)^2 = (AE^{(8)}\,\mathrm{d}x/\mathrm{d}s)^2 = \tfrac{1}{6}g^2(\mathrm{d}x/\mathrm{d}s)^2. \tag{8.26}$$

The above may be obtained by placing a hypothetical quark–antiquark singlet at the ends of a straight section of flux tube. Combining (8.25) with (8.26) yields

$$\sigma = g^2/6A. \tag{8.27}$$

Using a related but somewhat different procedure, Nambu [8.3] obtained the same equation. It is interesting that (8.27) can be obtained from a different picture. In order for a quark–antiquark potential to be a linear confining one, the electric flux between them must be confined to a tube of fixed area as the interquark distance, R, increases. The electrical energy stored in such a configuration is

$$\tfrac{1}{2}E^\alpha E^\alpha AR = (g^2/6A)R, \tag{8.28}$$

yielding a σ identical with that of (8.27). In view of the cavalier nature of the arguments leading to (8.25) this agreement is fortuitous.

Let us return for a moment to eq. (8.24). Although the sign is wrong, it is interesting to note that a calculation of this quantity using the instanton approximation to the path integral yields an acceptable value for the absolute value of σ [8.13, 8.14]. In section 6.1 we saw that instantons were expelled by electric fluxes. A reinterpretation of (8.24) is

$$\sigma = -\frac{g^2}{12}[\langle\langle \mathrm{Tr}\, F^2\rangle\rangle - \langle \mathrm{Tr}\, F^2\rangle], \tag{8.29}$$

where $\langle\langle F^2\rangle\rangle$ is the value of the F^2 due to instantons, inside the flux tube, namely zero. This gives the correct sign of σ.

Before closing this section it is worthwhile to point out some dangers in this whole approach. First, all these equations involve very singular operator products and a meaning has to be given to them. Some progress in this direction has been made [8.5, 8.9, 8.12]. Second, the approximate form, (8.21), has serious difficulties. Is there a theory that yields (8.21) as an exact loop equation? Formally, one can write down such a theory [8.15] but due to the large number of configurations, none of the expectation values exist, i.e., $W[C]$ is infinite. This is true even though the theory is constructed on a finite lattice. These infinities have nothing to do with difficulties in the ultraviolet or infrared, but rather with the number of configurations in loop space growing too rapidly.

9. Deconfinement at high temperatures

Should QCD truly confine quarks, there are good theoretical arguments that at finite temperatures quarks will be liberated. The critical temperature, T_c, above which this phenomenon occurs is of the order of hadronic energies, or of the order of 10^{12} °K. The interest in high temperature QCD is practical, pedagogical and philosophical.

On the practical side, the existence of a phase of free quarks may have a significant impact on astrophysics and the physics of dense nuclear matter. (A proper treatment of the latter should include a finite quark chemical potential.) In standard Big Bang cosmologies the temperature during the first 10^{-6} seconds exceeds T_c [9.1]. In the interiors of neutron stars or for very short times in heavy ion collisions, high effective temperatures may exist.

From a pedagogical point of view the nature of this transition emphasizes the importance of the center of the group in the discussion of confinement. At zero temperature the importance of the group center was discussed in section 7.

The third reason is, perhaps, somewhat subjective. It is interesting that the variables we use to describe the problem may, under certain circumstances, have direct physical manifestation. An analogy can be drawn with superconductivity. If the critical temperature for this phenomenon had been much higher, so that the superconducting state had been more common than the normal one, the idea that the electric current is carried by a Cooper pair of hypothetical electrons might have been suspect.

The demonstration that gauge theories cease to confine above a certain temperature is carried out on the strong coupling lattice version of this theory [9.2, 9.3]. It is just in this region that we are confident that the theory confines at zero temperature. Corrections due to magnetic terms tend to deconfine the theory without any help from finite temperature effects. Therefore, if we find a temperature above which the theory ceases to confine in the very strong coupling limit, we are confident it will do so, even at a lower temperature in the full theory. We will first discuss the thermodynamics of lattice gauge theories and then show the deconfinement in the strong coupling limit. Instanton effects at finite temperature have also been studied [9.4].

9.1. Thermodynamics of lattice gauge theories

The prime differences between a field theory at finite temperature and one at zero temperature is that rather than the expectation values of products of operators in the vacuum state we are interested in a weighted sum of expectation values in all the eigenstates of the theory. We define a partition function

$$Z = \mathrm{Tr}\, \mathrm{e}^{-\beta H} \tag{9.1}$$

with H the full Hamiltonian of the theory. β is the inverse temperature, and the thermal average of an operator O is

$$\langle O\rangle = Z^{-1}\,\mathrm{Tr}\, O\, \mathrm{e}^{-\beta H}. \tag{9.2}$$

We will first discuss gauge theories on a lattice with no external sources. The Hamiltonian is given in (3.15). The dynamical variables are the link variables $U_{x;\hat{e}}$ of (3.6). To an eigenstate n of H corresponds an eigenfunction $\psi_n[U_{x;\hat{e}}]$. The physical eigenstates are invariant under gauge transformations:

$$\psi_{\text{physical}}[U_{x;\hat{e}}] = S[g_x]\psi_{\text{physical}}[U_{x;\hat{e}}] = \psi_{\text{physical}}[g_x^+ U_{x;\hat{e}} g_{x+\hat{e}a}]. \tag{9.3}$$

$S[g_x]$ is a gauge transformation depending on a position dependent set of group elements g_x. The partition function is

$$Z = \sum_{n\text{-physical}} \int \prod_{x;\hat{e}} \mathrm{d}U_{x;\hat{e}} (\psi_n^*[U_{x;\hat{e}}] \exp(-\beta E_n)\, \psi_n[U_{x;\hat{e}}]). \tag{9.4}$$

The restriction of the summation to gauge invariant states is inconvenient and we may eliminate it by introducing a projection operator:

$$\begin{aligned} Z &= \sum_n \int \prod_{x;\hat{e}} \mathrm{d}U_{x;\hat{e}}\, \mathrm{d}g_x (\psi_n^*[U_{x;\hat{e}}] \exp(-\beta E_n)\, S[g_x]\, \psi_n[U_{x;\hat{e}}]) \\ &= \sum_n \int \prod_{x;\hat{e}} \mathrm{d}U_{x;\hat{e}}\, \mathrm{d}g_x (\psi_n^*[U_{x;\hat{e}}] \exp(-\beta E_n) \psi_n[g_x^+ U_{x;\hat{e}} g_{x+\hat{e}a}]). \end{aligned} \tag{9.5}$$

The thermodynamic interpretation of the partition function is

$$Z = \exp\{-\beta F(\beta)\}, \tag{9.6}$$

with $F(\beta)$ the free energy at a temperature corresponding to β. It plays the same role as the vacuum energy at zero temperature.

The presence of fixed external sources modifies the above. Consider a source transforming as the q representation located at point $\boldsymbol{R}$ and another transforming as the $\bar{\mathrm{q}}$ representation located at the origin. (For quarks, q is the 3 dimensional representation.) Under a gauge transformation

$$\psi_{\alpha\beta}[U_{x;\hat{e}}] \to U^{(\mathrm{q})^+}_{\alpha\alpha'}(g_0)\, U^{(\mathrm{q})}_{\beta\beta'}(g_{\boldsymbol{R}})\, \psi_{\alpha'\beta'}[g_x U_{x;\hat{e}} g_{x+\hat{e}a}], \tag{9.7}$$

and

$$\begin{aligned} Z(R) = \sum_n \prod_{x,\hat{e}} \mathrm{d}U_{x;\hat{e}}\, \mathrm{d}g_x \\ \times \{\psi^*_{n;\alpha\beta}[U_{x;\hat{e}}] \exp(-\beta E_n) [U^{+(\mathrm{q})}_{\alpha\alpha'}(g_0)\, U^{(\mathrm{q})}_{\beta\beta'}(g_R)\, \psi_{n;\alpha'\beta'}[g_x U_{x;\hat{e}} g_{x+\hat{e}a}]\}. \end{aligned} \tag{9.8}$$

The free energy of a configuration with a source at the origin and at $\boldsymbol{R}$ relative to that of the vacuum is

$$\exp\{-\beta[F(\beta; R) - F(\beta)]\} = Z^{-1} Z(R). \tag{9.9}$$

The effective potential energy of the two sources at a temperature β^{-1} is

$$E(\beta, R) = -\frac{1}{\beta}\ln\{Z(R)/Z\}. \tag{9.10}$$

We will show that at a certain temperature $E(\beta, R)$ ceases to have a linear growth in R.

9.2. *Strong coupling approximation at finite temperatures*

The strong coupling limit provides the simplification of explicitly knowing the wave functions and energies. The Hamiltonian separates into a sum of uncoupled link Hamiltonians as in (3.18) with $x = 0$. The wave functions are indicated in (3.17). In this approximation (9.5) is

$$Z = \int \prod_{x;\hat{e}} \mathrm{d}U_{x;\hat{e}}\, \mathrm{d}g_x \sum_{r,p,q} \frac{1}{d(r)} D^{*(r)}_{pq}[U_{x;\hat{e}}] \exp\left\{-\frac{\beta g^2}{2a} C^{(2)}(r)\right\} D^{(r)}_{pq}[g^+_x U_{x;\hat{e}} g_{x+\hat{e}a}]. \tag{9.11}$$

The orthogonality relations of group representations [3.3] permit a free integration over the link variables resulting in

$$Z = \int \prod_{x;\hat{e}} \mathrm{d}g_x \exp\left\{-\frac{\beta g^2}{2a} C^2(r)\right\} \chi^{*(r)}(g_x)\, \chi^{(r)}(g_{x+\hat{e}a}). \tag{9.12}$$

$\chi^{(r)}(g)$ is the character of the representation r, i.e. $\chi^{(r)} = \operatorname{tr} D^{(r)}$. We have reduced the partition function for a gauge system to that of a nearest neighbor coupling spin system. At each site we have a spin matrix $g(x) \in \mathrm{SU}(3)$ with a complicated interaction given in (9.12). Z is invariant under the discrete transformation

$$g \to \mathrm{e}^{\pm 2\pi\, \mathrm{i}/3}\, g, \tag{9.13}$$

namely, it is invariant under the center Z_3. In this approximation (9.8) becomes

$$Z(R) = \int \prod_{x;\hat{e}} \mathrm{d}g_x \exp\left\{-\frac{\beta g^2}{2a} C^{(2)}(r)\right\} \chi^{*(r)}(g_x)\, \chi^{(r)}(g_{x+\hat{e}a})\, \chi^{*(q)}(g_0)\, \chi^{(q)}(g_R) \tag{9.14}$$

or

$$E(\beta, R) = -\frac{1}{\beta}\ln\langle \chi^{*(q)}(g_0)\, \chi^{(q)}(g_R)\rangle. \tag{9.15}$$

If the Z_3 symmetry is not broken, then for $q = 3, \bar{3}, 6$ or $\bar{6}$

$$\langle \chi^{q}\rangle = 0 \tag{9.16}$$

and as $R \to \infty$ usual statistical mechanics arguments tell us that the correlation function behaves as

$$\langle \chi^{*(q)}(g_0)\, \chi^{(q)}(g_R)\rangle \underset{R\to\infty}{\sim} \mathrm{e}^{-\beta\sigma(\beta)R}, \tag{9.17}$$

and we may interpret $\sigma(\beta)$ as the temperature dependent interquark force. If the symmetry is broken, then

$$\langle\chi^{(q)}\rangle = M \neq 0 \tag{9.18}$$

and

$$\langle\chi^{*(q)}(g_0)\,\chi^{(q)}(g_R)\rangle \underset{R\to\infty}{\sim} |M|^2\,(1 - C(R)\,\mathrm{e}^{-\mu R}). \tag{9.19}$$

$C(R)$ is slowly varying. The energy of this two quark configuration is

$$E(\beta, R) = -\frac{2}{\beta}\ln|M| + \frac{C}{\beta}\,\mathrm{e}^{-\mu R}. \tag{9.20}$$

$-\ln|M|/\beta$ has the interpretation of a quark self energy and the second term provides a short range Yukawa potential between the liberated quarks.

We shall now argue that at sufficiently high temperatures the symmetry is broken. At first sight it is surprising that a symmetry should be broken at high temperatures. It is at low temperatures that ordered states occur. However, the temperature of the effective spin system of (3.12) is the reciprocal of the temperature of the gauge field theory. For β small (large temperature) the most important contribution to the partition function (9.12), comes from $\chi(g_x)$'s close to each other. In fact for $\beta = 0$, due to the completeness of the group characters on functions constant on the various characteristic classes of the group, Z is a "δ-function" of the group characters; for very large β only the trivial representation contributes, and the g_x's are relatively unconstrained. Effectively

$$Z = \int \prod_{x;\hat{e}} \mathrm{d}g_x \exp\left\{-\frac{a}{g^2\beta}\, V[\chi(g_x), \chi(g_{x+\hat{e}a})]\right\}. \tag{9.21}$$

The potential has a minimum when the two variables are equal. We see that the role of the temperature is the inverse of that of the gauge system. At low spin temperatures (small β) this system undergoes a phase transition to an ordered state and the situation envisioned in (9.18) occurs.

At what temperatures does this phase transition take place? From (9.12) or (9.21) the only parameter is $g^2\beta/a$ so we expect $T_c \sim g^2/a$. As the only dimensional physical parameter is the string constant, the full theory temperature will be

$$T_c = C\sqrt{\sigma} \tag{9.22}$$

with C of order unity (a more refined analysis finds C to be $\frac{1}{2}$ [9.4]) yielding the temperature mentioned at the beginning of this section. A Monte Carlo calculation confirms this idea [9.5].

What is the origin of this transition? In the strong coupling limit the eigenstates of the gauge system are closed flux contours. Neglecting interactions, a contour of L links contributes an energy

$$E_L = \frac{g^2}{2a}\frac{4}{3}L. \tag{9.23}$$

The partition function becomes roughly

$$Z = \sum_L N(L) \exp\left(-\frac{\beta g^2}{2a}\frac{4}{3}L\right). \tag{9.24}$$

$N(L)$ is the number of contours of length L. This number is easily evaluated [9.6]

$$N(L) = \exp[cL]. \tag{9.25}$$

For $T > 2g^2/(3ca)$, the system contains a condensate of flux contours of large lengths (in fact, in order to stabilize this system we must include interactions among intersecting strings). A quark–antiquark can attach themselves to a large contour and separate over a large distance.

10. "Unsolved" and speculative problems

Should all the previous hopes and approximations persist to higher orders in strong coupling expansions, better Monte Carlo calculations or even analytic arguments, and the assumption that in QCD with no dynamic quarks, non-singlets are confined, there are still some open problems. One is of immediate importance, namely, the inclusion of light quark fields, in addition to possibly static sources. The second question we will touch upon is more speculative, namely, how much of the foregoing discussion could survive the discovery of free quarks. We shall address these questions in turn.

10.1. Inclusion of dynamical quarks

Intuitively, the introduction of quark fields with a small, or even zero, mass term, implying pair annihilation and creation, does not change our criterion for confinement. Nonconfinement or confinement is equivalent to the existence or nonexistence of isolated color singlets. None of the previous tests to which we subjected the theory survive the presence of dynamic quarks and thus cannot be used to determine whether the theory confines. The simple reason is that if we try to pull a quark and an antiquark apart to see whether the separation potential increases with distance, at some point it pays for the system to create another quark–antiquark pair out of the vacuum, with the original quark attached to the new antiquark and vice versa. The two parts of the system go off as mesons (bound q–q̄ systems), with no further increase in energy. We shall systematically investigate how the various approaches fail to establish confinement in the presence of quark fields.

The obvious modification dynamic quarks make is the addition of

$$L_Q = i\bar{q}\gamma^\mu D_\mu q - m\bar{q}q \tag{10.1}$$

to the QCD Lagrangian. One lattice transcription of (10.1) is

$$L_Q = \sum_{\hat{e}} \bar{q}(x + ae)\, \gamma_{\hat{e}} U_{x+ae} q(x) - m\bar{q}(x)\, q(x). \tag{10.2}$$

Actually, this lattice version, for fermions, does not have the classical limit implied by (10.1) but reduces to a classical theory with more degrees of freedom [3.1]. Other lattice Lagrangians have been introduced [10.1]; for the purposes of the present discussion (10.2) is sufficient.

The lattice Hamiltonian approach fails for the reason mentioned earlier. The lowest energy state of a static quark pair separated by a large distance is not the one with a flux line connecting the sources, but with an extra dynamic quark–antiquark pair superimposed on the static ones. To low orders in the strong coupling expansion, the energy is independent of separation. The expectation value of the Wilson loop operator is always proportional to the exponential of the circumference. Again, the lowest order in the strong coupling expansion does not come from filling up the area enclosed by the loop with fundamental plaquettes as in fig. 3.3(a), but rather bringing down terms of the form $(\bar{q}(x+\hat{e}a)\,U\,q(x))$ from the action. These terms are needed only along the curve c itself and thus the lowest order contribution to $W[\mathrm{c}]$ is

$$\ln W[\mathrm{c}] \sim P[\mathrm{c}]/a, \tag{10.3}$$

with $P[\mathrm{c}]$ the circumference of c.

Similar arguments can be given for the instanton contribution to the quark potential. In (6.16) a quark–antiquark pair shield the external field $F^{\mathrm{ext}}_{\mu\nu}$ and the effective coupling does not grow with separation.

We do not expect any infrared singularities in the quark propagator; if this propagator can be defined at all it will probably have no poles or cuts, as by unitarity these would correspond to physical color triplet states. The inclusion of quark loops into the Dyson–Schwinger equation would eliminate a $1/k^4$ singularity the pure QCD theory might produce.

The duality arguments cannot be put forward as the quark fields transforming as color triplets are not invariant under gauge transformations characterizing a magnetic vortex. An electric flux can be undone by a quark pair. The presence of quarks invalidates the approximate loop equation as in (8.11) and (8.15) as there will be extra terms generated by $\mathrm{D}_\mu F_{\mu\nu}$.

We see that all previous approaches fail to distinguish confinement from nonconfinement. In the presence of dynamical quarks there may be no fundamental difference between a confining phase and a nonconfining Higgs region [7.4, 10.2]. The only approach left is the one based on the original observation that only integrally charged objects, or more generally, only SU(3)-flavor triality zero objects exist. The quark Lagrangian must have at least a global Z_3 symmetry, independent of the Z_3 subgroup of SU(3). We may then ask whether any states transforming nontrivially under this Z_3 have finite energy. A negative answer would imply confinement. To date no work in this direction has been done.

10.2. What if free quarks are found?

How much of the previous presentations can survive the discovery of free quarks?[7] We know that if such objects exist they are heavy and/or not produced very readily. Some experimental limits are discussed in section 1. In fig. 10.1 we show two possible interquark energies that ultimately lead to free quarks, but are consistent with a linear energy of separation at short distances. Case (a) in fig. 10.1 leads

[7] Part of the motivation for this section has been the author's fear that upon completion of this review positive evidence for free quarks will be presented.

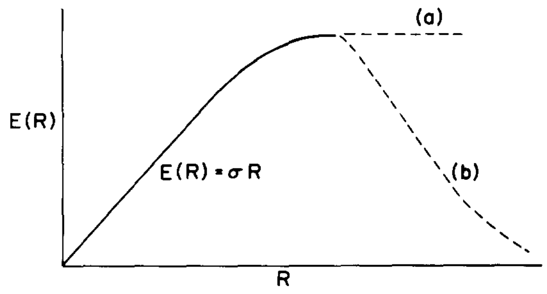


Fig. 10.1. The quark separation energies for two nonconfining scenarios consistent with a linear rising potential at short distances.

to deconfined but heavy quarks. A large amount of energy must be supplied to pull the quark–antiquark pair in a meson to a certain radius of separation. After that, no more energy is needed to pull them further apart. The total energy supplied (minus the mass of the initial meson) is the rest mass of the free quark pair. In case (b) this energy is recovered and the quark mass is small. Quarks can escape from hadrons only by tunneling or by acquiring a large amount of energy to drive the system over the potential barrier.

What modifications, and what could be the possible causes of such modifications, have to be made to the previous discussions to obtain energies of separations as in fig. 10.1? For this speculation we will use the formulation based on eqs. (6.25) to (6.28). The energy of separation (interquark potential plus self energies) is

$$E(\boldsymbol{r}) = \frac{1}{(2\pi)^3}\int d\boldsymbol{k}(1 - e^{i\boldsymbol{k}\cdot\boldsymbol{r}})\,\Delta_{00}(k^2)$$
$$= \frac{1}{2\pi^2}\int_0 dk\left(1 - \frac{\sin kr}{kr}\right)k^2\Delta_{00}(k^2). \tag{10.4}$$

To recapitulate section 6.2, a $1/k^4$ singularity in $\Delta_{00}(k^2)$ yields a linear, confining potential. We shall consider two modifications of $\Delta_{00}(k^2)$ near $k^2 = 0$. First, take

$$\Delta_{00}(k^2) \sim 1/(k^2 + \mu_1^2)(k^2 + \mu_2^2) \tag{10.5}$$

then

$$E(r) \sim \frac{1}{\mu_1 + \mu_2} + \frac{1}{r}[e^{-\mu_1 r} - e^{-\mu_2 r}]\frac{1}{\mu_1^2 - \mu_2^2}. \tag{10.6}$$

For $r < \min(1/\mu_1, 1/\mu_2)$ the potential is linear. Adjusting the constant of proportionality of this term to be the string tension σ the mass of the quark is

$$M_Q = \sigma/(\mu_1 + \mu_2). \tag{10.7}$$

In the $\mu_{1,2} \to 0$ limit we recover infinite mass quarks. The masses $\mu_{1,2}$ could arise from a Higgs type symmetry breaking giving the gluons a very small mass. A detailed analysis of such a picture based on the MIT bag model has been presented previously [10.3]. Results similar to (10.7) were obtained. For distances $r < 1/\mu_{1,2}$ much of the confinement pictures hold, as for example the existence of a flux string attached to a quark. This string is not infinite in length but extends to $r \sim 1/\mu$ leading to (10.7).

Case (b) (of fig. 10.1) may be obtained from a situation where $\int \mathrm{d}\boldsymbol{k}\, k^2 \Delta_{00}(k^2) = 0$. $E(r)$ would approach zero and the quark mass would have no large infrared component. An example is

$$\Delta_{00}(k^2) \sim \frac{1}{k^2}\left[\frac{1}{(k-\mathrm{i}\mu)^2} + \frac{1}{(k+\mathrm{i}\mu)^2}\right] \tag{10.8}$$

yielding

$$E(r) \sim \frac{\sigma}{r\mu^2}[1-(1+r\mu)\,\mathrm{e}^{-\mu r}]. \tag{10.9}$$

Again for $r < 1/\mu$ we reproduce the previous pictures. Beyond $1/\mu$ the energy of separation decreases and at large distances we find light mass quarks. By choosing μ sufficiently small both the energy required to knock out quarks from a hadron or have them tunneled through a hadron may be made large.

Even if free quarks are conclusively observed, the confinement picture could be valid for distances smaller than those characterizing deconfinement.

Appendix A. Functional integral formulation of gauge theories

In this appendix we develop the equivalence between the lattice Hamiltonian gauge theory and the Euclidian path integral formulation. Our starting point is the lattice Hamiltonian, (3.15). Ignoring the quark fields we may write this quantity as

$$\begin{aligned} H &= H_{\mathrm{E}} + H_{\mathrm{M}} \\ H_{\mathrm{E}} &= \frac{g^2}{2a}\sum_{\text{links}} E^2_{x;\hat{e}} \\ H_{\mathrm{M}} &= \frac{\mathrm{Tr}}{ag^2}\sum_{p}(2 - U_p - U_p^+). \end{aligned} \tag{A.1}$$

As discussed in section 3, the eigenstates of E^2 are denoted by $|r; p, p'\rangle$ with r referring to a particular representation of the group

$$E^2|r; p, p'\rangle = C^{(2)}(r)\,|r; p, p'\rangle, \tag{A.2}$$

and there are $d^2(r)$ ($d(r)$ is the dimensionality of the representation) states for each r. The eigenstates of H_{M} are labeled by a group element, $|U(b)\rangle$, and the overlap between these states is given by (3.17). If no ambiguity arises, we shall, for notational simplicity, write

$$\begin{aligned} |U(b)\rangle &= \prod_{x,\hat{e}} |U(b_{x,\hat{e}})\rangle \\ |r\rangle &= \prod_{x,\hat{e}} |r_{x,\hat{e}}; p_{x,\hat{e}}; p'_{x,\hat{e}}\rangle. \end{aligned} \tag{A.3}$$

For the moment, let us ignore Gauss' law. The object of interest for which we will develop a path integral is

$$\langle U(b')|\mathrm{e}^{-TH}|U(b)\rangle. \tag{A.4}$$

In order to evaluate this we rewrite [2.2]

$$\mathrm{e}^{-TH}=\prod_{j=1}^{N}\mathrm{e}^{-\epsilon H} \tag{A.5}$$

with $\epsilon\sim T/N$ and small. By neglecting terms of order ϵ^2 we ignore the noncommutativity of H_{E} and H_{M} and write

$$\mathrm{e}^{-\epsilon H}=\mathrm{e}^{-\epsilon H_{\mathrm{E}}}\,\mathrm{e}^{-\epsilon H_{\mathrm{M}}}. \tag{A.6}$$

As the states $|U(b)\rangle$ form a complete set, we can rewrite (A.4) using (A.5),

$$(\mathrm{A}.4)=\int \mathrm{d}U_1\cdots\mathrm{d}U_N\prod_{i=0}^{N}\langle U(b_{i+1})|\mathrm{e}^{-\epsilon H}|U(b_i)\rangle \tag{A.7}$$

with $b_0=b$ and $b_{N+1}=b'$. We can use (A.6) to help us evaluate each of the terms in (A.7)

$$\begin{aligned}\langle U(b_{i+1})|\mathrm{e}^{-\epsilon H}|U(b_i)\rangle&=\sum_{r,p,p'}\langle U(b_{i+1})|\mathrm{e}^{-\epsilon H_{\mathrm{E}}}|r\rangle\langle r|\mathrm{e}^{-\epsilon H_{\mathrm{M}}}|U(b_i)\rangle\\&=\sum_r\left\{\exp\left[\frac{-\epsilon g^2}{2a}\sum_{\text{links}}C^{(2)}(r_{\text{link}})-\frac{\epsilon\,\mathrm{Tr}}{2g^2a}\sum_p(2-U(b_{p,i})-U^{+}(b_{p,i}))\right]\prod_{\text{links}}\mathrm{Tr}\,D^{(r)}(U^{+}(b_{i+1})\,U^{(r)}(b_i))\right\}.\end{aligned} \tag{A.8}$$

Define

$$\tilde{e}[U]=\sum_r\exp\left\{\frac{-\epsilon g^2}{2a}C^{(2)}(r)\right\}\mathrm{Tr}\,D^{(r)}[U]. \tag{A.9}$$

We obtain

$$(\mathrm{A}.4)=\int\mathrm{d}U_1\cdots\mathrm{d}U_N\prod_{i=0}^{N}\left\{\left[\prod_{\text{links}}\tilde{e}(U^{+}(b_{i+1})\,U(b_i))\right]\exp\left[-\frac{\epsilon\,\mathrm{Tr}}{g^2a}\sum_p(2-U(b_{p,i})-U^{+}(b_{p,i}))\right]\right\}. \tag{A.10}$$

In the $\epsilon\to 0$ limit (A.10) is equivalent to (A.4). At this stage we have both a discrete space and a discrete time. (A.10) however, is not symmetric between these two directions. To obtain a symmetric expression we must find an approximation for $\tilde{e}$. This approximation will be valid in the continuum time limit, $\epsilon\to 0$. As for many of the calculations we discuss we will want to keep ϵ finite; the approximation will agree with (A.10) only in the small g limit. We believe that the continuum quantum mechanics based on (A.10), or the approximation we shall derive, is the same.

Let $L_\alpha^{(r)}$ be the representation of dimension $d(r)$ of the generators of the Lie algebra of SU(3). Then, if $U = \exp(\mathrm{i}L \cdot b)$

$$\tilde{e}[U] = \sum_r \mathrm{Tr}\exp\left\{-\frac{\epsilon g^2}{2a}(L(r))^2 + \mathrm{i}b \cdot L^{(r)}\right\}. \tag{A.11}$$

We are interested in the limit ϵ or g small as well as U close to unity; we do not expect nearby lattice variables to differ significantly from each other. In this situation the L's may be treated as classical variables with the trace and summation over representations replaced, up to an irrelevant normalization, by an integration over L. Thus

$$\tilde{e}[U] \approx \exp\left\{-\frac{a}{2\epsilon g^2} b^2\right\}. \tag{A.12}$$

We wish to extend this approximation to U away from unity, or b away from zero. We also wish to maintain the property that $\tilde{e}[U]$ does not change if b is transformed into itself along any closed orbit of the group. This is accomplished by setting

$$b^2 \approx 2\,\mathrm{Tr}(2 - D^{(3)}(b) - D^{(3)+}(b)). \tag{A.13}$$

Combining these results we find the final approximation to[8]

$$\tilde{e}[U] \approx \exp\left\{-\frac{a}{\epsilon g^2}\mathrm{Tr}(2 - U - U^+)\right\}. \tag{A.14}$$

The group variable appearing in the functions $\tilde{e}$ in (A.10) is the product of two link variables separated by one unit in the time direction. The gauge $A_0 = 0$ corresponds, in the lattice language, to setting the link variables for links parallel to the time direction equal to unity. Keeping this restriction in mind, one may write $U^+(b_{i+1})\,U(b_i)$ as U_p for a plaquette in a time-space plane. Choosing the time step ϵ to be the same as the lattice spacing results in a symmetric approximation to (A.4):

$$(\mathrm{A.4}) \approx \int \prod_{x;\hat{e}} \mathrm{d}U_{x;\hat{e}}\, \delta(U_{x,\hat{t}} - 1)\exp\left\{\frac{1}{g^2}\mathrm{Tr}\sum_p (U_p + U_p^+)\right\}. \tag{A.15}$$

In the above we dropped an overall factor of $\exp\{(1/g^2)\,\mathrm{Tr}\,2\}$ for each plaquette (cf. (A.10) and (A.13)). A boundary condition fixing the link variables at $t = 0$ and $t = T$, as given initially in (A.4), is implied.

We now return to the task of enforcing Gauss' law. Remember that this is the same as the requirement that all of our configurations are invariant under gauge transformations. This may be accomplished by performing all possible gauge transformations at each separate time slice. This amounts to transforming each link variable $U_{x;\hat{e}}$ ($\hat{e}$-spacial) to $g^{-1}(x)\,U_{x;\hat{e}}\,g(x + \hat{e})$. The g's are group elements and we integrate each over the whole group. The above transformation has no effect on spacial plaquettes. Its effect on plaquettes parallel to the time direction is to change the timelike links

[8] The relation between (A.14) and (A.9) is the analogue, for non-Abelian theories, of the Villain approximation [A.1] in Abelian spin or gauge theories.

from unity to $g^{-1}(x)\,g(x+\hat{t})$. This combination can be interpreted as a link variable $U_{x,\hat{t}}$ for timelike links. Changing variables from the g's to the $U_{x,\hat{t}}$'s amounts to rewriting (A.15) without the δ-function restriction on the timelike variables.[9]

The vacuum to vacuum transition amplitude is recovered by taking the large T limit; in this case the initial and final boundary conditions (discussed below (A.15)) may be dropped [2.2]. We recover the expression for Z of eq. (3.26). The expression for the expectation value of a product of operators can be obtained by repeating much of the discussion of this appendix. We would thus recover the other part of (3.26).

For other discussion we need the continuum version of (A.15). Equations (3.5) and (3.8) form the bridge between the lattice and continuum variables:

$$\langle A'_\mu(x)|\mathrm{e}^{-TH}|A_\mu(x)\rangle = \int [\mathrm{d}A^\alpha_\mu(x)]\,\delta[A^\alpha_0]\exp -\tfrac{1}{4}\left\{\int \mathrm{d}^4x\,F^\alpha_{\mu\nu}F^\alpha_{\mu\nu}\right\}. \tag{A.16}$$

As the gauge transformations in continuum gauge theories are no longer restricted to a compact group [2.2, A.2], we cannot trivially eliminate the δ function. To do so would introduce an infinite scale factor. We do not expect answers to depend on singling out the time direction; this is especially true for Euclidian theories. Choosing an arbitrary direction n, (A.16) takes on the form

$$\langle 0|\mathrm{e}^{-TH}|0\rangle = \int [\mathrm{d}A^\alpha_\mu]\,\delta(n\cdot A)\exp\left\{-\int \tfrac{1}{4}F^2\,\mathrm{d}^4x\right\}. \tag{A.17}$$

Standard terminology refers to the situation where n is timelike, as the timelike gauge, and when n is space-like, as the axial gauge.

In appendix D we will obtain the propagators and some of the vertices based on (A.17).

Appendix B. Perturbative β-function

It was stated several times that the quantization of a field theory necessitates the introduction of some sort of length scale. In principle this could be the lattice parameter a, or equivalently a momentum cut-off $\sim 1/a$. Physical quantities will depend on this parameter, and one can trade this cut-off in favor of such quantities. This is the essence of the renormalization schemes. For large couplings on the lattice it was possible to define a tension σ; keeping this quantity fixed provided us a dependence of the coupling constant on the lattice parameter a. For small couplings this tension does not exist at least perturbatively and we need some other quantity. The usual procedure is to define a renormalized coupling constant as the value of some vacuum expectation at specified external momenta. We shall follow a different approach, more in tune with the spirit of this article. We will calculate the energy of a static quark–antiquark singlet pair at a separation R [B.1]. This turns out to have the form

$$E(R) = -\frac{4}{3}\frac{1}{4\pi R}\,\bar{g}^2[g(a); R/a]. \tag{B.1}$$

[9] It should not be surprising that enforcing Gauss' law introduces timelike link variables. In the continuum limit this corresponds to letting A_0 vary. The equation of motion obtained by varying A_0 in the Lagrangian is just Gauss' law. Choosing the $A_0 = 0$ gauge forced us to impose this law as a supplementary constraint.

$g(a)$ is the coupling constant appropriate to a cut-off $\sim 1/a$. In this situation, we fix the energy at some given distance $R = 1/\Lambda$ and find how g has to depend on a to insure that $E(1/\Lambda)$ is independent of a. This implies

$$0 = a\frac{\mathrm{d}\tilde{g}}{\mathrm{d}a} = \frac{\partial\tilde{g}}{\partial a} - \frac{\partial\tilde{g}}{\partial g}\tilde{\beta}(g)$$
$$\tilde{\beta}(g) = -a\,\partial g/\partial a. \tag{B.2}$$

Therefore

$$\tilde{\beta}(g) = (\partial\tilde{g}/\partial g)^{-1}\,\partial\tilde{g}/\partial a. \tag{B.3}$$

$\tilde{\beta}(g)$ is the renormalization group function analogous to that of (4.2). We have temporarily placed a tilde over it to emphasize that as it is defined differently from the strong coupling lattice β, it does not have to be equal to it. We will show that in a power series expansion in g the first two terms are universal [4.2]. From now on we will drop the tilde over the β function.

The power series for $\tilde{g}$ turns out to have the form

$$\tilde{g}(R)^2 = g^2 + 2\beta_0 g^4 \ln(\gamma_1 R/a) + 2g^6(\beta_1 \ln(\gamma_2 R/a) + 2\beta_0^2(\ln \gamma_1 R/a)^2) + \mathrm{O}(g^8). \tag{B.4}$$

Using (B.3) we find

$$\beta(g) = -\beta_0 g^3 - \beta_1 g^5 + \mathrm{O}(g^7). \tag{B.5}$$

The constants γ_i depend on the details of the particular renormalization scheme. We will first show that β_0 and β_1 are universal and then evaluate β_0. The evaluation of β_1 is very lengthy and only the result will be presented.

To show that the first two terms do not depend on the definition of the coupling constant, we evaluate

$$\bar{\beta}(\bar{g}) = -a\,\mathrm{d}\bar{g}/\mathrm{d}a \tag{B.6}$$

for a coupling constant $\bar{g}$. $\bar{g}$ can be expanded as a power series in g

$$\bar{g} = g + cg^3 + \mathrm{O}(g^5) \tag{B.7}$$
$$g = \bar{g} - c\bar{g}^3 + \mathrm{O}(\bar{g}^5).$$

Thus

$$\begin{aligned}\bar{\beta}(\bar{g}) &= (1 + 3cg^2)\,\beta(g)\\ &= -\beta_0 g^3 - \beta_1 g^5 - 3c\beta_0 g^5 + \cdots \\ &= -\beta_0\bar{g}^3 - \beta_1\bar{g}^5 + \cdots,\end{aligned} \tag{B.8}$$

and the first two terms agree with (B.5).

In order to find β_0 and β_1 we return to (B.1) and evaluate $E(R)$ through ordinary perturbation theory [B.1]. We use the Hamiltonian formalism developed in section 2. Rewriting (2.5) and (2.6) we find

$$H = \tfrac{1}{2}\int \mathrm{d}^3x\, \{\boldsymbol{E}^\alpha \cdot \boldsymbol{E}^\alpha + \boldsymbol{B}^\alpha \cdot \boldsymbol{B}^\alpha\}, \tag{B.9}$$

and the states are annihilated by the operator

$$\nabla \cdot \boldsymbol{E}^\alpha + g f^{\alpha\beta\gamma} A_i^\beta E_i^\gamma + g\rho^\alpha. \tag{B.10}$$

ρ^α is the charge density due to the quarks $\rho = \rho_1 + \rho_2$, and we will seek the term in the energy proportional to $\rho_1\rho_2$. This is the term we can interpret as the interquark potential. In the first two orders of perturbation theory such a term comes from the longitudinal part of the electric field. Namely, with the help of (B.10), we separate the electric field into a transverse part and a longitudinal part:

$$\begin{aligned} &\boldsymbol{E}^\alpha = \boldsymbol{E}_\mathrm{T}^\alpha - \nabla\varphi^\alpha \\ &\nabla \cdot \boldsymbol{E}_\mathrm{T}^\alpha = 0. \end{aligned} \tag{B.11}$$

Using (B.10) we solve for φ^α (to order g^2)

$$\varphi^\alpha = g[\delta^{\alpha\gamma} + g f^{\alpha\beta\gamma}\nabla^{-2}A^\beta \cdot \nabla]\, \nabla^{-2}[f^{\gamma\delta\epsilon}A^\delta \cdot E_\mathrm{T}^\epsilon + \rho^\gamma]. \tag{B.12}$$

$\boldsymbol{E}_\mathrm{T}$ and $\boldsymbol{A}_\mathrm{T}$ are the dynamical variables. To order g^4, the interesting part of the Hamiltonian is

$$H \approx \tfrac{1}{2}\int \mathrm{d}^3x [\boldsymbol{E}_\mathrm{T}^\alpha \cdot \boldsymbol{E}_\mathrm{T}^\alpha + (\nabla \times \boldsymbol{A}^\alpha)\cdot(\nabla \times \boldsymbol{A}) - \varphi^\alpha\nabla^2\varphi^\alpha]. \tag{B.13}$$

The first order contribution to the energy is

$$\begin{aligned} E_1(\rho) &= -\frac{g^2}{2}\int \mathrm{d}^3x\, \{\rho\nabla^{-2}\rho + 3g^2 f^{\alpha\beta\gamma} f^{\delta\epsilon\gamma}\langle 0|\rho^\alpha\nabla^{-2}A^\beta \cdot \nabla\nabla^{-2}A^\delta \cdot \nabla\nabla^{-2}\rho^\epsilon|0\rangle\} \\ &= \int \frac{\mathrm{d}^3x\, \mathrm{d}^3y}{4\pi|x-y|}\rho^\alpha(x)\,\rho^\alpha(y)\left\{1 + \frac{3g^2}{4}\ln\frac{\gamma_1^2(x-y)^2}{a^2}\right\}. \end{aligned} \tag{B.14}$$

γ_1 is a finite constant. To the same g^4 order we have to evaluate a second order perturbation contribution to the energy. The intermediate state consists of two gauge particles with momenta $\boldsymbol{k}_i$ and polarizations λ_i, $i = 1, 2$,

$$E_2 = -g^4 \sum_{\lambda_1,\lambda_2}\int \mathrm{d}^3k_1\, \mathrm{d}^3k_2 \frac{\left|\langle 0|\int \mathrm{d}^3x\, f^{\alpha\beta\gamma}\rho^\alpha\nabla^{-2}\boldsymbol{A}^\beta \cdot \boldsymbol{E}_\mathrm{T}^\gamma|k_i, \lambda_i\rangle\right|^2}{|k_1| + |k_2|}. \tag{B.15}$$

In reality, to this order, higher powers of ρ appear. As discussed previously, it is the quadratic term that can be given a potential interpretation. Evaluating (B.15) we obtain

$$E_2(\rho) = -\frac{1}{2}\int \frac{\mathrm{d}^3x\, \mathrm{d}^3y}{4\pi|x-y|}\frac{g^4}{12\pi^2}\ln[\gamma_2^2(x-y)^2/a^2]\rho^\alpha(x)\,\rho^\alpha(y). \tag{B.16}$$

γ_2 is another finite constant. Combining E_1 and E_2, we can read of the interquark potential

$$E(R) = -\frac{4}{3}\frac{1}{4\pi R}\left\{g^2 + \frac{11g^4}{16\pi^2}\ln[\gamma_c^2 R^2/a^2]\right\}. \tag{B.17}$$

The subscript c in γ_c is there to emphasize that this calculation was carried out in a continuous formalism. Comparing with (B.4) we find [1.11]

$$\beta_0 = 11/(16\pi^2). \tag{B.18}$$

A calculation of β_1 is much more involved [B.2]. The result is

$$\beta_1 = 102/(16\pi^2)^2. \tag{B.19}$$

The last topic we wish to discuss is the dependence of the coupling constant on the finite part of the g^4 term. From (B.4) or (B.17) we can define a renormalized coupling constant at $R = 1/\Lambda$, fixed

$$g^2(\Lambda) = g^2 + \beta_0 g^4 \ln[(\gamma/\Lambda a)^2]. \tag{B.20}$$

The finite constant depends on the calculational scheme. Above we emphasized this point by placing a subscript c on γ. Had we performed a lattice calculation we would have obtained a $\gamma_L \neq \gamma_c$. The two calculations can be brought into agreement by choosing different Λ's such that

$$\Lambda_c/\Lambda_L = \gamma_c/\gamma_L. \tag{B.21}$$

Detailed calculations comparing these various renormalization schemes have been carried out for continuum renormalization prescriptions [B.3], and for the comparison between continuum and lattice cut-offs [B.4, B.5]. Comparison of the Euclidian four dimensional lattice cut-off to the momentum subtraction continuum scheme yields

$$\begin{aligned} \Lambda_{\text{mom}} &= 83.5\Lambda_{\text{E.L.}} \qquad [\text{B.4}] \\ \Lambda_{\text{mom}} &= 87.3\Lambda_{\text{E.L.}} \qquad [\text{B.5}]. \end{aligned} \tag{B.22}$$

The origin of the 5% discrepancy in these numbers is unknown. The ratio of the Euclidian lattice $\Lambda_{\text{E.L.}}$ to the three dimensional spacial lattice $\Lambda_{\text{S.L.}}$ is [B.6] [cf. Note added in proof]:

$$\Lambda_{\text{S.L.}} = 3.012\Lambda_{\text{E.L.}}. \tag{B.23}$$

A renormalization scheme used widely in phenomenological applications is the modified minimal subtraction scheme ($\overline{\text{MS}}$) [4.3]. The relation of $\Lambda_{\overline{\text{MS}}}$ to other ones is [B.3]

$$\Lambda_{\overline{\text{MS}}} = 0.37\Lambda_{\text{mom}}. \tag{B.24}$$

Appendix C. Instantons

Examining the QCD Hamiltonian, eq. (2.5), we see that minimum energy configurations correspond to vanishing of the magnetic term. This implies that the gauge potential A is a pure gauge. Namely

$$A = \frac{\mathrm{i}}{g} U^{+} \nabla U, \tag{C.1}$$

with U a time independent unitary matrix. These unitary matrices fall into discrete classes based on their behavior at spatial infinity. Two unitary matrices which differ from each other in a finite region of space are said to belong to the same class. Nontrivial behavior at large distances can be obtained from transformations of the form

$$U_\nu(\boldsymbol{r}) = \exp[-2\pi\, \mathrm{i} \nu r_a R_{a\alpha} \lambda_\alpha / 2\sqrt{r^2 + \rho^2}]. \tag{C.2}$$

$R_{a\alpha}$ is a group rotation with $a = 1, 2, 3$ and $\alpha = 1 \cdots 8$, ν is an integer, and ρ an arbitrary length. A pure gauge potential

$$A^{(\nu)} = \frac{\mathrm{i}}{g} U_\nu^{+} \nabla U_\nu \tag{C.3}$$

can be characterized by a topological quantum number equal to ν

$$\nu = \frac{-\mathrm{i} g^3}{24\pi^2} \operatorname{Tr} \int \mathrm{d}^3 x\, A^{(\nu)} (A^{(\nu)} \times A^{(\nu)}). \tag{C.4}$$

Any pure gauge potential generated by a unitary matrix which differs from U_ν in a finite region of space will have the same topological quantum number as $A^{(\nu)}$.

The next question we ask is whether there is a solution of the classical equations of motion, which for Euclidian $t \to -\infty$ belongs to class $\nu = 0$, and at $t \to +\infty$ changes to $\nu = \pm 1$. In between this solution will not be a pure gauge. We first give the form of this solution in the Lorentz gauge and then transform to the $A_0 = 0$ gauge. This potential configuration is the BPTS [6.1] instanton and anti-instanton

$$A_\mu^\alpha(r) = \frac{2}{g} R_{a\alpha} \eta_{a\mu\nu} r_\nu / (r^2 + \rho^2). \tag{C.5}$$

r is now a Euclidian four vector, and

$$\begin{aligned} &\eta_{aij} = \epsilon_{aij} \\ &\eta_{a0i} = -\eta_{ai0} = \pm \delta_{ai} \end{aligned} \tag{C.6}$$

with the $\pm$ referring to instantons or anti-instantons.

This potential can be transformed to the $A_0 = 0$ gauge by

$$S(\boldsymbol{r}, r_0) = \exp\left\{-2\mathrm{i} \int_{-\infty}^{r_0} \frac{\mathrm{d}y_0\, r_a}{(r_0 - y_0)^2 + r^2 + \rho^2} R_{a\alpha} \frac{\lambda_\alpha}{2}\right\}. \tag{C.7}$$

For $r_0 \to -\infty$, $S \to 1$ and as $r_0 \to +\infty$, S approaches $U_{(1)}(\boldsymbol{r})$ of (C.2). As promised, the instanton interpolates between $\nu = 0$ and $\nu = 1$. Similarly we can show that the anti-instanton goes from $\nu = 0$ to $\nu = -1$.

The classical Euclidian action for (C.5) is

$$S_0(g) = \tfrac{1}{4}\int F^{\alpha}_{\mu\nu}\,\mathrm{d}^4x = 8\pi^2/g^2. \tag{C.8}$$

Appendix D. Properties of propagators and vertices in the axial gauge

As usual, the propagator of a theory is obtained from the quadratic part of the exponent appearing in (A.17)

$$L_Q = \tfrac{1}{4}(\partial_\mu A^{\alpha}_{\nu} - \partial_\nu A^{\alpha}_{\mu})(\partial_\mu A^{\alpha}_{\nu} - \partial_\nu A^{\alpha}_{\mu}). \tag{D.1}$$

Because $\delta(n \cdot A)$ appears in (A.17) not all potentials are integrated over. The propagator is the inverse of L_Q, and due to the above, we restrict the inverse to the three dimensional space orthogonal to n. The Euclidian propagator is [D.1, 6.9]

$$\Delta^{(E)}_{\mu\nu}(q) = \frac{1}{q^2}\left[\delta_{\mu\nu} - \frac{n_\mu q_\nu + n_\nu q_\mu}{n \cdot q} + \frac{n^2 q_\mu q_\nu}{(n \cdot q)^2}\right]. \tag{D.2}$$

The corresponding free Minkowski space propagator is

$$\Delta^{(0)}_{\mu\nu}(q) = \frac{-\mathrm{i}}{q^2 + \mathrm{i}\epsilon}\left[g_{\mu\nu} - \frac{n_\mu q_\nu + n_\nu q_\mu}{n \cdot q} + \frac{n^2 q_\mu q_\nu}{(n \cdot q)^2}\right]. \tag{D.3}$$

Of course, $n_\mu \Delta_{\mu\nu} = 0$ and so the four-dimensional inverse of $\Delta_{\mu\nu}$ does not exist. In the space orthogonal to n, the inverse of $\Delta^{0}_{\mu\nu}$ is

$$\begin{aligned} \Pi_{\mu\nu} &= \left[\Pi^{(0)}_{\mu\nu} + \mathrm{i}\,\frac{(q_\mu n_\nu + n_\mu q_\nu) n \cdot q}{n^2} - \mathrm{i}\,\frac{n_\mu n_\nu}{n^2}\left(q^2 + \frac{(n \cdot q)^2}{-n^2}\right)\right] \\ \Pi^{(0)}_{\mu\nu} &= \mathrm{i}(q^2 g_{\mu\nu} - q_\mu q_\nu). \end{aligned} \tag{D.4}$$

In many instances $\Pi_{\mu\nu}$ is contracted with tensors orthogonal to n and only $\Pi^{(0)}_{\mu\nu}$ will be important. We will loosely refer to it as the "inverse" propagator

$$\begin{aligned} \Pi^{(0)}_{\mu\alpha}\Delta^{(0)}_{\alpha\nu} &= g_{\mu\nu} - n_\mu q_\nu / n \cdot q \\ \Delta^{(0)}_{\mu\alpha}\Pi^{(0)}_{\alpha\nu} &= g_{\mu\nu} - n_\nu q_\mu / n \cdot q. \end{aligned} \tag{D.5}$$

These relations will be generalized to the full propagator and its "inverse".

We shall also need some of the exact relations between the full propagators and the full vertices. The propagator is defined as

$$\delta^{\alpha\beta}\Delta_{\mu\nu}(q) = \frac{-\mathrm{i}}{(2\pi)^4}\int \mathrm{d}x\, \mathrm{e}^{\mathrm{i}q \cdot x}\langle 0|(A^{\alpha}_{\mu}(x)\, A^{\beta}_{\nu}(0))|0\rangle, \tag{D.6}$$

and the proper three point vertex as

$$\Delta_{\nu\alpha}(q)\,\Delta_{\lambda\beta}(r)\,\Gamma^{abc}_{\mu\alpha\beta}(p,q,r)\,\delta^4(p+q+r)$$
$$=\frac{\mathrm{i}}{(2\pi)^8}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\mathrm{d}^4z\;\mathrm{e}^{ipx}\,\mathrm{e}^{iqy}\,\mathrm{e}^{iyz}\,\langle 0|T(J^a_\mu(x)\,A^b_\nu(y)\,A^c_\lambda(z))|0\rangle. \tag{D.7}$$

J^a_μ is the color current.

The conservation law, $\partial_\mu J_\mu$, and the current potential commutation relations

$$\delta(x_0)[J^a_0(x), A^b_\nu(0)] = \mathrm{i}f^{abb'}\delta^4(x)\,A^{b'}_\nu(x), \tag{D.8}$$

provide the Ward identity [2.2, D.2]

$$\mathrm{i}p^\mu\Gamma^{abc}_{\mu\nu\lambda}(p,q,r) = f^{abc}[\Pi_{\nu\lambda}(r)-\Pi_{\nu\lambda}(q)]. \tag{D.9}$$

Analogous results hold for four point vertices. Γ is symmetric under the interchange of all the variables.

Acknowledgement

This work would have been impossible without the help of many colleagues. For discussions and/or correspondence involving preliminary results I am grateful to: J. Ambjørn, J. Ball, T.W. Chiu, M. Creutz, S. Drell, D. Gross, J. Kogut, G. Mack, R. Pearson, J. Richardson, K. Wilson and F. Zachariasen.

For cheerfully typing, retyping and proofreading, I would like to thank Mrs. J. Brown, Mrs. M. Golian and Mrs. E. Williams.

To Carol, I extend my thanks for not losing her good humor during the past several months.

Note added in proof

A. Hasenfratz and P. Hasenfratz [Budapest, KFK5-1981-15] have reevaluated the relation between Λ_{mom} and $\Lambda_{\mathrm{E.L.}}$, eq. (B.22), by the methods of ref. [B.5] and confirmed the result of ref. [B.4]. Thus $\Lambda_{\mathrm{mom}} = 83.5\Lambda_{\mathrm{E.L.}}$. As the difference between the various results was only 5% this does not affect any of the previous discussion.

Hasenfratz and Hasenfratz have also evaluated

$$\Lambda_{\mathrm{S.L.}} = 0.91_{\mathrm{E.L.}}$$

which is considerably different from the result presented in eq. (B.23). This introduces serious discrepancies between the Hamiltonian and Euclidian lattice perturbation theories as well as Monte Carlo calculations. The value of $\Lambda_{\mathrm{S.L.}}$ presented in eq. (5.5) yields

$$\Lambda_{\overline{\mathrm{MS}}} = (220 \pm 50)\ \mathrm{MeV}$$

in total disagreement with results of Euclidian perturbation theory, eq. (5.8) and the Monte Carlo results, eq. (5.14).

A possible resolution of these discrepancies may lie in the extrapolation procedures used to determine A (eq. (5.6)) from Hamiltonian perturbation theory. A sixth order result yielded $A = 170$. If this is the correct value rather than the extrapolated one then agreement with other methods would be restored.

References

[1.1] M. Gell-Mann, Phys. Letters 8 (1964) 214.
[1.2] G. Zweig, CERN Reports 8182/TH.401 and 8419/TH.412 (1964).
[1.3] J. Bjorken and S. Glashow, Phys. Letters 11 (1964) 84.
[1.4] J.J.J. Kokkedee, The Quark Model (W.A. Benjamin Inc., New York, 1969).
[1.5] L. Lyons, Oxford University Nuclear Physics Laboratory report 38/80.
[1.6] D. Cutts et al., Phys. Rev. Letters 41 (1978) 363.
[1.7] G. La Rue, W.M. Fairbank and A.F. Hebard, Phys. Rev. Letters 38 (1977) 1011;
G. La Rue, W.M. Fairbank and J.D. Phillips, Phys. Rev. Letters 42 (1979) 142.
[1.8] O.W. Greenberg, Phys. Rev. Letters 13 (1964) 598;
H. Fritzsch and M. Gell-Mann, in: Proc. Conf. on Duality and Symmetry in High Energy Physics, ed. E. Gotsman (Weizmann, Jerusalem, 1971);
W.A. Bardeen, H. Fritzsch and M. Gell-Mann, in: Scale and Conformal Symmetry in Hadron Physics, ed. R. Gatto (Wiley, New York, 1972).
[1.9] H. Fritzsch and M. Gell-Mann, in: Proc. XVI Intern. Conf. on High Energy Physics, Vol. 2 (National Accelerator Laboratory, 1972) p. 135;
H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Letters 74B (1973) 365.
[1.10] S. Weinberg, Phys. Rev. Letters 31 (1973) 494.
[1.11] D. Gross and F. Wilczek, Phys. Rev. Letters 30 (1973) 1343;
H.D. Politzer, Phys. Rev. Letter 30 (1973) 1346.
[1.12] J.D. Bjorken, Phys. Rev. 179 (1969) 1547.
[1.13] R.P. Feynman, in: High Energy Collisions, eds. C. Yang et al. (Gordon and Breach, New York, 1969).
[1.14] A.J. Buras, Rev. Mod. Phys. 52 (1980) 199.
[1.15] W. Marciano and H. Pagels, Physics Reports 36C (1978) 137.
[2.1] M. Gell-Mann and Y. Ne'eman, The Eightfold Way (W.A. Benjamin, Inc., New York, 1964).
[2.2] E.S. Abers and B.W. Lee, Physics Reports 9C (1973) 1.
[2.3] R.P. Feynman and A.R. Hibbs, Quantum Mechanics and Path Integrals (McGraw-Hill, New York, 1965);
F.A. Berezin, The Method of Second Quantization (Academic, New York, 1966).
[2.4] E. Eichten and K. Gottfried, Phys. Letters 66B (1977) 286;
E. Eichten, K. Gottfried, T. Kinoshita, K.D. Lane and T.M. Yan, Phys. Rev. D21 (1980) 203.
[2.5] K. Wilson, Phys. Rev. D10 (1974) 2455.
[2.6] J. Kogut and L. Susskind, Phys. Rev. D9 (1974) 3501;
G. 't Hooft, CERN Report No. TH/1902 (1974).
[3.1] J. Kogut and L. Susskind, Phys. Rev. D11 (1975) 395;
J. Kogut, Rev. Mod. Phys. 51 (1979) 659.
[3.2] R. Balian, J. Drouffe and C. Itzykson, Phys. Rev. D10 (1975) 3376; D11 (1975) 2098; D11 (1975) 2104.
[3.3] L. Pontrjagin, Topological Groups (Princeton University Press, 1959);
P. Carruthers, Introduction to Unitary Symmetry (Interscience, New York, 1966).
[3.4] J. Kogut, R. Pearson and J. Shigemitsu, Phys. Rev. Letters 43 (1979) 484.
[3.5] J. Kogut, R. Pearson and J. Shigemitsu, University of Illinois preprint, ILL-(TH)-80-41.
[3.6] G. Münster and P. Weisz, DESY 80/57.
[3.7] R. Pearson, preprint NSF-ITP-80-49.
[3.8] M. Lüscher, G. Münster and P. Weisz, DESY 80/63.
[3.9] C. Itzykson, M.E. Peskin and J.B. Zuber, Phys. Letters 95B (1980) 259.
[3.10] A. Hasenfratz, E. Hasenfratz and P. Hasenfratz, TH.2890-CERN (1980).
[4.1] E. Stueckelberg and A. Peterman, Helv. Phys. Acta 26 (1953) 499;
M. Gell-Mann and F.E. Low, Phys. Rev. 95 (1954) 1300.
[4.2] C.G. Callan, Phys. Rev. D2 (1970) 1541;
K. Symanzik, Comm. Math. Phys. 18 (1970) 227.
[4.3] W. Bardeen, A. Buras, D. Duke and T. Muta, Phys. Rev. D18 (1978) 3998.
[4.4] J. Ellis, M.K. Gaillard, D.V. Nanopolous and S. Rudaz, Annecy report, LAPP-TH-14 (1980).
[4.5] L.F. Abbott and R.M. Barnett, Annals. of Phys. 125 (1980) 276.

[4.6] K. Wilson, Rev. Mod. Phys. 47 (1975) 773;
T. Niemeijer and J. Van Leeuwen, in: Phase Transitions and Critical Phenomena, Vol. 6, eds. C. Domb and M. Green (Academic Press, New York, 1976) p. 425.
[4.7] K. Wilson, Cornell report-CLNS/80/442.
[5.1] J. Kogut and J. Shigemitsu, Phys. Rev. Letters 45 (1980) 410.
[5.2] M. Creutz, Phys. Rev. Letters 43 (1976) 553; 43 (1979) 890 (E).
[5.3] M. Creutz, Brookhaven report BNL-26847 (1979).
[5.4] M. Creutz, Phys. Rev. Letters 45 (1980) 313.
[6.1] C. Callan, R. Dashen and D. Gross, Phys. Rev. D17 (1978) 2717; D19 (1979) 1826.
[6.2] C. Callan, R. Dashen and D. Gross, Phys. Rev. D20 (1979) 3279.
[6.3] G. 't Hooft, Phys. Rev. D14 (1976) 3432;
A. Polyakov, Nucl. Phys. B120 (1977) 429.
[6.4] D. Gross, private communication.
[6.5] B. Berg and M. Lüscher, Nucl. Phys. B160 (1979) 281.
[6.6] R. Anishetty, M. Baker, S.K. Kim, J.S. Ball and F. Zachariasen, Phys. Letters 86B (1979) 52;
M. Baker, J.S. Ball, P. Lucht and F. Zachariasen, Phys. Letters 89B (1980) 211;
J.S. Ball and F. Zachariasen, Phys. Letters 95B (1980) 273.
[6.7] S. Mandelstam, Phys. Rev. D20 (1979) 3223.
[6.8] N.N. Bogoliubov and D.V. Shirkov, Introduction to the Theory of Quantized Fields (Interscience, London, 1959).
[6.9] J.S. Ball and T.W. Chiu, Univ. of Utah report, UU/HEP-80/3.
[6.10] S.K. Kim and M. Baker, Nucl. Phys. B164 (1980) 152.
[6.11] J.S. Ball, private communications.
[7.1] H. Kramers and G. Wannier, Phys. Rev. 60 (1941) 252;
R. Savit, Rev. Mod. Phys. 52 (1980) 453.
[7.2] S. Mandelstam, Phys. Rev. D19 (1979) 2391.
[7.3] G. 't Hooft, Nucl. Physics B138 (1978), Nucl. Physics B153 (1979) 141.
[7.4] G. 't Hooft, Lecture II at the Cargèse Summer Institute (1979).
[7.5] J. Ambjorn, B. Felsager and P. Olesen, Niels Bohr Institute preprint NBI-HE-80-22.
[7.6] G. Mach and V.B. Petkova, Ann. Phys. (N.Y.) 123 (1979) 442.
[7.7] G. Münster, Physics Letters 95B (1980) 59.
[7.8] G. Mack and E. Pietarinen, Phys. Letters 94B (1980) 397.
[7.9] J. Ambjorn and H. Flyvbjerg, Niels Bohr Institute preprint NBI-HE-80-30.
[7.10] S. Mandelstam, Phys. Lett. 53B (1975) 476; Phys. Reports 23C (1976) 245.
[7.11] T.T. Wu and C.N. Yang, in: Properties of Matter Under Unusual Conditions, eds. H. Mark and S. Fernbach (Interscience, New York, 1968).
[7.12] D. Foerster, Phys. Letters 76B (1978) 597.
[8.1] R.F. Streater and A.S. Wightman, PCT, Spin & Statistics, and all that (W.A. Benjamin, New York, Amsterdam, 1964).
[8.2] C. Rebbi, Phys. Reports 12C (1974) 1;
J. Scherk, Rev. Mod. Phys. 47 (1975) 123.
[8.3] Y. Nambu, Phys. Letters 80B (1979) 372.
[8.4] J.L. Gervais and A. Neveu, Phys. Lett. 80B (1979) 255; Ecole Normale Superieure preprint LPTENS-80/6.
[8.5] A.M. Polyakov, Phys. Letters 82B (1979) 247.
[8.6] A.M. Polyakov, Phys. Letters 82B (1979) 247; Los Alamos preprint.
[8.7] L. Durand and E. Mendel, Phys. Letters 85B (1979) 241.
[8.8] S. Mandelstam, Phys. Rev. 175 (1968) 1580.
[8.9] Yu.M. Makeenko and A.A. Migdal, Phys. Letters 88B (1979) 135; 89B (1980) 437(E); ITEP preprint-23-1980;
A.A. Migdal, Landau Institute preprint (1979); Landau Institute preprint-80-0647.
[8.10] T. Eguchi, Phys. Rev. Letters 44 (1980) 126.
[8.11] M. Lüscher, Phys. Letters 90B (1980) 277.
[8.12] M. Lüscher, K. Symanzik and P. Weisz, DESY preprint 80/63.
[8.13] G. Domokos and S. Kovesi-Domokos, Johns Hopkins preprint JHU-HET 7904 (1979).
[8.14] M. Shifman, A. Vainstein and V. Zakharov, Nucl. Phys. B147 (1979) 385, 448.
[8.15] D. Weingarten, Phys. Letters 90B (1980) 281.
[9.1] S. Weinberg, Gravitation and Cosmology (John Wiley and Sons, Inc., New York, 1972).
[9.2] A.M. Polyakov, Phys. Letters 72B (1978) 477.
[9.3] L. Susskind, Phys. Rev. D20 (1979) 2610.
[9.4] D.J. Gross, R.D. Pisarski and L.G. Yaffe, preprint.
[9.5] L.D. McLerran and B. Svetisky, SLAC-PUB-2572 (1980);
J. Kuti, J. Polónyi and K. Szlachányi, Budapest KFKI 1980-67.
[9.6] M. Stone and P.R. Thomas, Phys. Rev. Letters 41 (1978) 351.

[10.1] S.D. Drell, M. Weinstein and S. Yankielovicz, Phys. Rev. D14 (1976) 1677.
[10.2] E. Fradkin and L. Susskind, Phys. Rev. D17 (1978) 2637;
T. Banks and E. Rabinovici, Nucl. Phys. B160 (1979) 349.
[10.3] A. De Rujula, R.C. Giles and R.L. Jaffe, Phys. Rev. D17 (1978) 285;
H. Georgi, Phys. Rev. D22 (1980) 225;
A. De Rujula, R.C. Giles and R.L. Jaffe, Phys. Rev. D22 (1980) 227.
[A.1] J. Villain, J. Phys. 36 (1975) 581;
J.V. Jose, L.P. Kadanoff, S. Kirkpatrick and D.R. Nelson, Phys. Rev. B (1978) 1217.
[A.2] L.D. Fadde'ev and V.N. Popov, Phys. Letters 25B (1967) 29.
[B.1] V.N. Gribov, Lecture at the 12th Winter School of the Leningrad Nuclear Physics Institute (1977);
T. Appelquist, M. Dine and J.J. Muzinich, Phys. Letters 69B (1977) 231; Phys. Rev. D8 (1978) 2074;
S.D. Drell, private communication and Oxford University report.
[B.2] W. Caswell, Phys. Rev. Lett. 33 (1974) 244;
D.R.T. Jones, Nucl. Phys. B75 (1974) 531.
[B.3] W. Celmaster and R. Gonsalves, Phys. Rev. D20 (1979) 1420.
[B.4] A. Hasenfratz and P. Hasenfratz, Phys. Letters 93B (1980) 165.
[B.5] R. Dashen and D. Gross, preprint (1980).
[B.6] D. Gross, private communication.
[C.1] A.A. Belavin, A.M. Polyakov, A.S. Schwartz and Yu.S. Tyupkin, Phys. Letters 59B (1975) 85.
[D.1] W. Kummer, Acta Phys. Austriaca 41 (1975) 315.
[D.2] J.C. Taylor, Nucl. Phys. B33 (1971) 436;
A. Slavnov, Theor. and Math. Phys. 10 (1972) 153 (English translation 10 (1972) 99).