VARIATIONAL TECHNIQUES IN NON-PERTURBATIVE QCD

ALEX KOVNER

School of Mathematics and Statistics, University of Plymouth
Drake Circus, PL4 8AA, United Kingdom
and
Physics Department, University of Connecticut, 2152 Hillside Road,
Storrs, CT 06269-3046, USA

J. GUILHERME MILHANO

CENTRA, Instituto Superior Técnico (IST),
Av. Rovisco Pais, P-1049-001 Lisboa, Portugal
and
Department of Physics, CERN,
CH-1211 Genève 23, Switzerland

We review attempts to apply the variational principle to understand the vacuum of non-abelian gauge theories. In particular, we focus on the method explored by Ian Kogan and collaborators, which imposes exact gauge invariance on the trial Gaussian wave functional prior to the minimization of energy. We describe the application of the method to a toy model, confining compact QED in 2+1 dimensions, where it works wonderfully and reproduces all known non-trivial results. We then follow its applications to pure Yang-Mills theory in 3+1 dimensions at zero and finite temperature. Among the results of the variational calculation are dynamical mass generation and the analytic description of the deconfinement phase transition.

Table of Contents

This review is dedicated to the memory of Ian Kogan. Ian was an extraordinary physicist with an almost unimaginable breadth of interests. The non-perturbative domain of QCD was but one of the many problems that interested him. As in all areas that he worked in, Ian left a lasting mark in this subject. During the last 10 years of his life, Ian periodically returned to the variational approach described in this review, always with new ideas of what to do next and how to do it better. Ian's enthusiasm and bubbling energy was the driving force that caused the initial embryonic idea to develop into a solid calculational approach which has already produced many interesting results, and will hopefully keep on developing and improving. Ian's untimely death is a severe blow to all of his many friends. We miss him...

1. Introduction

Quantum Chromodynamics (QCD) was formulated more than 30 years ago. There is very little doubt that it is, indeed, the correct theory of strong interactions. An impressive amount of experimental data is successfully described by QCD calculations in the high momentum transfer domain. For these processes it is the short time - short distance dynamics that is relevant. Fortunately, in this domain QCD is weakly coupled and thus precise quantitative information can be obtained by analytic calculations.

Of course, high momentum transfer processes are not the only, and perhaps not the most interesting, domain governed by QCD dynamics. The vacuum (low energy, large distance) sector of the strong interactions exhibits a wealth of phenomena which are both qualitatively striking at the experimental level and notoriously difficult to establish within the fundamental framework of QCD. Understanding low energy phenomena in QCD, such as confinement and chiral symmetry breaking or, in more general terms, the strong coupling problem and the ground state structure of an asymptotically free non-abelian gauge theory is, without doubt, one of the main problems in modern quantum field theory. In spite of years of attempts to answer these questions we are still far from achieving this goal.

Many routes have been tried in approaching this problem. They range from the highly computer intensive numerical programme of the lattice gauge theory [1], through the universality based concepts of effective field theories [2] to the monopole and vortex inspired searches for the confinement mechanism [3,4]. In recent years, there have also been attempts to approach the non-perturbative physics of QCD utilizing information about supersymmetric gauge theories [5].

While all these ideas are interesting and productive, each has its own drawbacks. The lattice gauge theory aims, in principle, at producing the complete set of numbers that characterize the QCD spectrum, condensates, various matrix elements, etc. This goal is, however, still not within reach. Also often one would like to understand the underlying physics rather than just calculate a given number albeit with an accuracy of a fraction of a percent, and this is difficult within a numerical approach. The effective field theory and dual superconductor approaches are based on simple physical pictures, but have their starting points rather far from actual QCD, so that making quantitative contact with QCD dynamics is very difficult. The SUSY motivated route suffers from the same basic problem, as it is not in principle clear how much the QCD dynamics is distorted by supersymmetry.

One should therefore welcome any method which attempts to analytically obtain dynamical information directly from QCD, even if this information may be partial and incomplete. Such a method should be, of course, intrinsically non-perturbative. Unfortunately, the arsenal of non-perturbative methods to tackle strongly interacting quantum field theories (QFT) is, to say the least, very limited. Methods that perform very well in simple quantum mechanical problems are much more difficult to use in QFT. This is true, for example, of the variational approach. In quantum mechanics it is usually enough to know a few simple qualitative features in order to set up a variational Ansatz which gives pretty accurate results, not only for the energy of a ground state, but also for various other vacuum expectation values (VEV). In QFT one is immediately faced with several difficult problems when trying to apply this method, as discussed insightfully by Feynman [6].

Nevertheless over the years there have been several attempts to apply different versions of the variational Rayleigh-Ritz method to QCD. The purpose of this article is to review the results of this approach. We will concentrate most of our attention on the recent incarnation of the variational method pioneered by Ian Kogan with collaborators [14], which in terms of applications and results made much more headway than any of the previous attempts. This is not to say that it is free from problems and immune to criticism. Still, we believe that this is a good time to summarize its status for two reasons. First because some interesting quantitative results have already been achieved; and, second, because there is plenty of room for improving the method, and some aspects can be improved with relative ease, so that clearly the method has not yet outlived its usefulness.

As mentioned earlier, when applying a variational method to QFT one is faced with many difficulties. First of all there is a problem of generality of a trial state. The trial state ought to be general enough to allow, through variation of its parameters, the relevant physics to be spanned. In quantum mechanics the task can be put down simply to identifying a few critical physical properties and consequently writing a compliant trial state. On the other hand, the Hilbert space of QFT is enormous, and it is much more difficult to identify "by pure thought" the relevant characteristics that have to be probed.

Then there is the problem of calculability. That is, even if one had a very good guess at the form of the vacuum wave functional (or, for that matter, even knew its exact form) one would still have to evaluate expectation values of various operators in this state,

$$\langle\mathcal{O}\rangle = \int D\phi\Psi^*[\phi]\mathcal{O}\Psi[\phi]\,. \tag{1.1}$$

A calculation of this kind is, obviously, tantamount to the evaluation of a Euclidean functional integral with the square of the wave functional (WF) playing the role of the partition function. One should therefore be able to solve exactly a d-dimensional field theory with the action

$$S[\phi] = -\log\Psi^*[\phi]\Psi[\phi]\,. \tag{1.2}$$

In quantum mechanics such a concern plays only a background role. Whatever the chosen trial state might be, the calculation to be performed involves integrals of functions. The evaluation of any such integral can be tackled, if not always analytically then numerically, without major complications. In QFT, where our ability to evaluate path integrals is, to say the least, limited, the calculability restriction on the trial wave functional is very severe. Since in dimension $d > 1$ the only theories one can solve exactly are free field theories, the requirement of calculability almost unavoidably restricts the possible form of the trial WF to a Gaussian state

$$\Psi[\phi] = \exp\left\{-\frac{1}{2}\int d^3x d^3y\,[\phi(x)-\zeta(x)]\,G^{-1}(x,y)\,[\phi(y)-\zeta(y)]\right\} \tag{1.3}$$

with $\zeta(x)$ and $G(x,y)$ being c-number functions.

Another serious problem is that of "ultraviolet modes". The main motivation of a variational calculation in a strongly interacting theory is to learn about the distribution of the low momentum modes of the field in the vacuum wave functional. However, the VEV of the energy (and all other intensive quantities) is dominated entirely by contributions of high momentum fluctuations, for the simple reason that there are infinitely more ultraviolet modes than modes with low momentum. Therefore, even if one has a very good idea of how the WF at low momenta should look, if the ultraviolet part of the trial state is even slightly incorrect the minimization of energy may lead to absurd results. Due to the interaction between the high and low momentum modes, there is a good chance that the infrared (IR) variational parameters will be driven to values which minimize the interaction energy, and have nothing to do with the dynamics of the low momentum modes themselves.

Finally, in gauge theories there is an additional complication: that of gauge invariance. The allowed wave functions must be invariant under the time independent gauge transformations. If one does not impose the Gauss' law on the states exactly, one is not solving the right problem. The QCD Hamiltonian is only defined on the gauge invariant states, and its action on non gauge invariant states can be modified at will. Thus, by minimizing a particularly chosen Hamiltonian without properly restricting the set of allowed states, one is taking the risk of finding a "vacuum" which has nothing to do with the physical one, but is only picked due to a specific form of the action of the Hamiltonian outside the physical subspace. There are two ways to approach this problem. One is to solve Gauss' law operatorially *à la* Dirac [7,8]. This leads to a "completely gauge fixed" formalism. Once this is done any state can be chosen. The price to be paid is that the Hamiltonian in any completely fixed gauge is very complicated. The calculation of the expectation value of energy then is not analytically manageable. Another way is to let the Gauss' law constraint be, but write down trial states which are explicitly gauge invariant. The Hamiltonian in such a calculation is simple, but the trial states are usually very complicated. Thus the problem of gauge invariance is linked very strongly with the problem of calculability.

2. The variational setup

So given that the only path integral that we know how to calculate analytically is that of a Gaussian wave function, what kind of progress can one make with variational calculations in QFT? Surprisingly enough in QFT without gauge symmetry a Gaussian ansatz can take one a long way. The famous BCS theory of superconductivity is nothing but a variational analysis of the interacting QFT using a Gaussian variational state [9]. The Gaussian Ansatz has also been applied to self-interacting relativistic scalar and spinor theories [10], where it gives non-trivial exact results in the large N limit. The reason it works is quite simple. A Gaussian wave functional is the exact ground state in quantum field theories of non-interacting fields, massless or massive. As a trial state it is thus flexible enough to probe the existence of a mass gap in the spectrum. Therefore, whenever the main effect of the interaction is to dynamically generate (or significantly change) the mass of the particles, the Gaussian Ansatz is adequate and informative. Put another way, one can hope that the Gaussian WF is useful if the non-perturbative physics is dominated by a single condensate, that of the lowest possible dimension. From this point of view, it would seem that, it is perfectly reasonable to try a similar variational Ansatz in the Yang-Mills theory. After all, it is strongly suggested by the QCD sum rules [15] that the pure glue sector is strongly dominated by the lowest dimensional non-perturbative condensate $\langle F^2 \rangle$.

The difficulty comes from the necessity to impose gauge invariance. It is very easy to see that in a non-abelian theory it is impossible to write a Gaussian WF that satisfies the constraint of gauge invariance. The SU(N) gauge theory is described by a Hamiltonian

$$\mathsf{H} = \int d^3x \left[\frac{1}{2} E_i^{a2} + \frac{1}{2} B_i^{a2} \right] \tag{2.1}$$

where

$$\begin{aligned} E_i^a(x) &= i \, \frac{\delta}{\delta A_i^a(x)} \,, \\ B_i^a(x) &= \frac{1}{2} \, \epsilon_{ijk} \{ \partial_j A_k^a(x) - \partial_k A_j^a(x) + g f^{abc} A_j^b(x) A_k^c(x) \} \,, \end{aligned} \tag{2.2}$$

and all physical states must satisfy the constraint of gauge invariance

$$G^a(x)\Psi[A] = \Big[\partial_i E_i^a(x) - g f^{abc} A_i^b(x) E_i^c(x) \Big] \, \Psi[A] = 0 \,. \tag{2.3}$$

Under a gauge transformation U – generated by $G^a(x)$ – the vector potential transforms as

$$A_i^a(x) \rightarrow A_i^{U\,a}(x) = S^{ab}(x) A_i^b(x) + \lambda_i^a(x) \,, \tag{2.4}$$

where

$$S^{ab}(x) = \frac{1}{2} \, \mathrm{tr} \left(\tau^a U^\dagger \tau^b U \right) \,, \quad \lambda_i^a(x) = \frac{i}{g} \, \mathrm{tr} \left(\tau^a U^\dagger \partial_i U \right) \,, \tag{2.5}$$

and τ^a are traceless Hermitian N by N matrices satisfying $\mathrm{tr}(\tau^a \tau^b) = 2\delta^{ab}$. A Gaussian wave functional

$$\begin{aligned} \Psi[A_i^a] = \exp\Big\{ - \frac{1}{2} \int d^3x \, d^3y \, [A_i^a(x) - \zeta_i^a(x)] & \\ \times (G^{-1})_{ij}^{ab}(x,y) \Big[A_j^b(y) - \zeta_j^b(y) \Big] \Big\} & \end{aligned} \tag{2.6}$$

transforms under the gauge transformation as

$$\Psi[A_i^a] \to \Psi[(A^U)_i^a]\,. \tag{2.7}$$

In the abelian case it is enough to take $\partial_i G_{ij}^{-1} = 0$ to satisfy the constraint of gauge invariance. In the non-abelian case, however, due to the homogeneous piece in the gauge transformation Eq. (2.4), no gauge invariant Gaussian WF exists. Thus one has to abandon the notion of a simple Gaussian variational Ansatz on the Hilbert space spanned by the canonical variables $\{A_i^a(x), E_i^a(x)\}$.

One option is to attempt to solve the constraint Eq. (2.3) operatorially. Such a procedure leads to a "completely gauge fixed" formalism. For example if one solves Eq. (2.3) for E_3^a and then performs Dirac quantization, one obtains the Hamiltonian in the axial gauge. Another popular choice is to express $\partial_i E_i^a$ in terms of the transverse components of electric field and vector potential. This leads to the Hamiltonian in the Coulomb gauge. In either case, after the Hilbert space has been reduced to the physical space, the Gaussian trial state in the remaining degrees of freedom can be considered. However, the actual calculations are extremely difficult since the Hamiltonian in any completely fixed gauge is very complicated. Attempts to perform such a calculation in the axial gauge were made in [16, 17]. The problem encountered here is that the calculation is plagued by a spurious infrared pole of the form $1/k_3$. The results depend very much on the way one treats this pole, and there is no clear understanding as to the proper way to do so. The Coulomb gauge calculation is even less straightforward, primarily because the Coulomb gauge is known to suffer from Gribov ambiguity [18]. Thus the expression for the Hamiltonian that one obtains in this gauge is only formal, and has to be very carefully defined by considering zeros of the Fadeev-Popov operator. Non-perturbative analytic calculations in this gauge are therefore all but impossible [19]. Attempts have been made to perform numerical variational minimization taking account of the Gribov horizon [20, 21], but the interpretation of the results is very difficult.

Another route, pursued in [22], is to modify the Gaussian trial state in a "minimal" way to make it compatible with gauge invariance. Thus one considers the state

$$\Psi[A] = \exp\left\{-\int d^3x\, d^3y\, B_i^a(x) W^{ab}(x,y) B_j^b(y) G_{ij}(x-y)\right\} \tag{2.8}$$

where W^{ab} is the Wilson path ordered integral along the straight line connecting points x and y taken in the adjoint representation. Explicit introduction of the Wilson line indeed makes the state gauge invariant. Some semi-quantitative arguments were given in [22] to the effect that the best variational state of this type will have a short range variational function $G(x-y)$ and therefore will be confining at large distances. However the state is so complicated that no reasonable way to get the actual calculation going was ever found.

Finally, another series of works attempts to initially disregard the Gauss' law constraint, and subsequently calculate corrections due to its implementation perturbatively [23–27]. A clear discussion of this method and the analogy with nuclear physics calculations is given in a recent review paper [28]. Although the method is mathematically very elegant, it is not suited for systems whose energy may be lowered by a non-perturbative amount due to the exact implementation of Gauss' law. As we believe this to be the case for the Yang-Mills theory, we strongly doubt the usefulness of this method to obtain realistic results.[a]

In this review we will therefore restrict ourselves to the discussion of the approach proposed in [14]. The scheme devised in [14] is rather straightforward. One opts for strictly preserving gauge invariance, and constructs a gauge invariant wave functional by projecting the Gaussian wave functional of Eq. (2.6) onto the gauge invariant sector. Restricting ourselves to the case of zero classical fields ($\zeta = 0$), the variational Ansatz proposed in [14] is

$$\Psi[A_i^a] = \int DU(x) \exp\left\{-\frac{1}{2}\int d^3x\, d^3y\, A_i^{Ua}(x) G_{ij}^{-1ab}(x,y)\; A_j^{Ub}(y)\right\} \quad (2.9)$$

with A_i^{Ua} defined in Eq. (2.4) and the integration is performed over the space of special unitary matrices with the $SU(N)$ group invariant measure.

The trial state defined in this way is explicitly gauge invariant. It is not obvious at this stage that we will be able to deal with it analytically. But as we will see in the following, some headway can be made using analytical approximation schemes to calculate various expectation values in this state. In this review we will discuss the applications of this technique to the pure Yang-Mills theory at zero as well as at finite temperature.

But before plunging head first into the discussion of non-abelian Yang-Mills theories, we will consider the much simpler, but nevertheless informative, case of a compact U(1) gauge theory in 2+1 dimensions [30]. The theory is known to be confining, and our aim is to see whether the confining properties of the ground state can be captured with our simple Ansatz. It

---

[a] There has also been some work on the application of the variational principle to determine excited states, rather than the ground state [29]. In this review we restrict ourselves to the problem of the vacuum.

will also illustrate how the use of different field theoretical techniques allows one to carry out the variational estimation of the ground state.

3. Compact QED in 2+1 dimensions

We start by setting up the Hamiltonian description of U(1) compact QED. First of all, we need to determine what the Hilbert space of admissible states is. It is clear that Gauss' law should be implemented, and thus all the physical states should satisfy

$$\exp\left\{i\int d^2x\,\partial_i\phi(x)E_i(x)\right\}|\Psi\rangle = |\Psi\rangle \tag{3.1}$$

for an arbitrary continuous function $\phi(x)$. This is true for both compact and non-compact QED. There is however a crucial difference between Gauss' law in the compact theory and in the non-compact one. In the non-compact theory Eq. (3.1) should be satisfied only for regular functions ϕ. For example, the operator

$$V(x) = \exp\left\{\frac{i}{g}\int d^2y\,\epsilon_{ij}\,\frac{(x-y)_j}{(x-y)^2}\,E_i(y)\right\} \tag{3.2}$$

which has the form of Eq. (3.1) with the function ϕ proportional to the planar angle θ, i.e. $\phi = \frac{1}{g}\theta(x)$, does not act trivially on the physical states. In fact, this operator creates a point-like magnetic vortex with magnetic flux $2\pi/g$ and therefore changes the physical state on which it acts.

In the compact theory the situation in this respect is quite different. Point-like vortices with quantized magnetic flux $2\pi n/g$ cannot be detected by any measurement. This translates into the requirement that the creation operator of a point-like vortex must be indistinguishable from the unit operator. In other words, the operator Eq. (3.2) generates a transformation which belongs to the compact gauge group, and should therefore act trivially on all physical states. Eq. (3.1) should therefore be satisfied also for these operators.

Accordingly, the Hamiltonian of the compact theory must be also invariant under these transformations. The magnetic field itself, defined as $B = \epsilon_{ij}\partial_i A_j$, does not commute with $V(x)$ (cf. Ref. [11]),

$$V^\dagger(x)B(y)V(x) = B(y) + \frac{2\pi}{g}\,\delta^2(x-y)\,. \tag{3.3}$$

The Hamiltonian should therefore contain not B^2 but rather a periodic function of B. We will choose our Hamiltonian to be

$$H = \frac{1}{2}\, a^2 \sum E^2_{\mathbf{n}i} - \frac{1}{g^2 a^2} \sum \cos g a^2 B_{\mathbf{n}}\,. \tag{3.4}$$

Since we will need in the following an explicit ultraviolet regulator, we switched temporarily to lattice notation. Here a is the lattice spacing, and the sums are respectively over the links and plaquettes of the two-dimensional spatial lattice. The coefficients of the two terms in the Hamiltonian are chosen so that in the weak coupling limit, upon formal expansion to lowest order in g^2, the Hamiltonian reduces to the standard free Hamiltonian of 2+1 dimensional electrodynamics. Following Polyakov [32], we work in the weakly coupled regime. Since the coupling constant g^2 in 2+1 dimensions has dimension of mass, weak coupling means that the following dimensionless ratio is small

$$g^2 a \ll 1\,. \tag{3.5}$$

Our aim now is to find variationally the vacuum wave functional of this theory.

3.1. *The variational Ansatz*

As our variational trial ground state we choose the Gaussian wave function of A projected onto the gauge singlet. The projection has to be performed with respect to the full compact gauge group of Eq. (3.1).

To facilitate this, we define a vortex field $A^V(x)$ that satisfies (we suppress the lattice spacing a henceforth)

$$\nabla \times A^V(x) = \frac{2\pi}{g}\, \delta^2(x)\,, \qquad \nabla \cdot A^V = 0\,. \tag{3.6}$$

This is the vector potential corresponding to a magnetic field that is zero everywhere except at $x = 0$, where it takes the value $\frac{2\pi}{g}$. The explicit solution of Eq. (3.6) is

$$A^V_i(x) = -\frac{1}{g}\, \epsilon_{ij}\, \frac{x_j}{x^2}\,. \tag{3.7}$$

The compact gauge invariance requires the variational wave function $\psi[A]$ to be invariant under shifts $A \to A + A^V$. This is, of course, consistent with the periodicity of H under $B \to B + \frac{2\pi}{g}$.

Hence we define a field, shifted by a non-compact gauge transformation $\phi_{\mathbf{n}}$ and by an integer valued vortex distribution $m(x)$,

$$A^{(\phi,m)}(x) = A(x) - \nabla\phi(x) - \sum_y m(y)A^V(x-y) \tag{3.8}$$

or, for short,

$$A^{(\phi,m)} = A - \nabla\phi - A^V \cdot m \ . \tag{3.9}$$

We choose the gauge invariant and periodic trial wave function as

$$\psi[A] = \sum_{\{m(x)\}} \int [d\phi] \, \exp\left[-\frac{1}{2}\int d^2x d^2y A_x^{(\phi,m)} G^{-1}(x-y) A_y^{(\phi,m)}\right] \ . \tag{3.10}$$

Under a gauge transformation,

$$\psi[A+\nabla\lambda] = \psi[A] \tag{3.11}$$

since λ can be absorbed into a shift in ϕ. The simple rotational structure of $G_{ij} = \delta_{ij}G$ that appears in the variational wave function Eq. (3.10) is consistent with perturbation theory. We also take $G(x)$ to be a real function.

Our task now is to calculate the expectation value of the Hamiltonian in this state, and to minimize it with respect to the variational function G. We start by considering the normalization of the wave function.

3.2. *The normalization integral*

The norm of $|\psi\rangle$ is

$$\begin{aligned} Z \equiv \langle\psi|\psi\rangle = &\sum_{\{m,m'\}} \int [d\phi][d\phi'][d\mathbf{A}] \\ &\times \exp\left[-\frac{1}{2}A^{(\phi,m)}G^{-1}A^{(\phi,m)}\right] \exp\left[-\frac{1}{2}A^{(\phi',m')}G^{-1}A^{(\phi',m')}\right] \ . \end{aligned} \tag{3.12}$$

We shift A by $\nabla\phi' + A^V \cdot m'$ and absorb the shift into ϕ and m, obtaining

$$Z = \sum_{\{m\}} \int [d\phi][d\mathbf{A}] \, e^{-\frac{1}{2}A^{(\phi,m)}G^{-1}A^{(\phi,m)}} e^{-\frac{1}{2}AG^{-1}A} \, . \tag{3.13}$$

Now we combine the exponents according to

$$\begin{aligned} A^{(\phi,m)}G^{-1}A^{(\phi,m)} + AG^{-1}A = &\, 2A^{(\phi/2,m/2)}G^{-1}A^{(\phi/2,m/2)} \\ &+ \frac{1}{2}S(\phi,m)G^{-1}S(\phi,m) \, , \end{aligned} \tag{3.14}$$

where

$$S \equiv \nabla\phi + A^V \cdot m \ . \tag{3.15}$$

It is easy to see that $SG^{-1}S$ contains no cross terms between m and ϕ. We shift A by $\nabla\phi/2 + A^V \cdot m/2$, and all the fields decouple. We have then

$$Z = Z_A Z_\phi Z_v \,, \tag{3.16}$$

where

$$Z_A = \det \pi G \,, \tag{3.17}$$

$$Z_\phi = \int [d\phi]\, e^{-\frac{1}{4}\nabla\phi\cdot G^{-1}\cdot\nabla\phi} = \left(\det 4\pi \frac{1}{\nabla^2} G\right)^{1/2} \,, \tag{3.18}$$

$$Z_v = \int_{\{m_{\mathbf{n}'}\}} \exp\left[-\frac{1}{4g^2}\int d^2x\, d^2y\, m(x) D(x-y) m(y)\right] . \tag{3.19}$$

Here, Z_v is the vortex "partition function", with the "vortex-vortex interaction" D given by

$$D(x'-y') = g^2 \int d^2x d^2y\, A^V(x-x') \cdot G^{-1}(x-y) A^V(y-y') \,. \tag{3.20}$$

We can split off the $x = y$ terms in Eq. (3.19) and write

$$Z_v = \int [dm(x)] \exp\left[-\frac{1}{4g^2}\int_{x\neq y} m(x) D(x-y) m(y)\right] \prod_y z^{m(y)^2} \,, \tag{3.21}$$

where we have defined the vortex fugacity

$$z = e^{-\frac{1}{4g^2}D(0)} \,. \tag{3.22}$$

We expect the ultraviolet behavior of the variational function G to be the same as in the free theory, *viz.* (for the Fourier transform),

$$G^{-1}(k) \sim k \,, \tag{3.23}$$

so that

$$D(0) \sim \int^\Lambda \frac{d^2k}{(2\pi)^2}\, \frac{4\pi^2}{k^2}\, G^{-1}(k) \sim 2\pi\Lambda \tag{3.24}$$

and thus

$$z \sim e^{-\frac{\pi}{2}\frac{\Lambda}{g^2}} \,, \tag{3.25}$$

where we use the momentum space cutoff $\Lambda = a^{-1}$.

In the weak coupling region we have $z \ll 1$, which allows us to restrict possible values of m to $0, \pm 1$ in Eqs. (3.19, 3.21).

3.3. *Expectation values*

We are now ready to calculate the expectation value of the Hamiltonian Eq. (3.4). Using the definition Eq. (3.10), we obtain

$$V^{-1} \left\langle \int E^2 \, d^2x \right\rangle = -\frac{1}{V} \left\langle \psi \left| \sum_{\mathbf{n},i} \frac{\partial^2}{\partial A^2_{\mathbf{n},i}} \right| \psi \right\rangle \\ = \frac{1}{2} \int \frac{d^2k}{(2\pi)^2} G^{-1}(k) - \frac{\pi^2}{g^2} \int \frac{d^2k}{(2\pi)^2} k^{-2} G^{-2}(k) K(k) \,, \tag{3.26}$$

where $K(k)$ is the correlation function of the vorticity

$$K(k) \equiv \int d^2x \, e^{ikx} \left\langle m(x) m(0) \right\rangle \,. \tag{3.27}$$

To calculate correlation functions of m we use a duality transformation [31, 33–35]. We add an $iJ \cdot m$ term to the exponent in Eq. (3.19) and use the formula

$$e^{-\frac{1}{4g^2} m \cdot D \cdot m} = const \int [d\chi] \, e^{-g^2 \chi \cdot D^{-1} \cdot \chi} e^{i\chi \cdot m} \tag{3.28}$$

to obtain

$$Z_v[J] = \int [d\chi] \, e^{-g^2 \chi \cdot D^{-1} \cdot \chi} \prod_x [1 + 2\cos(\chi(x) + J(x))] \,. \tag{3.29}$$

Noting that [b]

$$\cos(\chi + J) = \langle \cos\chi \rangle_0 :\cos(\chi + J): = z :\cos(\chi + J): \,, \tag{3.30}$$

we have

$$Z_v = \int [d\chi] \, e^{-g^2 \chi \cdot D^{-1} \cdot \chi} \prod [1 + 2z :\cos(\chi + J):] \\ \simeq \int D\chi \exp\left[-g^2 \chi D^{-1} \chi + 2z \int d^2x \; :\cos\Big(\chi(x) + J(x)\Big): \right] \,. \tag{3.31}$$

[b] The normal ordering is performed relative to the free theory defined by the quadratic action in Eq. (3.29).

Correspondingly [30],

$$\langle m(x)m(y)\rangle = 2g^2 D^{-1}(x-y) - 4g^4 \left\langle D^{-1}\chi(x)\, D^{-1}\chi(y)\right\rangle \tag{3.32}$$

and

$$K(k) = 2z + O(z^2)\,, \tag{3.33}$$

which, in this approximation, does not depend on momentum.

The propagator of χ is also easily calculated. To first order in z,

$$\begin{aligned}\int d^2x\, e^{ikx}\,\langle\chi(x)\chi(0)\rangle &= \frac{1}{2g^2D^{-1}(k)+2z}\\ &= \frac{D(k)}{2g^2} - z\,\frac{D^2(k)}{2g^4} + O(z^2)\,.\end{aligned} \tag{3.34}$$

Thus we have

$$V^{-1}\left\langle \int E^2\, d^2x\right\rangle = \frac{1}{2}\int \frac{d^2k}{(2\pi)^2}\, G^{-1}(k) - \frac{2\pi^2}{g^2} z \int \frac{d^2k}{(2\pi)^2}\, k^{-2}G^{-2}(k)\,. \tag{3.35}$$

The magnetic part is easily calculated since it has an exponential form and, therefore, with our trial wave function leads to a simple Gaussian integral. We find

$$\left\langle e^{ingB_{\mathbf{n}}}\right\rangle = \exp\left[-\frac{1}{4}n^2g^2\int\frac{d^2k}{(2\pi)^2}\,k^2G(k)\right]\left\langle e^{in\pi m_{\mathbf{n}}}\right\rangle\,. \tag{3.36}$$

The second factor is different from unity only for odd values of n. Using Eq. (3.31) we find $\left\langle e^{i\pi m}\right\rangle = e^{-4z}$. Expanding to leading order in g^2 and z, we get

$$\left\langle -\frac{1}{g^2}\cos gB\right\rangle = \frac{1}{4}\int\frac{d^2k}{(2\pi)^2}\,k^2G(k) + \frac{4}{g^2}\,z\,, \tag{3.37}$$

where we have dropped an additive constant. Finally, the expression for the variational vacuum expectation value of the energy to first order in z is

$$\frac{1}{V}\langle H\rangle = \frac{1}{4}\int\frac{d^2k}{(2\pi)^2}\left[G^{-1}(k) + k^2G(k) - \frac{4\pi^2}{g^2}\,z\left(k^{-2}G^{-2}(k) - \frac{4}{\pi^2}\right)\right]\,. \tag{3.38}$$

3.4. *Determination of the ground state*

The expression Eq. (3.38) has to be minimized functionally with respect to $G(k)$. From Eqs. (3.22, 3.24) we find

$$\frac{\delta z}{\delta G(k)} = \frac{1}{4g^2}\, k^{-2} G^{-2}(k)\, z\,. \tag{3.39}$$

The variation of Eq. (3.38) gives [c]

$$k^2 - G^{-2}(k) = \frac{4\pi^4}{g^4}\, z k^{-2} G^{-2}(k) \int \frac{d^2p}{(2\pi)^2} \left[p^{-2} G^{-2}(p) - \frac{4}{\pi^2} \right]. \tag{3.40}$$

Equation (3.40) has the solution

$$G^{-2}(k) = \frac{k^4}{k^2 + m^2}\,, \tag{3.41}$$

where

$$m^2 = \frac{4\pi^4}{g^4} z \int \frac{d^2k}{(2\pi)^2} \left[k^{-2} G^{-2}(k) - \frac{4}{\pi^2} \right]. \tag{3.42}$$

The main contribution to the integral in the gap equation, Eq. (3.42), comes from momenta $k^2 \gg m^2$. For these momenta $k^2 G^{-2}(k) = 1$. Therefore, we see that Eq. (3.42) has a non-trivial solution. Using Eqs. (3.22, 3.24, 3.41) we obtain

$$m^2 = \frac{4\pi^4}{g^4} \exp\left(-\frac{\pi^2}{g^2} \int \frac{d^2p}{(2\pi)^2} \frac{1}{\sqrt{p^2+m^2}} \right) \int \frac{d^2k}{(2\pi)^2} \left[\frac{k^2}{k^2+m^2} - \frac{4}{\pi^2} \right], \tag{3.43}$$

which for $g^2 a = g^2/\Lambda \ll 1$ can be simplified to (cf. eq (3.25))

$$m^2 = 4\pi^2\, \frac{(\pi^2 - 4)\Lambda^4}{g^4}\, \exp\left(-\frac{\pi\Lambda}{2g^2} \right), \tag{3.44}$$

where we have restored the ultraviolet cutoff dependence explicitly. The resulting m is the mass gap of the theory, in the sense that it is the inverse of the spatial correlation length. Calculating, for example, the propagator of magnetic field, we find

$$\left\langle e^{igB_{\mathbf{m}}} e^{-igB_{\mathbf{n}}} \right\rangle = \left| \left\langle e^{igB} \right\rangle \right|^2 e^{\frac{g^2}{2} \nabla^2 G(\mathbf{m}-\mathbf{n})}\,, \tag{3.45}$$

[c] We have dropped a term $-\frac{8\pi^2}{g^2} z k^{-2} G^{-3}(k)$ from the right-hand side of Eq. (3.40) since it is smaller by a factor of $\frac{g^2 k}{\Lambda^2}$ than the term retained when one assumes $G \sim k^{-1}$ at large k.

and at large distances (neglecting power-like prefactors),

$$\nabla^2 G(x) = -\int \frac{d^2k}{(2\pi)^2}(k^2+m^2)^{1/2} e^{i\mathbf{k}\cdot\mathbf{x}} \sim e^{-mx}\,. \tag{3.46}$$

This dynamically generated mass is Polyakov's result [36, 37]. Thus, we recover in the Hamiltonian approach the first important result known about compact QED: a finite mass gap m, as well as its correct dependence on the coupling constant.

3.5. *Spatial Wilson loops*

We also want to see whether the charges are confined in our best variational state. The simplest quantity that is related to confinement is the expectation value of the Wilson loop. Therefore, we will calculate it in our ground state

$$W_l[C] = \left\langle \exp\left(ilg \oint_C \mathbf{A}\cdot d\mathbf{x}\right)\right\rangle = \left\langle \exp\left(ilg \int_\Sigma B\, dS\right)\right\rangle\,, \tag{3.47}$$

where l is an integer and the integral is over the area Σ bounded by the loop C. We have

$$W_l[C] = \left\langle \prod_S e^{il\pi m_{\mathbf{n}}}\right\rangle Z_A^{-1}\int D\mathbf{A}\exp\left(-AG^{-1}A + ilg\int_\Sigma B\, dS\right)\,. \tag{3.48}$$

The second factor is a Gaussian integral, which gives

$$W_A = \exp\left(\frac{l^2g^2}{4}\int_\Sigma d^2x\int_\Sigma d^2y\, \nabla^2 G(\mathbf{x}-\mathbf{y})\right)\,. \tag{3.49}$$

In the limit of large Σ the leading behavior of the exponent is

$$-\frac{l^2}{4}\, g^2\Sigma \lim_{k\to 0} k^2 G(k) = -\frac{l^2}{4}\, g^2 m\Sigma\,. \tag{3.50}$$

This gives the area law with the string tension

$$\sigma = \frac{l^2}{4}\, g^2 m\,. \tag{3.51}$$

The first factor in Eq. (3.48) is different from unity only for odd l. It can be easily calculated but gives only subleading corrections to the string tension [30].

3.6. *Potential between external charges*

We have thus found that in the best variational state the Wilson loop has an area law behavior. This usually signals confinement, and one is naturally inclined to conclude that we have indeed found confinement with the string tension related in the expected way to the dynamically generated scale, $\sigma \propto g^2 m$. One must be however a tad more careful at this stage, since the spatial Wilson loop does not directly give the potential between external charges. Although in the Euclidean formulation there is no difference between spatial and time-like Wilson loops, in the Hamiltonian approach this is not obvious. It is, therefore, desirable to calculate directly the potential between external charges.

How does one do this? Obviously one has to introduce into the theory the source corresponding to the pair of external charges. The result is a modification of Gauss' law,

$$\partial_i E_i(x) = g\rho(x)\,, \tag{3.52}$$

with $\rho(x) = \delta(x - x_1) - \delta(x - x_2)$. As the external charges do not have dynamics of their own, the Hamiltonian remains unchanged. The seemingly simplest option seems to be to take the same Gaussian variational state which minimizes the energy in the vacuum sector, and project it with the modified projection operator corresponding to the new Gauss' law. This "minimally modified" state would be

$$\psi[A] = \sum_{\{m(x)\}} \int [d\phi] \exp\Big[-\frac{1}{2} \int d^2x d^2y A_x^{(\phi,m)} G^{-1}(x-y) A_y^{(\phi,m)} + ig\rho(x)\phi(x)\Big]\,. \tag{3.53}$$

One could then calculate the energy expectation value in this state and take this as an estimate of the interaction potential. This procedure was suggested in [38–40]. However it turns out that the estimate obtained for the interaction energy in this way is very unreliable. For example, in the compact QED_3 case this calculation was performed in [41] with the resulting energy being not even infrared finite. Instead, to get a reasonable estimate for the energy one has to introduce additional variational parameters. In [41] the variational Ansatz was extended to

$$\psi[A] = \sum_{\{m(x)\}} \int [d\phi] \exp\left[-\frac{1}{2}\int d^2x d^2y A_x^{(\phi,m)} G^{-1}(x-y) A_y^{(\phi,m)} \right. \\ \left. + ie_i(x) A_i^{(\phi,m)} + ig\rho(x)\phi(x) \right] . \quad (3.54)$$

Here, the "classical" electric field profile $e_i(x)$ is to be varied so as to minimize the energy. It turns out that with this extra variational parameter the calculation gives very satisfactory results [41]. Without giving the details of this calculation, we note that it confirms the expectation that the potential between external charges is a linear function of the separation. The string tension calculated this way coincides within ten percent with Eq. (3.51).

What have we learned from this toy model? First of all it is very encouraging that the Gaussian projected Ansatz is good enough to reproduce all known results and, in particular, confinement and dynamical mass generation with parameterically correct values of string tension and mass.

We have also seen that we need some ingenuity to be able to carry out the calculations with the variational wave function. We were able to do it in this simple case as the coupling constant was small and we could utilize existing methods for treating weakly interacting two dimensional systems. It is quite clear that in 3+1 dimensions no such techniques will be available and consequently the situation will be, in this respect, much more complicated. Finally, we learned that we could not impose the Gauss' law constraint on the state perturbatively. Perturbatively, the operator V, Eq. (3.2), simply vanishes. However, without including it into the projection procedure of the Gaussian Ansatz, we would not get any non-trivial results; we would have found the vacuum of a free massless photon without a dynamical mass gap and with vanishing string tension. We fully expect that this aspect will remain important in the application to QCD, to which we turn in the next section.

4. The Yang-Mills theory

The dynamics of the pure glue sector of QCD is described by the Yang-Mills (also often called gluodynamics) Hamiltonian

$$\mathsf{H} = \int d^3x \left[\frac{1}{2} E_i^{a2} + \frac{1}{2} B_i^{a2} \right] , \quad (4.1)$$

where

$$E_i^a(x) = i\,\frac{\delta}{\delta A_i^a(x)}\,,$$
$$B_i^a(x) = \frac{1}{2}\,\epsilon_{ijk}\left\{\partial_j A_k^a(x) - \partial_k A_j^a(x) + g f^{abc} A_j^b(x) A_k^c(x)\right\}, \tag{4.2}$$

and all physical states must satisfy the constraint of gauge invariance

$$G^a(x)\Psi[A] = \left[\partial_i E_i^a(x) - g f^{abc} A_i^b(x) E_i^c(x)\right]\Psi[A] = 0\,. \tag{4.3}$$

We thus have to choose a set of variational states which are invariant under the action of Eq. (4.3). We start then with a Gaussian state

$$\Psi_0[A_i^a] = \exp\left\{-\frac{1}{2}\int d^3x\, d^3y\, A_i^a(x)(G^{-1})_{ij}^{ab}(x,y)A_j^b(y)\right\}, \tag{4.4}$$

where the set of functions $G_{ab}^{ij}(x)$ are variational parameters.

We hope that the freedom allowed by variation of G^{-1} is sufficiently wide to probe the non-perturbative physics of the Yang-Mills vacuum. As discussed in the introduction, the states of the form Eq. (4.4) are not gauge invariant and, therefore, as such do not belong to the physical Hilbert space of gluodynamics. To remedy this problem we project the states onto the gauge invariant subspace by gauge transforming them and integrating over the whole gauge group [d],

$$\Psi[A_i^a] = \int DU(x)\exp\left\{-\frac{1}{2}\int_{x,y} A_i^{Ua}(x)\, G_{ij}^{-1ab}(x,y)\, A_j^{Ub}(y)\right\} \tag{4.5}$$

with A_i^{Ua} defined as

$$A_i^{Ua}(x) = S^{ab}(x)A_i^b(x) + \lambda_i^a(x)\,, \tag{4.6}$$

and

$$S^{ab}(x) = \frac{1}{2}\,\mathrm{tr}\left(\tau^a U^\dagger \tau^b U\right), \quad \lambda_i^a(x) = \frac{i}{g}\,\mathrm{tr}\left(\tau^a U^\dagger \partial_i U\right) \tag{4.7}$$

with the τs, which are traceless $N\times N$ hermitian matrices, normalized by

$$\mathrm{tr}\,(\tau^a\tau^b) = 2\delta^{ab}\,. \tag{4.8}$$

The integration in Eq. (4.5) is performed over the space of special unitary matrices with the $SU(N)$ group invariant measure. This integration

[d] Hereafter we use the notational shorthand $\int_x \equiv \int d^3x$.

projects the original Gaussian state onto a color singlet. Due to the projection operation, the calculation of expectation values in this state is much more involved than in the case of a simple Gaussian. A full functional minimization with respect to the variational functions $G^{-1ab}_{ij}(x,y)$ is, therefore, beyond our calculational abilities.

In order to render calculations possible, we will impose several restrictions on the form of G^{-1}, which will lead to considerable simplifications. First, we require the state to be translationally invariant, that is we assume that Lorentz symmetry is not spontaneously broken in the ground state, restricting the form of G^{-1} to

$$G^{-1}(x,y) = G^{-1}(x-y)\,. \tag{4.9}$$

Further, we will only consider matrices G of the form

$$G^{ab}_{ij}(x-y) = \delta^{ab}\delta_{ij}G(x-y)\,. \tag{4.10}$$

This form is certainly the right one in the perturbative regime. In the leading order in perturbation theory, the non-abelian character of the gauge group is not important, and the integration in Eq. (4.5) is basically over the $U(1)^{N^2-1}$ group. The δ^{ab} structure is then obvious; there is a complete democracy between different components of the vector potential. The δ_{ij} structure arises in the following way. If it were not for the integration over the group, G^{-1}_{ij} would be precisely the (equal time) propagator of the electric field. However, due to the integration over the group, the actual propagator is the transverse part of G^{-1}. It is easy to check that the longitudinal part $\partial_i G^{-1}_{ij}$ drops out of all physical quantities. At the perturbative level, therefore, one can take $G_{ij} \sim \delta_{ij}$ without any loss of generality. Outside the perturbative framework Eq. (4.10) is a genuine restriction on the Ansatz, and we will adopt this form of the matrix G in order to simplify our variational calculation.

We can use additional perturbative information to further restrict the form of G. The theory of interest is asymptotically free. This means that the short distance asymptotics of correlation functions must be the same as in perturbation theory. Since G^{-1} in perturbation theory is directly related to correlation functions of gauge invariant quantities (e.g. E^2), we must have

$$G^{-1}(x) \to \frac{1}{x^4}\,, \quad x \to 0\,. \tag{4.11}$$

Finally, we expect the theory non-perturbatively to have a gap. In other words, the correlation functions should decay to zero at some distance scale

$$G(x) \sim 0\,, \quad x > \frac{1}{M}\,. \tag{4.12}$$

We will build this into our variational Ansatz in a fairly naive way. We will take M to be our only variational parameter. This can be done by choosing for $G(x)$ a particular form that has the UV and IR asymptotics given by Eqs. (4.11) and (4.12) such as, for example, a massive scalar propagator with mass M. We find another parameterization slightly more convenient. The form that will be used throughout this calculation has the Fourier transform

$$G^{-1}(k) = \begin{cases} \sqrt{k^2} & \text{if } k^2 > M^2, \\ M & \text{if } k^2 < M^2. \end{cases} \tag{4.13}$$

Using a massive propagator instead changes the results very little. Equation (4.5), together with Eqs. (4.10, 4.13), defines our variational Ansatz.

We now have to calculate the energy expectation value in these states and minimize it with respect to the only variational parameter left – the scale M. Note that the perturbative vacuum is included in this set of states and corresponds to $M = 0$. A non-zero result for M would therefore mean a non-perturbative dynamical scale generation in the Yang-Mills vacuum. Now we have to face up to the question of how to calculate averages in the state Eq. (4.5).

4.1. *The effective σ-model: the running coupling and the low momentum limit*

The expectation value of an arbitrary gauge invariant operator $\mathcal{O}$ is given by the functional integral

$$\begin{aligned} \langle\mathcal{O}\rangle &= \frac{1}{Z}\int DUDU'\langle\mathcal{O}\rangle_A\,, \\ \langle\mathcal{O}\rangle_A &= \int DA\, e^{-\frac{1}{2}\int_{x,y} A_i^{Ua}(x)G^{-1}(x-y)A_i^{Ua}(y)} \\ &\qquad\qquad \times \mathcal{O}\, e^{-\frac{1}{2}\int_{x',y'} A_j^{U'b}(x')G^{-1}(x'-y')A_j^{U'b}(y')}\,, \end{aligned} \tag{4.14}$$

where Z is the norm of the trial state. Two simplifications are immediately obvious. First, for gauge invariant operators $\mathcal{O}$, one of the group integrations is redundant. Performing the change of variables $A \to A^U$ (and remembering that both integration measures DU and DA are group invariant), we obtain (omitting the volume of $SU(N)$ factor $\int dU$)

$$\begin{aligned}
\langle\mathcal{O}\rangle &= \frac{1}{Z}\int DU\langle\mathcal{O}\rangle_A\,,\\
\langle\mathcal{O}\rangle_A &= \int DA\, e^{-\frac{1}{2}\int_{x,y} A_i^{Ua}(x)G^{-1}(x-y)A_i^{Ua}(y)}\\
&\qquad\qquad\qquad\qquad \times\mathcal{O}\, e^{-\frac{1}{2}\int_{x',y'} A_j^b(x')G^{-1}(x'-y')A_j^b(y')}\,. \qquad (4.15)
\end{aligned}$$

Also, since the gauge transform, Eq. (4.6), of a vector potential is a linear function of A, for fixed $U(x)$ this is a Gaussian integral, and can therefore be performed explicitly for any reasonable operator $\mathcal{O}$. We are then left only with a path integral over one group variable $U(x)$. But this is not easy.

Let us consider first the normalization factor Z. After integrating over the vector potential we obtain

$$Z = \int DU \exp\{-\Gamma[U]\} \qquad (4.16)$$

with

$$\Gamma[U] = \frac{1}{2}\,\mathrm{Tr}\ln\mathcal{M} + \frac{1}{2}\,\lambda\,[G + SGS^T]^{-1}\lambda\,, \qquad (4.17)$$

where multiplication is understood as matrix multiplication with the indices: color a, space i and position (the values of space coordinates) x, i.e.

$$\begin{aligned}
(AB)_{ik}^{ac}(x,z) &= \int_y A_{ij}^{ab}(x,y)B_{jk}^{bc}(y,z)\,,\\
\lambda O\lambda &= \int_{x,y}\lambda_i^a(x)O_{ab}^{ij}(x-y)\lambda_j^b(y)\,. \qquad (4.18)
\end{aligned}$$

The trace Tr is understood as a trace over all three types of indices. In Eq. (4.17) we have defined

$$\begin{aligned}
S_{ij}^{ab}(x,y) &= S^{ab}(x)\delta_{ij}\delta(x-y)\,,\\
\mathcal{M}_{ij}^{ab}(x,y) &= \left[S^{Tac}(x)S^{cb}(y) + \delta^{ab}\right]G^{-1}(x-y)\delta_{ij}\,, \qquad (4.19)
\end{aligned}$$

where $S^{ab}(x) = \frac{1}{2}\mathrm{tr}\left(\tau^a U^\dagger \tau^b U\right)$ and $\lambda_i^a(x) = \frac{i}{g}\mathrm{tr}\left(\tau^a U^\dagger\partial_i U\right)$ were defined in Eq. (4.7) and tr is a trace over color indices only. Using the completeness condition for $SU(N)$,

$$\tau_{ij}^a\tau_{kl}^a = 2\left(\delta_{il}\delta_{jk} - \frac{1}{N}\delta_{ij}\delta_{kl}\right)\,, \qquad (4.20)$$

one can see that S^{ab} is an orthogonal matrix

$$S^{ab}S^{cb} = \frac{1}{4}\,\tau^b_{ij}\tau^b_{kl}\left(U\tau^a U^\dagger\right)_{ji}\left(U\tau^c U^\dagger\right)_{kl} = \frac{1}{2}\,\mathrm{tr}\left(\tau^a\tau^b\right) = \delta^{ab}\,. \tag{4.21}$$

We have written Eq. (4.17) in a form which suggests a convenient way of thinking about the problem. The functional integral Eq. (4.16) defines a partition function of a non-linear sigma model with the target space $SU(N)/Z_N$ in three dimensional Euclidean space. The fact that the target space is $SU(N)/Z_N$ rather than $SU(N)$ follows from the observation that the action Eq. (4.17) is invariant under local transformations belonging to the center of $SU(N)$. This can be trivially traced back to invariance of A^a_i under gauge transformations that belong to the center of the gauge group.

Note that the quantity $U(x)$ has a well defined gauge invariant meaning, and it is *not* itself a matrix of a gauge transformation. A contribution of a given $U(x)$ to the partition function Eq. (4.16) and to other expectation values corresponds to the contribution to the same quantity from the off-diagonal matrix element between the initial Gaussian and the Gaussian gauge rotated by $U(x)$. Consequently, if matrices $U(x)$ which are far from unity give significant contribution to the partition function, it means that the off-diagonal contribution is large, and therefore that the simpleminded non gauge invariant Gaussian approximation (which neglects the off diagonal elements) misses important physics.

The action of this sigma model is rather complicated. It is a non-local and a non-polynomial functional of $U(x)$. There are, however, two observations that will help us devise an approximation scheme to deal with the problem. First, remembering that the bare coupling constant of the Yang-Mills theory is small, let us see how it enters the sigma model action. It is easy to observe that the only place it enters is in the second term in the action Eq. (4.17), because $\lambda^a_i(x)$ has an explicit factor $1/g$. Moreover, it enters in the same way as a coupling constant in a standard sigma model action. We can, therefore, easily set up perturbation theory in our sigma model. With the standard parameterization

$$U(x) = \exp\left\{i\,\frac{g}{2}\,\phi^a\tau^a\right\} \tag{4.22}$$

one gets $\lambda^a_i(x) = -\partial_i\phi^a(x) + O(g),\ \ S^{ab}(x) = \delta^{ab} + O(g)$ and the leading order term in the action becomes

$$\frac{1}{16}\int_{x,y}\partial_i\phi^a(x)G^{-1}(x-y)\partial_i\phi^a(y)\,. \tag{4.23}$$

This is a free theory albeit with a non-standard propagator which at large momenta behaves as

$$D(k) \sim G(k)\,\frac{1}{k^2} \sim \frac{1}{|k|^3}\,. \tag{4.24}$$

It is easy to see that in this sigma model perturbation theory the coupling constant renormalizes logarithmically. The first order diagram that contributes to the coupling constant renormalization is the tadpole. In a sigma model with a standard kinetic term this diagram diverges linearly as $\int d^3k/k^2$, a sign of perturbative non-renormalizability. In our model, however, due to a non-standard form of the kinetic term Eq. (4.24), the diagram diverges only logarithmically as $\int d^3k/k^3$. The form of the β-function is, therefore, very similar to the β-function in ordinary QCD perturbation theory. It is a straightforward albeit tedious matter to calculate the one loop graphs in the σ-model perturbation theory and to extract the renormalization of the coupling constant [42]. The result is

$$\beta(g) = -\frac{g^3}{(4\pi)^2}\,4N\,. \tag{4.25}$$

For comparison, the pure Yang-Mills β-function is

$$\beta(g) = -\frac{g^3}{(4\pi)^2}\left(4 - \frac{1}{3}\right)N\,. \tag{4.26}$$

The two almost coincide. The first contribution to the Yang-Mills β-function which is reproduced by Eq. (4.25) is due to the longitudinal gluons, or in other words to the implementation of Gauss' law. Since we have implemented Gauss' law exactly on our trial wave function, this (major) part of the β-function is correctly reproduced by the σ-model renormalization group. The second contribution, which is not present in Eq. (4.25), is due to the dynamics of transverse gluons. The fact that this contribution is missing in the σ-model suggests that our Ansatz is not perfect in the ultraviolet. However, as this contribution is relatively small, we will not be discouraged at this stage. We will simply think of Eq. (4.25) as representing the complete one loop Yang-Mills β-function, keeping in mind that it would indeed be very interesting to eliminate this discrepancy by perhaps exploring a less simplistic form of the variational propagator $G(k)$ [38].

Since the σ-model is asymptotically free, perturbation theory becomes worse and worse as we go to lower momenta, and at some point becomes inapplicable.

Now, however, let us look at the other side of the coin: let us see how the action looks for the matrices $U(x)$ which are slowly varying in space. Due to the short range of $G(x)$, for $U(x)$ which contain only momenta lower than the variational scale M the action is local. In fact, with our Ansatz Eq. (4.13), it becomes the standard local action of the non-linear σ-model

$$\Gamma_L[U] = \frac{M}{2g^2}\,\mathrm{tr}\int_x \partial_i U^\dagger(x)\partial_i U(x) + \dots \ . \tag{4.27}$$

In this low-momentum approximation we also neglected the space dependence of $S^{ab}_{ij}(x)$ in the term SGS^T in Eq. (4.17); then, using the fact that S is an orthogonal matrix Eq. (4.21), one gets $SGS^T \to G$.

Strictly speaking, due to the Z_N local symmetry of Eq. (4.17), the action for the low momentum modes is slightly different. The derivatives should be understood as Z_N covariant derivatives. The most convenient way to write this action would be to understand $U(x)$ as belonging to $U(N)$ rather than $SU(N)$ and introduce a $U(1)$ gauge field by

$$\Gamma_L = \frac{1}{2}\frac{M}{g^2}\,\mathrm{tr}\int_x (\partial_i - ib_i)U^\dagger(x)(\partial_i + ib_i)U(x)\,. \tag{4.28}$$

This defines a sigma model on the target space $U(N)/U(1)$, which is isomorphic to $SU(N)/Z_N$. This subtlety is unimportant for large N and will not play a significant role in our analysis.

The action Eq. (4.27) does not look too bad. Even though it still cannot be solved exactly, it is amenable to analysis by standard methods, such as the mean field approximation, which in 3 dimensions and for large number of fields should give reasonably reliable results.

We adopt therefore the following strategy for dealing with the integration over the $SU(N)$ group. We integrate perturbatively the high momentum modes of the field $U(x)$. This is the renormalization group (RG) transformation. We would like to integrate out all modes with momenta $k^2 > M^2$. This procedure will necessarily generate a *local* effective action for the low momentum modes. At the same time, because of the (presumed) equivalence of the RG flows in QCD and our effective sigma-model, the effective coupling constant will be the the running QCD coupling constant $\alpha_{QCD}(M)$ at scale M. This part of the theory can then be solved in the mean field approximation. Clearly, in order for the perturbative RG transformation to be justified, the QCD running coupling constant at the scale M must be small enough. Our procedure makes sense provided the energy is minimized at a value of the variational parameter M for which

$$\alpha_{QCD}(M) < 1\,. \tag{4.29}$$

We will check whether this consistency condition is satisfied at the end of the calculation. In the next section we will calculate the expectation value of the Hamiltonian in the lowest order of this approximation scheme, and perform the minimization with respect to M.

4.2. *Solving the variational equations; dynamical mass generation*

We will now calculate the energy, i.e. the expectation value of the Hamiltonian Eq. (4.1). We first perform the Gaussian integrals over the vector potential at fixed $U(x)$. Let us consider, for example, the calculation of the chromoelectric energy,

$$\begin{aligned}\int_x \langle E_i^{a2}\rangle_A &= \int_x \left\langle -\frac{\delta}{\delta A_i^a(x)}\frac{\delta}{\delta A_i^a(x)}\right\rangle_A \\ &= \mathrm{Tr}\, G^{-1} - \int_{x,y,z} G^{-1}(x-y)G^{-1}(x-z)\langle A_i^a(y)A_i^a(z)\rangle_A\,. \end{aligned}\tag{4.30}$$

The Gaussian averaging over A is easily performed. Defining for convenience

$$a_i^a(x) = \int_{y,z} \lambda_i^b(y)G^{-1}(y-z)S^{bc}(z)(\mathcal{M}^{-1})^{ca}(z,x)\,, \tag{4.31}$$

one gets

$$\begin{aligned}\int_x \langle E_i^{a2}\rangle_A = 3(N^2-1)\int_x G^{-1}(x,x) - \int_x (G^{-1}\mathcal{M}^{-1}G^{-1})^{aa}_{ii}(x,x) \\ - \int_{x,y} a_i^a(x)G^{-2}(x-y)a_i^a(y)\,. \end{aligned}\tag{4.32}$$

For the chromomagnetic contribution the calculation is straightforward and one gets

$$\langle(\epsilon_{ijk}\partial_j A_k^a)^2\rangle_A = (\epsilon_{ijk}\partial_j a_k^a)^2 + \epsilon_{ijk}\epsilon_{ilm}\partial_i^x\partial_l^y(\mathcal{M}^{-1})^{aa}_{km}(x,y)|_{x=y}\,, \tag{4.33}$$

$$\begin{aligned}\langle\partial_j A_k^a A_l^b A_m^c\rangle_A &= \partial_j a_k^a a_l^b a_m^c + \partial_j a_k^a(\mathcal{M}^{-1})^{bc}_{lm}(x,x) \\ &+ a_l^b\partial_j^x(\mathcal{M}^{-1})^{ac}_{km}(x,y)|_{x=y} + a_m^c\partial_j^x(\mathcal{M}^{-1})^{ab}_{kl}(x,y)|_{x=y}\,, \end{aligned}\tag{4.34}$$

and

$$\begin{aligned}\epsilon_{ijk}\epsilon_{ilm}f^{abc}f^{ade}\langle A^b_j A^c_k A^d_l A^e_m\rangle_A &= 2f^{abc}f^{ade}a^b_j a^c_k a^d_l a^e_m \\ &\quad + 8f^{abc}f^{ade}a^b_i a^d_i(\mathcal{M}^{-1})^{ce}(x,x) \\ &\quad + 12f^{abc}f^{ade}(\mathcal{M}^{-1})^{bd}(x,x)(\mathcal{M}^{-1})^{ce}(x,x)\,. \end{aligned} \tag{4.35}$$

Here, we have used the obvious notation $\mathcal{M}^{ab}_{ij} = \mathcal{M}^{ab}\delta_{ij}$. The next step is to decompose the matrix field $U(x)$ into low and high momentum modes. In general this is a non-trivial problem. However, since we are only going to integrate over the high momenta in the lowest order in perturbation theory, for the purposes of our calculation we can write

$$U(x) = U_L(x)U_H(x)\,, \tag{4.36}$$

where U_L contains only modes with momenta $k^2 < M^2$, and U_H has the form $U_H = 1 + ig\tau^a\phi^a_H$ and ϕ_H contains only momenta $k^2 > M^2$. This decomposition is convenient, since it preserves the group structure. Also, since the measure DU is group invariant, we can write it as $DU_L DU_H$. With this decomposition we have

$$\lambda^a_i(x) = S^{ab}_H(x)\lambda^b_{iL}(x) + \lambda^a_{iH}(x)\,. \tag{4.37}$$

Further simplifications arise, since we only have to keep the leading piece in ϕ^a_H. In this approximation:

$$\begin{aligned} S^{ab}(x) &= S^{ab}_L(x)\,, \\ \mathcal{M}^{ab}(x,y) &= 2\delta^{ab}G^{-1}(x-y)\,, \\ \lambda^a_i(x) &= \lambda^a_{iL}(x) + \lambda^a_{iH}(x)\,, \\ a^a_i(x) &= \frac{1}{2}\,\lambda^a_{iL}(x) + \frac{1}{2}\,\lambda^b_{iH}(x)S^{ba}_L(x)\,. \end{aligned} \tag{4.38}$$

The chromoelectric part of the energy can then be written

$$\begin{aligned}\int_x \langle E^{a2}_i\rangle_A = \frac{3(N^2-1)}{2}\int_x G^{-1}(x,x) &- \frac{1}{4}\int_{x,y}\lambda^a_{iL}(x)G^{-2}(x-y)\lambda^a_{iL}(y) \\ &- \frac{1}{4}\int_{x,y}\lambda^a_{iH}(x)G^{-2}(x-y)\lambda^a_{iH}(y)\,. \end{aligned} \tag{4.39}$$

The cross term vanishes since to this order, as we shall see, there is a decoupling between the high and the low momentum modes in the action, and therefore the product factorizes, and also $\langle\lambda^a_{iH}\rangle = 0$. The Ansatz Eq. (4.13) allows us to simplify this expression further. Recalling that $\lambda_L(x)$ contains only momenta below M, it is immediate to see that

$$\int_{x,y} \lambda^a_{iL} G^{-2}(x-y)\lambda^a_{iL}(y) = M^2 \int_x \lambda^a_{iL}(x)\lambda^a_{iL}(x)\,. \tag{4.40}$$

We can then rewrite Eq. (4.39) as

$$\begin{aligned}\int_x \langle E^{a2}_i\rangle_A &= \frac{3(N^2-1)}{2}\int G^{-1}(x,x)\\ &\quad - \frac{M^2}{4}\int_x \lambda^a_{iL}(x)\lambda^a_{iL}(x) - \frac{1}{4}\int_{x,y}\lambda^a_{iH}(x)G^{-2}(x-y)\lambda^a_{iH}(y)\,.\end{aligned} \tag{4.41}$$

The contribution of the magnetic term to the energy is very simple. All cross terms between the low and high momentum modes drop out. Some vanish for the same reason as the cross terms in Eq. (4.39), and others because they are explicitly multiplied by a power of the coupling constant. Since our approximation is the lowest order in g, except for the non-analytic contributions that come from the low mode effective action, those terms do not contribute. In fact, the entire low momentum mode contribution drops out of this term. The reason for this is that the only term which could give a leading order contribution, i.e.

$$\int (\epsilon_{ijk}\partial_j \lambda^a_{kL})^2 \tag{4.42}$$

can be rewritten as

$$(f^a_{ijL})^2 + O(g^2)\,, \tag{4.43}$$

where f^a_{ijL} is the "magnetic field" corresponding to the "vector potential" λ^a_{iL}. However, λ_L has the form of a pure gauge vector potential. Therefore $f^a_{ijL} = 0$, and the contribution of this term is higher order in g^2. One can check a posteriori that including this term indeed changes the energy density in the best variational state by a small amount ($O(10\%)$), but has no effect at all on the best value of the variational parameter M. The entire magnetic field contribution to the energy is then

$$\frac{1}{2}\langle B^{a2}_i\rangle_A = \frac{1}{8}(\epsilon_{ijk}\partial_j\lambda^a_{kH})^2 + \frac{N^2-1}{2}\partial^x_i\partial^y_i G(x-y)|_{x=y}\,. \tag{4.44}$$

The last step is to perform an averaging over the U-field. For convenience, we rewrite the complete expression for the energy density (here $V = \int d^3x$

is a space volume)

$$\frac{\langle 2H\rangle}{V} = \frac{3(N^2-1)}{2}\,G^{-1}(x,x) + (N^2-1)\,\partial_i^x\partial_i^y G(x-y)|_{x=y}$$
$$-\frac{1}{4V}\int_{x,y}\langle\lambda^a_{iH}(x)G^{-2}(x-y)\lambda^a_{iH}(y)\rangle_U + \frac{1}{4}\langle(\epsilon_{ijk}\partial_j\lambda^a_{kH})^2\rangle_U$$
$$-\frac{M^2}{4V}\int_x\langle\lambda^a_{iL}(x)\lambda^a_{iL}(x)\rangle_U\,, \tag{4.45}$$

where the averaging over the U-field should be performed with the sigma model action Eq. (4.17). In our approximation this action has a simple form. Using Eq. (4.38) we obtain

$$\Gamma = \frac{1}{4}\int_{x,y}\lambda^a_{iH}(x)G^{-1}(x-y)\lambda^a_{iH}(y) + \frac{M}{4}\int_x\lambda^a_{iL}(x)\lambda^a_{iL}(x)\,. \tag{4.46}$$

The low momentum mode part is precisely equal to Γ_L in Eq. (4.28). The only difference is that the coupling constant that appears in this action should be understood as the running coupling constant at the scale M. This obviously is the only $O(1)$ effect of the high momentum modes on the low momentum effective action

$$\Gamma_L = \frac{1}{2}\,\frac{M}{g^2(M)}\,\mathrm{tr}\int_x(\partial_i - ib_i)U^\dagger(x)(\partial_i + ib_i)U(x)\,. \tag{4.47}$$

We are now in a position to evaluate the VEV of the energy. The contribution of the high momentum modes is immediately calculable. Using the parameterization $U_H(x) = 1 - \frac{i}{2}g\phi^a\tau^a$, we find that ϕ^a are free fields with the propagator

$$\langle\phi^a(x)\phi^b(y)\rangle = 2\delta^{ab}[\partial_i^x\partial_i^y G^{-1}(x-y)]^{-1}\Big|_{p^2>M^2}\,. \tag{4.48}$$

Also to this order $\lambda^a_{iH}(x) = \partial_i\phi^a(x)$ and therefore $\epsilon_{ijk}\partial_j\lambda^a_{kH} = 0$. Using Eq. (4.48) one finds

$$\frac{1}{4}\int_{x,y}\langle\lambda^a_{iH}(x)G^{-2}(x-y)\lambda^a_{iH}(y)\rangle_U = V\,\frac{N^2-1}{2}\int_M^\Lambda\frac{d^3k}{(2\pi)^3}\,G^{-1}(k)\,, \tag{4.49}$$

where Λ is the ultraviolet cutoff, and the contribution of the high momentum modes to the energy (first two lines in Eq. (4.45)) is

$$\begin{aligned}\frac{2E_0}{V} &= (N^2-1)\left\{\int_0^\Lambda \frac{d^3k}{(2\pi)^3}\left[G^{-1}(k)+k^2G(k)\right]+\frac{1}{2}\int_0^M \frac{d^3k}{(2\pi)^3}\,G^{-1}(k)\right\} \\ &= \frac{N^2-1}{2\pi^2}\left\{\int_0^M k^2dk\left[\frac{3}{2}M+\frac{k^2}{M}\right]+2\int_M^\Lambda k^3dk\right\} \\ &= \frac{N^2-1}{10\pi^2}\,M^4+\dots\,. \end{aligned} \tag{4.50}$$

Terms omitted in Eq. (4.50) depend on Λ, but are independent of the variational scale M.

We now have to evaluate the contribution of the low momentum modes. It is clear from the form of the action Eq. (4.47) that this contribution, as a function of M, will not be featureless. The most convenient way to think about it is from the point of view of classical statistical mechanics. Comparing Eqs. (4.45) and (4.47), we see that we have to evaluate the internal energy of the sigma model (with the UV cutoff M) at a temperature proportional to the running coupling constant $g^2(M)$. For large M, the coupling constant is small, which corresponds to the low temperature regime of the sigma model. In this regime the global $SU(N)\otimes SU(N)$ symmetry group of the model is spontaneously broken. Lowering M, we raise $g^2(M)$, and therefore the temperature. At some critical value g_C, the σ-model undergoes a phase transition into the unbroken (disordered) phase. Clearly, in the vicinity of the phase transition all thermodynamical quantities will vary rapidly, and therefore this is a potentially interesting region of coupling constants.

Before analyzing the phase transition region let us calculate $E(M)$ for large M. In this regime the low momentum theory is weekly coupled. The calculation is straightforward, and to lowest order in g^2 gives

$$\frac{1}{4}\,M^2\langle\lambda^a_{iL}(x)\lambda^a_{iL}(x)\rangle = \frac{N^2-1}{12\pi^2}\,M^4\,. \tag{4.51}$$

Putting this together with the high momentum contribution, we find

$$\frac{E(M)}{V} = \frac{N^2-1}{120\pi^2}\,M^4, \qquad M\gg\Lambda_{QCD}\,. \tag{4.52}$$

This indeed is the expected result. The energy density monotonically increases as M^4, with a slope which is given by the standard perturbative expression. Note, however, that the slope is very small, and the contribution of the low momentum modes to the energy is negative. Therefore, if the internal energy of the sigma model grows significantly in the phase transition region, the sign of $E(M)$ could be reversed [e] and the energy will then be minimized for M in this region.

To see whether this indeed happens, we will now study the low momentum sigma model in the mean field approximation. We rewrite the partition function by introducing a (hermitian matrix) auxiliary field σ which imposes a unitarity constraint on $U(x)$

$$Z = \int DUD\sigma Db_i \exp\left(-\Gamma[U,b,\sigma]\right) , \tag{4.53}$$

with

$$\Gamma[U,b,\sigma] = \frac{M}{2g^2(M)} \mathrm{tr} \int_x \Big[(\partial_i - ib_i)\, U^\dagger(x)\, (\partial_i + ib_i)\, U(x) \\ + \sigma \left(U^\dagger U - 1\right)\Big] \, . \tag{4.54}$$

The role of the vector field b_i is to impose a $U(1)$ gauge invariance and, thereby, to eliminate one degree of freedom. As far as the thermodynamical properties are concerned, its effect is only felt as an $O(1/N^2)$ correction. At the level of accuracy of the mean field approximation, we can safely disregard it, which we do in the following. The mean field equations are

$$\langle U^\dagger U \rangle = 1 \, , \tag{4.55}$$

$$\langle \sigma U \rangle = 0 \, . \tag{4.56}$$

From Eq. (4.56) it follows that either $\langle\sigma\rangle = 0$, $\langle U\rangle \neq 0$ (the ordered, broken symmetry phase with massless Goldstone bosons), or $\langle\sigma\rangle \neq 0$, $\langle U\rangle = 0$ (the disordered, unbroken phase with massive excitations). We are mostly interested in the disordered phase, since there the mean field approximation should be reliable. Since the symmetry is unbroken, the expectation value of σ should be proportional to a unit matrix

$$\langle \sigma_{\alpha\beta} \rangle = \sigma^2 1_{\alpha\beta} \, . \tag{4.57}$$

Equation (4.55) then becomes

$$2N^2 \frac{g^2(M)}{M} \int_0^M \frac{d^3k}{(2\pi)^3} \frac{1}{k^2+\sigma^2} \\ = \frac{N^2 g^2(M)}{\pi^2} \left(1 - \frac{\sigma}{M} \arctan \frac{M}{\sigma}\right) = N \, . \tag{4.58}$$

[e] The energy, of course, never becomes negative, since Eq. (4.50) contains a divergent M-independent piece. Here we concentrate only on the M-dependence of E.

The gap equation, Eq. (4.58), has solution only for couplings (temperatures) $g^2(M)$ larger than the critical coupling (temperature) g_C^2, which is determined by the condition that $\sigma = 0$,

$$\alpha_C = \frac{g_C^2}{4\pi} = \frac{\pi}{4}\,\frac{1}{N}\,. \tag{4.59}$$

The low momentum mode contribution to the ground state energy is

$$N^2 M \int_0^M \frac{d^3k}{(2\pi)^3}\,\frac{k^2}{k^2+\sigma^2} = \frac{N^2}{2\pi^2}\,M\left[\frac{1}{3}M^3 - \sigma^2 M + \sigma^3 \arctan\frac{M}{\sigma}\right]\,. \tag{4.60}$$

The final mean field expression for the ground state energy density is [f]

$$E = \frac{N^2}{4\pi^2}\,M^4\left[-\frac{2}{15} + \frac{\sigma^2}{M^2}\frac{\alpha_C}{\alpha(M)}\right]\,, \tag{4.61}$$

where $\alpha(M)$ is the QCD coupling at the scale M, α_C is given by Eq. (4.59), and σ is determined by

$$\frac{\sigma}{M}\arctan\frac{M}{\sigma} = \frac{\alpha(M)-\alpha_C}{\alpha(M)}\,. \tag{4.62}$$

The energy as a function of M is plotted in Fig. 1 for $N = 3$. Qualitatively it is the same for any N. The minimum of the energy is obviously at the point $\alpha(M) = \alpha_C$. Using the one-loop Yang-Mills β function and $\Lambda_{QCD} = 150\,\mathrm{MeV}$, we find for $N = 3$

$$M = \Lambda_{QCD}\,e^{\frac{24}{11}} = 8.86\,\Lambda_{QCD} = 1.33\,\mathrm{GeV}\,. \tag{4.63}$$

We thus find that the best variational state is non-perturbative and is characterized by a dynamically generated mass scale. To get a rough idea whether the value of this mass is reasonable we have calculated the value of the gluon condensate $(\alpha/\pi)\langle F^a_{\mu\nu}F^a_{\mu\nu}\rangle = (2\alpha/\pi)\big(\langle B_i^{a\,2}\rangle - \langle E_i^{a\,2}\rangle\big)$. These calculations are straightforward, and yield [g]

$$\langle E_i^{a\,2}\rangle = -\frac{1}{24\pi^2}\,N^2M^4\,, \qquad \langle B_i^{a\,2}\rangle = -\frac{1}{40\pi^2}\,N^2M^4\,, \tag{4.64}$$

[f] We do not distinguish between N^2 and N^2-1 since we have neglected the contribution of the $U(1)$ gauge field. The errors are of order $1/N^2$ and even for $N = 2$ are very likely smaller than the error introduced by using the mean field approximation in the first place.

[g] We have again kept only the M-dependent pieces. Each one of the quantities $\langle E_i^{a\,2}\rangle$ and $\langle B_i^{a\,2}\rangle$ is of course positive, due to positive UV divergent, but M independent, pieces. It is easy to check that the energy density $E = 1/2(\langle E_i^{a\,2}\rangle + \langle B_i^{a\,2}\rangle) = -(1/30\pi^2)N^2M^4$ coincides with the first term in Eq. (4.61), as it must.

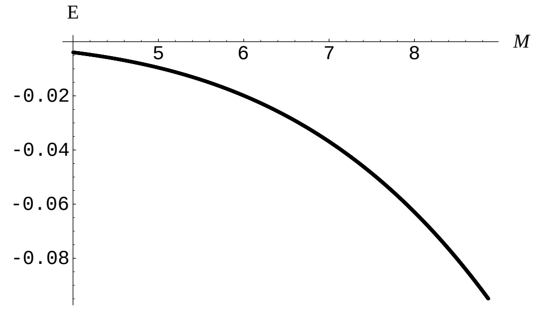

Figure 1. Energy density of a variational trial state as a function of the variational parameter M in units of Λ_{QCD}. The energy is only shown for $M < 8.86\Lambda_{QCD}$, which corresponds to the disordered phase of the effective low momentum σ-model.

and finally

$$\frac{\alpha}{\pi}\,\langle F^a_{\mu\nu}F^a_{\mu\nu}\rangle = \frac{N}{120\pi^2}\,M^4 = 0.008\ \mathrm{GeV}^4\,. \tag{4.65}$$

The preferred phenomenological value of this condensate is 0.012 GeV^4 [15], although the uncertainty in this number is large. Considering this, and the unsophisticated nature of our calculation, the agreement is quite reasonable.

A natural question is of course whether one can assign to M directly the meaning of some physical mass? In the initial Gaussian wave function, before the projection, it appears as the "gluon mass". However the projection changes the wave functional very considerably, and the direct meaning of M (apart from it being a dimensional variational parameter) is not so clear. Nevertheless, naively one expects that, as the gauge invariant operators $E^2 - B^2$ and EB to leading order in g^2 are quadratic in the gluon field, the scale $1/2M$ should appear as the correlation length in these correlators. Thus we are tempted to identify $2M$ with the glueball mass. An attempt to calculate glueball correlation functions was made in [43]. The result is somewhat unexpected. It was found in [43] that the scale $2M$ does indeed dominate the long distance behavior of the correlation function of the pseudoscalar glueball. However it appears that the scalar glueball correlator at long distances is dominated by the scale σ – the gap of the effective σ-model. If correct, this would mean that the phenomenologically acceptable scalar glueball mass can only arise in our calculation if the phase transition in the effective σ-model is of the first order with significant latent heat (see next subsection). This point clearly warrants further investigation.

4.3. *Is it safe?*

For $N = 3$, the value of the QCD coupling constant at the variational scale is $\alpha_C = 0.26$. It is reasonably small, so that the consistency condition for the perturbative integration of the high momentum modes is satisfied. However, it is not so small that higher order corrections would be negligible. We expect therefore that including higher orders in perturbation theory could give corrections to our result for $\alpha(M)$ of order 25%. Since M depends exponentially on $\alpha(M)$, such change in α may change the value of M by a factor of $2 - 3$. Consequently our result for the dynamical scale M and other dimensional quantities should be taken only as an order of magnitude estimate. In particular, the value of the condensate $\langle F^2 \rangle$ is proportional to the fourth power of M, and would change dramatically as a result of a moderate change of α_c. This is the normal state of affairs in theories with a logarithmically running coupling constant. The best accuracy is always achieved for dimensionless quantities, since those are usually slowly varying functions of α. The overall scale depends on α exponentially, and therefore always has the largest error.

The use of the mean field approximation to analyze the σ-model introduces uncertainties into the result. As a rule, the mean field approximation gives a good estimate of the critical temperature. Sometimes, however, it gives wrong predictions for the order of the phase transition. We believe that this is indeed the case here. The mean field approximation indicates that the phase transition is second order. The mass gap in the sigma model vanishes continuously at the critical point. The universality class describing the symmetry breaking pattern $(SU(N) \otimes SU(N))\,/SU(N)$ was considered in the context of finite temperature chiral phase transition in QCD. The results of ϵ expansion [44] and also numerical simulations [45, 46] strongly suggest that the phase transition is of first order except for $N = 2$. In our case there is an additional Z_N symmetry in the game. However its presence is likely only to increase the latent heat rather than turn the transition into a second order one. The reason is that the Z_N gauge invariant theory allows for the existence of topological defects, the Z_N strings, and condensation of topological defects frequently leads to discontinuous phase transitions.

Nevertheless, we believe that the bulk of our results is robust against this uncertainty. The mean field approximation should be reliable in the regime where the mass gap in the sigma model is not too small. At the point $M = 4.5\Lambda_{QCD}$ we find

$$\sigma = 0.23M, \qquad \alpha(M) = 0.38\,. \tag{4.66}$$

Since the gap of the σ-model is at this point of the order of the UV cutoff, the mean field approximation should be reliable here. Perturbation theory is also still reasonable at this value of α. The fact that the energy is negative and has a minimum for some $\alpha(M) < 0.38$, seems to be, therefore, unambiguous.

In fact, independently of the mean field calculation, it is physically very plausible that the energy is minimized precisely at the critical temperature, on the disordered side of the phase transition (if it is of the first order). Consider first, the contribution of the high momentum modes to the ground state energy, Eq. (4.50). It is proportional to M^4 with a fixed (M-independent) proportionality coefficient $x = (N^2-1)/10\pi^2$. Consider now the low momentum contribution in the large M region, Eq. (4.51). It is again proportional to M^4 with the coefficient $y_0 = N^2/12\pi^2$. The proportionality coefficient of the low momentum contribution at the phase transition point, according to our calculation, is twice as big $y_C = 2N^2/12\pi^2$. This is physically quite transparent. In the large M, low-temperature, regime the global $SU(N)\otimes SU(N)$ symmetry of the sigma model is broken down spontaneously to $SU(N)$. This leads to the appearance of $N^2 - 1$ massless Goldstone bosons. At zero temperature, those are the only propagating degrees of freedom in the model. All the rest have masses of the order of the UV cutoff, and therefore do not contribute to the internal energy. When the temperature is raised (M is lowered), the Goldstone bosons remain massless and other excitations become lighter. If the transition is second order, at the phase transition point the symmetry is restored, and one should have a complete multiplet of the $SU(N)\otimes SU(N)$ symmetry of massless particles. The dimensionality of this multiplet is $2N^2$. The contribution of every degree of freedom to the internal energy is still roughly the same as at zero temperature. This is so, since, although at the phase transition the particles are interacting, critical exponents of scalar theories in 3 dimensions are generally very close to their values in a free theory [47]. The internal energy of the σ-model at this point therefore should be roughly twice its value at zero temperature. Moving now to higher temperatures, all the particles become heavier, and therefore their contribution to internal energy decreases. The internal energy therefore should have a maximum at the phase transition temperature.

Note that the ground state energy of the Yang-Mills theory is the difference between the high momentum contributions and the internal energy of the low mode sigma model. Already at zero temperature, these two contributions differ only by 20%, which is why the coefficient in the expression Eq. (4.52), even though positive, is so small. At the critical point, where the low momentum mode internal energy is twice as large, the chances of the slope becoming negative are very good. This is what happens in our mean field analysis but, according to the previous argument, this in large measure is independent of the approximation. If the phase transition is first order one should be more careful. The internal energy then changes discontinuously across the phase transition. The particles in the disordered phase are always massive, and the internal energy is smaller than in the case of the second order phase transition. However, if the transition is only weakly first order the same argument still holds (the fact that the mean field predicts a second order phase transition may be an indication that if it is in fact first order, it is only weakly so). It does seem quite likely that the Yang-Mills ground state energy will become negative, since it only needs the internal energy of the σ-model to grow by 20% at the phase transition relative to its zero temperature limit. In this case there will be a finite latent heat, which means that the internal energy in the disordered (high temperature) phase is higher. The Yang-Mills variational ground state energy, therefore, will have its minimum in the disordered phase.

There is good reason to believe, therefore, that these results are qualitatively correct, and will survive the improvement of the approximation.

4.4. *Instantons*

Instantons are believed to play a very important role in the non-perturbative dynamics of QCD [13]. It is, therefore, interesting to see if our variational calculation has any relation to the instanton approach.

The first thing is to understand how we expect to see instantons in this formalism. Instantons are localized, finite-action classical solutions of the field equations of QCD in Euclidean space-time. Physically they represent tunneling processes between topologically distinct vacuum sectors with the exponent of the instanton action being equal to the transition probability between two of these vacuum states. Although the notion of the instanton is intrinsically Euclidean, the tunneling between different vacuum sectors can be formulated both in the Hamiltonian and the Lagrangian languages. In fact, the gauge projected variational approach we discuss here is very well suited for this purpose. The projection of the initial Gaussian onto the gauge invariant subspace is achieved by the integration over the gauge group. The effective σ-model arises as an integration over the relative gauge transformations between the two Gaussian states in the linear superposition Eq. (4.5). The Boltzmann factor $\exp(-\Gamma[U])$ for a given matrix U is therefore just the overlap of the initial and the gauge rotated state or, in other words, the transition amplitude between the two states [h]. The instanton transition is precisely a transition of this type, where the two states (at $t \to -\infty$ and $t \to +\infty$) are related by a large gauge transformation. The matrix of this large gauge transformation must carry a non-zero topological charge $\Pi_3(SU(N))$.

The integration measure over U indeed includes integration over topologically non-trivial configurations. The finiteness of the action Eq. (4.17) requires that the matrix U approaches a constant value at infinity. This identifies all points at spatial infinity, hence the physical space of the model is S^3. Field configurations are maps from S^3 into the manifold of $SU(N)$ and are classified by their winding number, or topological charge, which is an element of the homotopy group $\Pi_3(SU(N)) = Z$. The σ-model action in a given topological sector is minimized on some configuration which is a solution of classical σ-model equations of motion. In particular, the solution with a unit topological charge is expected to have a "hedgehog" structure much like the topological soliton in the Skyrme model [12]. The integral over U in the steepest descent approximation is saturated by these classical solutions.

These σ-model configurations that belong to a non-trivial topological sector with a unit winding number represent QCD transitions between the topologically distinct sectors. The topologically non-trivial classical soliton solutions of the σ-model are therefore the three dimensional images of the QCD instantons.

The QCD instantons are defined in space-time and are therefore four dimensional point-like objects. The σ-model solutions are intrinsically three dimensional. Nevertheless, there is a natural simple relation between the two. For a given Yang-Mills instanton solution $A^{\rm inst}(x_\mu)$ one can find a three dimensional $SU(N)$ matrix $U(x_i)$ by the procedure discussed by Atiyah and Manton [48],

$$U_{\rm AM}(x_i) = P \exp\left(i \int_C dx^\mu A^{\rm inst}_\mu \right) , \tag{4.67}$$

where the contour of integration C is a straight line $x_i = const$, $-\infty < x_0 < \infty$. The matrix $U_{\rm AM}$ gives the relative gauge transformation between the initial trivial vacuum at $x_0 \to -\infty$ and the topologically non-trivial vacuum at $x_0 \to +\infty$ or, in other words, between the initial and final states of the instanton transition. Clearly, its meaning is precisely the same as

[h] Since the space of the matrices U is continuous, strictly speaking the Boltzmann factor is the differential rather than the total amplitude.

that of the classical soliton solution of the effective σ-model Eq. (4.17). Also, the QCD instanton action and the σ-model soliton action have the same physical meaning. They both give the transition probability between different topological sectors in QCD. We will therefore refer to the σ-model solitons as instantons in the following.

Although the QCD and the σ-model instantons have the same physical meaning, it is not assured that the numerical value for their respective actions is the same. They both approximate the value of the transition probability in QCD, but the approximations involved are quite different. The QCD instanton action is the result of the standard WKB approximation which is valid at weak coupling and therefore for small instantons, but breaks down for instantons of large size. The σ-model instanton action on the other hand is the value of this transition probability in a particular Gaussian variational approximation. It is natural to expect that the variational calculation underestimates the value of the transition probability at very weak coupling. The transition probability is given by the overlap of the "ground state" wave functions in two topological sectors. For simplicity let us consider a quantum mechanical system with two vacua at $x_\pm$. If the area below the barrier separating the vacua is large, the standard WKB instanton calculation is applicable. The wave function of each of the vacua below the barrier has essentially an exponential fall off $\exp\{i\int^x \sqrt{E - V(x - x_\pm)}\}$. The instanton calculation is the calculation of the overlap of these functions. Our variational calculation corresponds to approximating the respective "ground states" at $x_\pm$ by Gaussian wave functions. The tails of the Gaussians fall off much faster away from the minimum than the actual wave function and the overlap is therefore expected to be smaller. When the coupling constant is not too small (or when the area below the barrier is not too large) the overlap between the two states is no longer determined by the behavior of the "tails" of the wave functions. In this situation one can expect the Gaussian approximation to do much better, since the overlap region contributes significantly to the energy and therefore plays an important role in the minimization procedure.

In fact, implicitly, the instantons played a very important role already in the energy minimization described in this section. As we have seen above, the energy is minimized for the value of the mass parameter M at which the σ-model is in the disordered phase. The transition between ordered and disordered phases in a statistical mechanical system can usually be described as a condensation of topological defects. This is a standard description of the phase transition in the Ising and XY models [49]. In the σ-model Eq. (4.17) the relevant topological defects are none other than the instantons. In this sense the appearance of the dynamical mass in the best variational state is itself driven by the condensation of instantons.

Perhaps the most significant difference between the kinks in the Ising model and the QCD instantons, is that the former have a fixed size, while the latter come in a variety of sizes. This is a direct consequence of the dilatation symmetry of the classical Yang-Mills action. It is not necessary for instantons of all sizes to condense in order to drive the transition. The naive expectation therefore is that the large size instantons (larger than $1/M$) condense, while the smaller ones should still exist as semi-classical solutions in the effective σ-model action.

The simple qualitative argument to this effect is the following. Consider the effective σ-model action for very large size instantons. In such a configuration only field modes with small momentum $k < M$ are present. For these momenta the action is the standard local σ-model where M plays the role of an ultraviolet cutoff

$$\Gamma = \frac{1}{2}\frac{M}{g^2(M)}\mathrm{tr}\int_x \partial_i U^\dagger(x)\partial_i U(x)\,. \tag{4.68}$$

If the large size instantons are stable at all, they should also be present as stable solutions in this local action Eq. (4.68). However this is not the case as can be easily seen by the standard Derrick type scaling argument. Take an arbitrary configuration $u(x)$ in the instanton sector and scale all the coordinates by a common factor λ. Then obviously

$$\Gamma[u(\lambda x)] = \lambda^{-1}\Gamma[u(x)]\,. \tag{4.69}$$

The dependence of the action on λ is monotonic and is minimized at $\lambda \to \infty$. This means that the instantons in the local σ-model shrink to the ultraviolet cutoff $1/M$. For instantons smaller than the inverse cutoff we cannot use the local action anymore. However, the behavior of these small size instantons is already familiar. We know that classically they exist at arbitrary size, but that when the running of the coupling is taken into account, these instantons are pushed to the large size. This is the familiar infrared problem of large instantons. In our variational state, the coupling constant stops running at the scale M. The picture is therefore very simple. The small size instantons are pushed to larger size by the effect of the coupling constant, while the large size instantons are pushed to smaller size by the effect of the local σ-model scaling. We therefore expect that the instanton size will be stabilized somewhere in the vicinity of $\rho \sim 1/M$.

The behavior of the instantons in the variational ansatz Eq. (4.5) was studied in detail in [50]. The results are indeed very much in line with the expectations just outlined. The action of a small size instanton in the σ-model was found to be independent of its size (neglecting the running coupling effects), and numerically equal to

$$\Gamma = 1.96\,\frac{8\pi^2}{g^2}\,. \tag{4.70}$$

This is about twice the value of the instanton action in QCD, $\Gamma_{inst} = \frac{8\pi^2}{g^2}$. Thus, as expected, the tunneling transition amplitude is underestimated in the Gaussian approximation for small instantons. Interestingly enough, however, the actual configuration of the σ-model field that minimizes the action in the one instanton sector was found to be practically indistinguishable from the Atiyah-Manton expression calculated on the QCD instanton. This means that even though the value of the transition probability is underestimated in the Gaussian approximation, the actual field configurations into which the tunneling is most probable are identified correctly; they are precisely the same as in the WKB calculation.

As for the large size instantons, when the running of the coupling constant is taken into account their size is stabilized at about $\rho = (1 - 1.5)/M$. The uncertainty is to do with the way the running of the coupling constant is modified at $k < M$. It is in fact interesting to note that this instanton size is consistent with the average size of the instantons in the instanton liquid model of [51–55]. For the case of $SU(2)$, the average instanton size, in units of the gluon condensate obtained in the instanton liquid model, turns out to be [53–55],

$$\rho\, \left(\langle F^a_{\mu\nu}F^a_{\mu\nu}\rangle\alpha/\pi\right)^{1/4} \sim 0.4\,. \tag{4.71}$$

In our case, taking the value of the gluon condensate obtained in the variational approach, we find

$$\rho\, \left(\langle F^a_{\mu\nu}F^a_{\mu\nu}\rangle\alpha/\pi\right)^{1/4} \sim 0.2 - 0.3\,. \tag{4.72}$$

The relation to the instanton liquid model is an interesting question which deserves further study.

We also note that the variational Ansatz which has been considered so far corresponds to a zero value of the QCD θ-parameter, since we have integrated over the entire gauge group without any extra phases. As is well known, the general θ-vacuum is defined as

$$|\theta\rangle = \sum_n e^{in\theta}|n\rangle\,, \tag{4.73}$$

where n labels the topological sectors in the configuration space (space of all potentials $A_i^a(x)$). Generalization of our trial wave functions to non-zero θ is trivial; all we need to do is to insert in the integrand of Eq. (4.5) the extra phase factor

$$\exp\left\{i\frac{\theta}{24\pi^2}\int dx\epsilon_{ijk}\mathrm{tr}\left[(U^\dagger\partial_i U)(U^\dagger\partial_j U)(U^\dagger\partial_k U)\right]\right\}\,. \tag{4.74}$$

The integrand here is a properly normalized topological charge, and it takes integer values for topologically non-trivial configurations $U(x)$, i.e. this factor reproduces the $\exp(in\theta)$ term in Eq. (4.73). This phase factor can be also obtained if one remembers that usually the θ-dependence of the wave functional is given by the $\exp\left[i\theta S_{CS}(A)\right]$, where $S_{CS}(A)$ is a Chern-Simons term, which under the gauge transformation U transforms as

$$S_{CS}(A^U) = S_{CS}(A) + \frac{1}{24\pi^2}\int dx\epsilon_{ijk}\mathrm{tr}\left[(U^\dagger\partial_i U)(U^\dagger\partial_j U)(U^\dagger\partial_k U)\right]\,, \tag{4.75}$$

so that integrating over U leads precisely to the phase factor Eq. (4.74). The state thus constructed is an eigenstate of an operator of the large gauge transformation with eigenvalue $e^{i\theta}$. This modification results in the addition of the same topological term to the effective action Eq. (4.17). It is amusing to note that for $\theta = \pi$, the "Skyrmions" in the effective theory will be "fermions".

While the extension of the variational calculation to non-vanishing θ-term is quite straightforward, it has not been performed thus far.

4.5. *Confinement?*

The most interesting question is of course whether our variational state is confining. In the toy model in 2+1 dimensions discussed in the previous section, we were able to answer this question by calculating both the expectation value of the spatial Wilson loop, and the potential between static charges. Unfortunately, in the Yang-Mills theory the calculation is much more complicated and the answer is not known. Although there are some arguments that the state is indeed confining (see next section), it has not been proved or disproved by a direct calculation. The calculation of a potential between static charges *á la* [41] has not been attempted. As for the calculation of the Wilson loop, some progress has been made in reducing this calculation to the σ-model level, but no final result has been obtained.

The difficulty in the calculation of the Wilson loop,

$$W(C) = \left\langle \mathrm{tr}\ P \exp\left(i\frac{g}{2} \oint_C dx_i A_i^a \tau^a \right) \right\rangle , \tag{4.76}$$

is to take into account the P-ordering of the exponent. One way of doing so is to introduce new degrees of freedom living on the contour C which, after quantization, become the $SU(N)$ matrices τ^a [56]. We briefly describe the construction in the case of the $SU(2)$ group – the generalization of this construction to an arbitrary Lie group has been discussed in [56].

The construction is based on the observation, made in [57, 58], that instead of considering the ordered product of τ^a matrices one can consider the correlation function

$$\begin{aligned}\left\langle \frac{\tau^a(t_1)}{2} \frac{\tau^b(t_2)}{2} \dots \frac{\tau^c(t_k)}{2} \right\rangle &\longrightarrow \langle n^a(t_1) n^b(t_2) \dots n^c(t_k) \rangle \\ &= \int Dn(t)\, n^a(t_1) n^b(t_2) \dots n^c(t_k) \\ &\times \exp\left[i(S+1/2) \int_\Sigma d^2\xi \epsilon_{\mu\nu} \epsilon^{abc} n^a \partial_\mu n^b \partial_\nu n^c \right] ,\end{aligned} \tag{4.77}$$

where S is the spin of the representation, i.e. for the fundamental representation $S = 1/2$; $n^a(t)$ is a unit vector $n^a n^a = 1$ living on a contour C (t is a coordinate on the contour); and Σ is an arbitrary two-dimensional surface with the boundary $C = \delta\Sigma$. The two-dimensional action,

$$S[n] = \int_\Sigma d^2\xi\, \epsilon_{\mu\nu} \epsilon^{abc} n^a \partial_\mu n^b \partial_\nu n^c , \tag{4.78}$$

depends only on values $n^a(t)$ at the boundary.

It can be shown that the Wilson loop can be rewritten as

$$\begin{aligned}W(C) = \Bigg\langle \int Dn(t) \exp&\left[i \int_\Sigma d^2\xi\, \epsilon_{\mu\nu} \epsilon^{abc} n^a \partial_\mu n^b \partial_\nu n^c \right] \\ &\times \exp\left(ig \oint_C dx_i\, A_i^a(x(t)) n^a(t) \right) \Bigg\rangle .\end{aligned} \tag{4.79}$$

The average over A_i can be performed using Eqs. (4.17, 4.31)

$$\left\langle \exp\left(ig\oint_C dx_i\, A_i^a(x(t))n^a(t)\right)\right\rangle_A = \exp\left(-ig\oint_C dx_i\, a_i^a(x(t))n^a(t)\right)$$
$$\times \exp\left(-\frac{1}{2}\oint_C\oint_C dt_1 dt_2\, \dot{x}_i(t_1)\dot{y}_i(t_2)n^a(t_1)n^b(t_2)(\mathcal{M}^{-1})^{ab}(x,y)\right)\,, \quad (4.80)$$

where a_i^a was defined in Eq. (4.31). The Wilson loop can be calculated as the average over two scalar fields: $U(x)$ living in the whole space and $n^a(\xi)$ living on a two-dimensional surface Σ such that $C = \delta\Sigma$

$$W(C) = \int DU \int Dn\, \exp\left(-\Gamma[U] + iS[n]\right) \exp\left(-ig\oint_C dx_i\, a_i^a(x(t))n^a(t)\right)$$
$$\times \exp\left(-\frac{1}{2}\oint_C\oint_C dt_1 dt_2\, \dot{x}_i(t_1)\dot{y}_i(t_2)n^a(t_1)n^b(t_2)(\mathcal{M}^{-1})^{ab}(x,y)\right)\,. \quad (4.81)$$

In the infrared limit one can use Eq. (4.38) to simplify Eq. (4.81) and get

$$\begin{aligned} W(C) = &\int DU_L \int Dn\, \exp\left(-\Gamma_L[U] + iS[n]\right) \\ &\times \exp\left(-i\frac{g}{2}\oint_C dx_i\, \lambda_i^a{}_L(x(t))n^a(t)\right) \\ &\times \exp\left(-\frac{1}{4}\oint_C\oint_C dt_1 dt_2\, \dot{x}_i(t_1)\dot{y}_i(t_2)n^a(t_1)n^a(t_2)G(x-y)\right) \\ &\times \int DU_H\, \exp\left(-\Gamma_H[U]\right) \exp\left(-i\frac{g}{2}\oint_C dx_i\, \lambda_i^b{}_H(x(t))S_L^{ba}n^a(t)\right)\,. \end{aligned} \quad (4.82)$$

Integrating over U_H one obtains

$$\begin{aligned} W(C) = &\int Dn\, \exp\left(iS[n]\right) \\ &\times \exp\left(-\frac{1}{2}\oint_C\oint_C dt_1 dt_2\, \dot{x}_i(t_1)\dot{y}_i(t_2)n^a(t_1)n^a(t_2)G(x-y)\right) \\ &\times \int DU\, \exp\left(-\Gamma[U]\right) \exp\left[\frac{1}{2}\oint_C dx_i\, \mathrm{tr}\left(\tau^a U^\dagger \partial_i U\right) n^a(t)\right]\,, \end{aligned} \quad (4.83)$$

where the integration DU is over the low momentum modes only and $\Gamma[U]$ is the corresponding low momentum action. Since $G(x-y)$ is short range, the term

$$\exp\left(-\frac{1}{2}\oint_C\oint_C dt_1 dt_2\, \dot{x}_i(t_1)\dot{y}_i(t_2)n^a(t_1)n^a(t_2)G(x-y)\right) \tag{4.84}$$

gives only perimeter dependence and can be neglected when calculating the string tension. Now reverting back from n^a to τ^a we find

$$W(C) = \left\langle \mathrm{tr}\, P\exp\left(\frac{1}{2}\oint_C dl_i\, U^\dagger \partial_i U\right)\right\rangle_U , \tag{4.85}$$

where the averaging is performed with the low momentum σ-model action. This is reminiscent of the average of the monodromy operator

$$M = \mathrm{tr}\, P\exp\left(\oint_C dl_i\, U^\dagger\partial_i U\right) \tag{4.86}$$

and one might expect that the result is similar. Since the target space of the sigma model is $\mathcal{M} = SU(N)/Z_N$, and $\Pi_1(\mathcal{M}) = Z_N$, the monodromy can take on values $\exp(i2\pi n/N)$. It has a natural interpretation in terms of the topological defects in the sigma model. As mentioned above, the topology of the σ-model allows for the existence of Z_N strings (the soliton-instantons discussed in the previous subsection do not play any special role in the monodromy calculation). The string creation operator and the operator M satisfy the commutation relations of the t' Hooft algebra [59]. Therefore, in the presence of a string, the operator M has expectation value $\exp(i2\pi n/N)$, where n is the linking number between the loop C and the string. As we have argued, the sigma model is in the disordered phase and the disordering can be thought of as the condensation of the topological defects. We have thus far discussed skyrmions (instantons) as the relevant defects, but it is quite plausible that the Z_N strings are condensed as well. That would mean that the vacuum of the sigma model has a large number of strings and also that the fluctuations in this number are large. In this situation the VEV of $\mathcal{M}$ must average to zero very quickly, and for large loops will have an area law. We then may expect that the Wilson loop will also have an area law $W(C) \sim \exp(-\alpha' A)$.

While this argument is not implausible, currently we have no quantitative method of estimating Eq. (4.85).

Even though the question about confining properties of the state remains unanswered, the results of the variational calculation so far are quite interesting. It yields the dynamical generation of the scale which is of the right magnitude, a reasonable value of the gluon condensate and a neat relation to the instanton physics. All these results are intrinsically non-perturbative.

Apart from the vacuum structure, there is another mysterious domain of the QCD physics which is not accessible with perturbative tools: the deconfining phase transition. We may hope that the variational method can give us a handle to understanding the deconfinement physics. In the next section we describe its application to the Yang-Mills theory at finite temperature and the study of the deconfinement phase transition.

5. The Yang-Mills theory at finite temperature

Attempts to understand the nature of the deconfining phase transition in QCD date back almost 30 years. Since the pioneering work of Polyakov [60] and Susskind [61], much effort has been made to study the basic physics as well as the quantitative characteristics of the transition. The high temperature phase of QCD is widely believed to resemble an almost free plasma of quarks and gluons. At asymptotically high temperatures this is confirmed by explicit perturbative calculations of the free energy [62]. Perturbation theory in its simplest form, however, is valid only at unrealistically high temperatures. In recent years a different and promising avenue has been explored. This incorporates analytical resummation of the effects of the gluon screening mass into the 3D effective Lagrangian, which is then solved numerically by 3D lattice gauge theory methods [63–66]. The results of this approach seem to be in agreement [67] with direct 4D lattice gauge theory calculations [68, 69] all the way down to $2\mathsf{T_c}$. Although we are quite advanced in the understanding of the high temperature phase, the transition region itself is very poorly understood. This region of temperatures, $\mathsf{T_c} < \mathsf{T} < 2\mathsf{T_c}$, is of course the most interesting one, since it is in this region that the transition between "hadronic" and "partonic" degrees of freedom occurs. Interestingly enough, the numerical results indicate that although asymptotically the free energy does approach that of the free partonic plasma, the deviations from the Stefan-Boltzmann law even at temperatures of order $10\mathsf{T_c}$ are quite sizable, of order of 15%. This is an indication that the interesting physics of the transition region remains important even at these high temperatures. The study of the transition region itself is a complicated and inherently non-perturbative problem.

The purpose of this section is to study the deconfining phase transition in a pure $SU(N)$ Yang-Mills theory using the variational approach described in the previous sections suitably extended to finite temperature [70]. We will minimize the relevant thermodynamic potential at finite temperature, i.e. the Helmholtz free energy, on a set of gauge invariant density matrices.

5.1. *The variational Ansatz for the density matrix*

The equilibrium state of a quantum mechanical system at finite temperature is not a pure state, but is described by a mixed density matrix. Thus in order to extend the variational analysis to finite temperature we have to generalize our ansatz Eq. (4.5) so that it includes mixed states. In scalar theories the Gaussian approximation has a long history of applications at finite temperature [71, 72]. We generalize our Ansatz along the same lines.

We start by considering the density matrices which in the field basis have Gaussian matrix elements [70]

$$\tilde{\varrho}[A, A'] = \exp\Big\{ -\frac{1}{2}\int_{x,y} A_i^a(x) G_{ij}^{-1ab}(x,y) A_j^b(y) \\ + A_i'^a(x) G_{ij}^{-1ab}(x,y) A_j'^b(y) - 2A_i^a(x) H_{ij}^{ab}(x,y) A_j'^b(y) \Big\}. \quad (5.1)$$

As before, we take the variational functions diagonal in both color and Lorentz indices, and translationally invariant,

$$\begin{aligned} G_{ij}^{-1ab}(x,y) &= \delta^{ab}\delta_{ij} G^{-1}(x-y)\,, \\ H_{ij}^{ab}(x,y) &= \delta^{ab}\delta_{ij} H(x-y)\,. \end{aligned} \quad (5.2)$$

Then we find

$$\tilde{\varrho}[A, A'] = \exp\Big\{ -\frac{1}{2}\int_{x,y} \big(AG^{-1}A + A'G^{-1}A' - 2AHA' \big) \Big\}. \quad (5.3)$$

For $H = 0$ this density matrix represents a pure state, since it can be written in the form

$$\tilde{\varrho} = |\Psi[A]\,\rangle\langle\,\Psi[A]|\,, \quad (5.4)$$

with $\Psi[A]$ a Gaussian wave function, Eq. (4.4). At non-zero H the density matrix is, however, mixed. The magnitude of H, therefore, determines the entropy of this trial density matrix.

We now make an additional simplification in our ansatz. First, we restrict the functions $G^{-1}(x)$ to the same functional form as at zero temperature Eq. (4.13), i.e.

$$G^{-1}(k) = \begin{cases} \sqrt{k^2} & \text{if } k^2 > M^2, \\ M & \text{if } k^2 < M^2. \end{cases} \quad (5.5)$$

Further, we will take $H(k)$ to be small and non-vanishing only at low momenta,

$$H(k) = \begin{cases} 0 & \text{if } k^2 > M^2, \\ H \ll M & \text{if } k^2 < M^2. \end{cases} \tag{5.6}$$

The logic behind this choice of Ansatz is the following. At finite temperature we expect $H(k)$ to be roughly proportional to the Bolzmann factor $\exp\{-\mathsf{E}(k)\beta\}$. In our ansatz, the role of one particle energy is played by the variational function $G^{-1}(k)$. We will be interested only in temperatures close to the phase transition, and those we anticipate to be small, $\mathsf{T_c} \leq M$. For those temperatures, one particle modes with momenta $k \geq M$ are not populated, and we thus can put $H(k) = 0$. For $k \leq M$ the Bolzmann factor is non-vanishing, but small. Further, it depends only very weakly on the value of the momentum. We will have, of course, to verify a posteriori that our assumptions about the smallness of $\mathsf{T_c}$ and H are justified.

As before, we explicitly impose gauge invariance by projecting ϱ onto the gauge invariant sector

$$\varrho[A, A'] = \int DU'DU'' \exp\Bigg\{ -\frac{1}{2}\int_{x,y} A^{U'}G^{-1}A^{U'} \\ + A'^{U''}G^{-1}A'^{U''} - 2A^{U'}HA'^{U''} \Bigg\}, \tag{5.7}$$

where A^U is given by Eqs. (4.6, 4.7). One of the group integrations in Eq. (5.7) is redundant, since we will only calculate the quantities of the form $\mathrm{Tr}\,\varrho O$, with O being gauge invariant. Our Ansatz for the density matrix then is

$$\varrho[A, A'] = \int DU \exp\left\{ -\frac{1}{2}\int_x AG^{-1}A + A'^U G^{-1}A'^U - 2AHA'^U \right\}. \tag{5.8}$$

This expression is not explicitly normalized to unity. Nevertheless, we find it convenient to refer to it as density matrix while explicitly inserting a normalization factor whenever necessary. Thus the average of a gauge invariant operator $\mathcal{O}$ is given by

$$\begin{aligned} \langle \mathcal{O} \rangle_{A,U} &= Z^{-1}\mathrm{Tr}(\varrho\mathcal{O}) \\ &= Z^{-1}\int DUDA\, \mathcal{O}(A, A') \\ &\quad \cdot \exp\left\{ -\frac{1}{2}\int_x AG^{-1}A + A'^U G^{-1}A'^U - 2AHA'^U \right\}\Bigg|_{A'=A}, \end{aligned} \tag{5.9}$$

where Z is the normalization of the trial density matrix ϱ, i.e.

$$\begin{aligned} Z = \mathrm{Tr}\varrho &= \int DUDA \exp\Big\{ -\frac{1}{2}\int_x AG^{-1}A + A^U G^{-1} A^U - 2AHA^U \Big\} \\ &= \int DUD\tilde{A} \exp\Big\{ -\frac{1}{2}\int_x \tilde{A}\Delta\tilde{A} + \lambda\big(G^{-1} - \omega\Delta^{-1}\omega^T\big)\lambda \Big\}, \end{aligned} \quad (5.10)$$

with

$$\tilde{A} = A + \lambda\omega\Delta^{-1}, \quad (5.11)$$

$$\Delta = 2G^{-1}\Big(1 - \frac{HG}{2}(S + S^T)\Big), \quad (5.12)$$

$$\omega = (G^{-1}S - H). \quad (5.13)$$

The $\tilde{A}$ integration can be performed to yield

$$\mathrm{Tr}\varrho = \int DU \exp\Big\{ -\frac{1}{2}\lambda\big(G^{-1} - \omega\Delta^{-1}\omega^T\big)\lambda - \frac{3}{2}\mathrm{Tr}\ln\frac{\Delta}{2} \Big\}. \quad (5.14)$$

We now adopt the same strategy for treating the high momentum modes of U as at $\mathsf{T} = 0$. Namely, they are integrated perturbatively to one loop accuracy. The result is the effective σ-model for the matrices U with momenta below M. The coupling constant g of this σ-model gets renormalized as before according to the one loop Yang-Mills β-function, and thus has to be understood as $g(M)$. Additionally, due to independence of H on momentum, for low momentum modes of U the function $H(x-y)$ is equivalent to $H\delta^3(x-y)$.

The final approximation has to do with the fact that H is assumed to be small. For arbitrarily large H the variational calculation is forbiddingly complicated even with all the above mentioned simplifications. This is because the gauge projection renders the calculation of entropy in the general case unfeasible. However, at small H we only need to calculate the leading term in the entropy. This calculation can indeed be done, and is described below. Since we are only calculating the leading order contribution in H, we only have to consider corrections to the σ-model action of first order in H. With this in mind, the normalization factor becomes

$$\mathrm{Tr}\varrho = \int DU \exp\Big\{ -\frac{1}{2}\lambda\Big(\frac{G^{-1}}{2} + \frac{H}{4}\big(S + S^T\big)\Big)\lambda + \frac{3}{4}HG\,\mathrm{tr}\big(S + S^T\big) \Big\}. \quad (5.15)$$

5.2. *The effective σ-model*

Just as at zero temperature, the normalization Z can be interpreted as the generating functional for a theory defined by the action $\mathcal{S}(U)$,

$$Z = \mathrm{Tr}\varrho = \int DU e^{-\mathcal{S}(U)}\,, \tag{5.16}$$

where

$$\mathcal{S}(U) = \frac{M}{4}\,\lambda\lambda + \frac{1}{8}\lambda\, H(S+S^T)\,\lambda - \frac{1}{4\pi^2}\,HM^2 \mathrm{tr}\,S\,. \tag{5.17}$$

We simplify this expression using

$$\lambda\lambda = \frac{2}{g^2}\,\mathrm{tr}(\partial U\partial U^\dagger)\,, \tag{5.18}$$

$$\lambda S^T\lambda = \lambda S\lambda = -\frac{1}{2g^2}\,\mathrm{tr}\Big[(U^\dagger\partial U - \partial U^\dagger U)(\partial U U^\dagger - U\partial U^\dagger)\Big]\,, \tag{5.19}$$

$$\mathrm{tr}\,S = \mathrm{tr}\,S^T = \mathrm{tr}\,U^\dagger \mathrm{tr}\,U - 1\,. \tag{5.20}$$

Inserting these into the action we get

$$\mathcal{S}(U) = \frac{M}{2g^2}\,\mathrm{tr}\,(\partial U\partial U^\dagger) - \frac{H}{8g^2}\,\mathrm{tr}\Big[(U^\dagger\partial U - \partial U^\dagger U)(\partial U U^\dagger - U\partial U^\dagger)\Big] - \frac{1}{4\pi^2}\,HM^2\,\mathrm{tr}\,U^\dagger\,\mathrm{tr}\,U\,, \tag{5.21}$$

where U-independent terms have been dropped.

At this point it is useful to relate our effective σ-model with a standard tool used in finite temperature calculations, namely the effective action for the Polyakov loop. The matrix U plays a similar role to the Polyakov loop P at finite temperature – the functional integration over U projects out the physical subspace of the large Hilbert space on which the Hamiltonian of gluodynamics is defined. The effective σ-model Eq. (5.21) therefore is a close analogue of the effective theory for the low momentum modes of the Polyakov loop variable. Its status and applicability region are however different from the usual perturbative effective actions, see e.g. [73]. The standard effective action is calculated in perturbation theory and is valid at high temperature. Our effective action Eq. (5.21) depends on the variational parameters M and H, and in a sense is a variational effective action. Also due to our restrictions to small values of H, a priori we do not expect it to be valid at high temperatures but, rather, it should represent correctly the physics in the phase transition region.

Another important difference is that our effective σ-model does not have the local gauge invariance $U(x) \to V^\dagger(x)U(x)V(x)$ which is usually associated with the effective action for the Polyakov loop. The reason for this is that our setup is different from that of the standard finite temperature calculation. The way this gauge invariance usually appears is the following. Consider the calculation of any gauge invariant observable in the equilibrium density matrix at finite temperature

$$\langle \mathcal{O} \rangle = \int DU \,\mathrm{Tr}\left[\exp\{-\beta \mathsf{H}\}\mathcal{O}g(U)\right], \tag{5.22}$$

where $g(U)$ is the second quantized operator of the gauge transformation represented by the matrix U. This expression for fixed U can be compared to the same expression but with U gauge transformed,

$$\begin{aligned}\mathrm{Tr}\left[\exp\{-\beta \mathsf{H}\}\mathcal{O}g(V^\dagger U V)\right] &= \mathrm{Tr}\left[\exp\{-\beta \mathsf{H}\}\mathcal{O}g(V^\dagger)g(U)g(V)\right] \\ &= \mathrm{Tr}\left[\exp\{-\beta \mathsf{H}\}\mathcal{O}g(U)\right]. \end{aligned}\tag{5.23}$$

The last equality here follows from the fact that both $\mathcal{O}$ and $\exp\{\beta \mathsf{H}\}$ are gauge invariant, and thus the operator $g(V^\dagger)$ can be commuted all the way to the left. The only effect of the transformation is then to change the basis over which the trace is being taken, which obviously leaves the trace invariant.

Our variational setup is somewhat different. Expectation values are calculated as

$$\int DU \mathrm{Tr}\left[\tilde{\varrho} g(U)\mathcal{O}\right] \tag{5.24}$$

with $\tilde{\varrho}$ defined in Eq. (5.3). This expression is altogether gauge invariant, since the integral over U correctly projects only the contribution of gauge singlet states. However the operator $\tilde{\varrho}$ is not itself explicitly gauge invariant. For that reason the gauge transformation operator $g(V^\dagger)$ cannot be commuted through it, and thus

$$\mathrm{Tr}\left[\tilde{\varrho}\mathcal{O}g(V^\dagger U V)\right] \neq \mathrm{Tr}\left[\tilde{\varrho}\mathcal{O}g(U)\right] \tag{5.25}$$

even for gauge invariant operators $\mathcal{O}$. This manifests itself as absence of local gauge invariance in the action of the effective σ-model, Eq. (5.21).

Nevertheless, we stress again that since the integration over the $SU(N)$ valued field U projects out the physical Hilbert space, its meaning in this sense is the same as that of the Polyakov loop.

5.3. *The calculation of the free energy*

To find the best variational density matrix we have to minimize the free energy with respect to the variational parameters M and H. The Helmholtz free energy F of the density matrix ϱ is given by

$$\mathsf{F} = \langle \mathsf{H} \rangle - \mathsf{T}\mathsf{S}\,, \tag{5.26}$$

where H is the standard Yang-Mills Hamiltonian Eq. (4.1), S is the entropy, and T is the temperature.

Thus

$$\mathsf{F} = \frac{1}{2}\Big(\mathrm{Tr}(E^2\varrho) + \mathrm{Tr}(B^2\varrho)\Big) + \mathsf{T}\cdot \mathrm{Tr}(\varrho \ln \varrho)\,. \tag{5.27}$$

First of all we need to perform the integration over the gauge fields, and reduce this expression to the average of a U-dependent operator in the effective σ-model. In fact, as we shall see soon, to leading order in H the only non-trivial calculation we need to perform is that of the entropy.

We will calculate the entropy up to the first non-trivial order in H. As we now show, the leading term at small H is $O(H \ln H)$. Let us denote by ϱ_0 the density matrix of the pure state with $H = 0$,

$$\varrho_0 = |0\rangle\langle 0|\,. \tag{5.28}$$

Here $|0\rangle$ does not denote necessarily the actual ground state, but rather a projected Gaussian state with arbitrary M. Now, since the matrix elements of the density matrix can be expanded in powers of H, to leading order we can write

$$\varrho = \varrho_0 + \delta\varrho\,, \tag{5.29}$$

where $\delta\varrho$ is $O(H)$.

Imagine that we have diagonalized ϱ. It will have one large eigenvalue $\alpha_0 = 1 - O(H)$, which corresponds to the eigenstate

$$|0'\rangle = |0\rangle + O(H)\,. \tag{5.30}$$

All the rest of the eigenvalues α_i are at most $O(H)$. Then the entropy can be written as

$$\mathsf{S} = -\mathrm{Tr}(\varrho \ln \varrho) = -\alpha_0 \ln \alpha_0 - \sum_{i=1}^{\infty} \alpha_i \ln \alpha_i\,. \tag{5.31}$$

The second term is $O(H \ln H)$, and it is the coefficient of this term that we will now calculate. Neglecting $O(H)$ corrections, we can substitute $\alpha_i =$

H/M under the logarithm. Thus to leading logarithmic order

$$\mathsf{S} = -\sum_i \alpha_i \ln H/M\,. \tag{5.32}$$

Thus we have to calculate $\sum_i \alpha_i$. Let

$$|0'\rangle = |0\rangle + H|x\rangle\,. \tag{5.33}$$

Then we have

$$\varrho = \alpha_0 |0'\rangle\langle 0'| + \sum_{i=1}^{\infty} \alpha_i |x_i\rangle\langle x_i| \tag{5.34}$$

with

$$\langle x_i|0'\rangle = 0\,. \tag{5.35}$$

Note that $\langle 0|x\rangle \neq 0$, but

$$\langle 0|x\rangle + \langle x|0\rangle = 0\,, \tag{5.36}$$

since $|0'\rangle$ has to be normalized at $O(H)$. Also

$$\langle x_i|0\rangle + H\langle x_i|x\rangle = 0\,. \tag{5.37}$$

Thus the overlap $\langle x_i|0\rangle$ is $O(H)$, and we have

$$\varrho = \alpha_0|0\rangle\langle 0| + H\Big(|0\rangle\langle x| + |x\rangle\langle 0|\Big) + \alpha_i |x_i\rangle\langle x_i|\,. \tag{5.38}$$

Multiplying this by ϱ_0 we get

$$\begin{aligned} \varrho_0 \cdot \varrho &= \alpha_0 \cdot \varrho_0 + H|0\rangle\langle x| + H\langle 0|x\rangle|0\rangle\langle 0|\,, \\ \varrho \cdot \varrho_0 &= \alpha_0 \cdot \varrho_0 + H|x\rangle\langle 0| + H\langle x|0\rangle|0\rangle\langle 0|\,, \end{aligned} \tag{5.39}$$

and thus,

$$\varrho_0 \cdot \varrho + \varrho \cdot \varrho_0 - \varrho = \alpha_0 \cdot \varrho_0 - \alpha_i |x_i\rangle\langle x_i|\,. \tag{5.40}$$

Multiplying again by ϱ_0, we get rid of $|x_i\rangle\langle x_i|$ to $O(H)$

$$\alpha_0 \cdot \varrho_0 = \varrho_0 \cdot \varrho + \varrho_0 \cdot \varrho \cdot \varrho_0 - \varrho_0 \cdot \varrho = \varrho_0 \cdot \varrho \cdot \varrho_0\,. \tag{5.41}$$

Then we have

$$\alpha_0 = \mathrm{Tr}(\varrho_0 \cdot \varrho)\,. \tag{5.42}$$

Since $\mathrm{Tr}\,\varrho = 1$ we have

$$\sum_i \alpha_i = 1 - \alpha_0 = \mathrm{Tr}\left(\varrho_0(1-\varrho)\right) \tag{5.43}$$

which, inserted into Eq. (5.31), gives

$$\mathsf{S} = -(1 - \mathrm{Tr}\,(\varrho_0 \cdot \varrho)) \ln H/M\,. \tag{5.44}$$

The derivation has been given for the normalized density matrices ϱ_0 and ϱ. In terms of our Gaussian matrices we should restore the normalization factors Z and Z_0, so that finally we have

$$\mathsf{S} = \Big(\frac{\mathrm{Tr}\,(\varrho_0 \cdot \varrho)}{\mathrm{Tr}\,\varrho_0 \cdot \mathrm{Tr}\,\varrho} - 1\Big) \ln H/M\,. \tag{5.45}$$

It is easy to check that to $O(H)$

$$\mathrm{Tr}\,(\varrho_0 \cdot \varrho) = (\mathrm{Tr}\,\varrho_0)^2\,, \tag{5.46}$$

and

$$\mathsf{S} = \Big(\frac{\mathrm{Tr}\,\varrho_0}{\mathrm{Tr}\,\varrho} - 1\Big) \ln H/M\,. \tag{5.47}$$

From Eq. (5.15), it is clear that

$$\mathrm{Tr}\,\varrho = \Big[1 + H\Big(\frac{1}{4\pi^2}\,M^2\,\mathrm{tr}S\, - \frac{1}{4}\lambda S \lambda\Big)\Big] \cdot \mathrm{Tr}\,\varrho_0\,. \tag{5.48}$$

Using Eqs. (5.19, 5.20) we finally get

$$\mathsf{S} = -\Big[\Big\langle \frac{1}{8g^2}\,\mathrm{tr}\,(U^\dagger \partial U - \partial U^\dagger U)(\partial U U^\dagger - U \partial U^\dagger) + \frac{1}{4\pi^2}\,M^2(\mathrm{tr}\,U^\dagger \mathrm{tr}\,U - 1)\Big\rangle_U\Big] H \ln H/M\,. \tag{5.49}$$

To leading order in H, the averaging over U in this expression has to be performed with the σ-model action with $H = 0$.

The expression of Eq. (5.49) has the following striking property. For $M < M_c$ it vanishes identically. The reason is very simple. The first term in Eq. (5.49) is the product of the left handed $SU(N)$ current and the right handed $SU(N)$ current in the σ-model. Thus it transforms as an adjoint representation under each one of the $SU(N)$ factors of the $SU_L(N) \otimes SU_R(N)$ transformation. The same is also true for the second term in Eq. (5.49). The σ-model action at $H = 0$ is itself obviously invariant under the whole $SU_L(N) \otimes SU_R(N)$ group. Now, at $M < M_c$, the symmetry group is not spontaneously broken, and thus any operator which is not a scalar has a vanishing expectation value. It follows immediately that the entropy has an $O(H \ln H/M)$ contribution only for $M > M_c$, when the $SU_L(N) \otimes SU_R(N)$ group is spontaneously broken down to $SU_V(N)$.

This observation makes our task considerably simpler. Since for $M < M_c$ the entropy is zero, we do not have to consider at all the disordered phase of the effective σ-model. In this disordered phase the free energy coincides with energy, and thus the calculation is identical to the calculation at zero temperature presented in Sec. (4).

Thus we only need to consider the effective σ-model in the ordered phase. As at $\mathsf{T} = 0$ we perform the calculations in the ordered phase to leading order in α_s. Since there are no $O(H \ln H/M)$ corrections to energy at this order, the result for the energy in the disordered phase is identical to the result at zero temperature, Eq. (4.52). Thus our expression for the free energy in the ordered phase of the σ-model is

$$\mathsf{F} = \frac{N^2}{120\pi^2} M^4 + \mathsf{T}\left(\Big\langle \frac{1}{8g^2} \operatorname{tr}(U^\dagger \partial U - \partial U^\dagger U)(\partial U U^\dagger - U \partial U^\dagger) + \frac{1}{4\pi^2} M^2 (\operatorname{tr} U^\dagger \operatorname{tr} U - 1)\Big\rangle_U\right) H \ln H/M \,. \quad (5.50)$$

We now average over U in the leading order perturbation theory.

5.4. *The σ-model perturbation theory; minimization of the free energy and the Debye mass*

For the purpose of the perturbative σ-model calculation we parameterize the U matrices as

$$U = \exp\left\{\frac{i}{2}\, g\phi^a \tau^a\right\}. \quad (5.51)$$

Although we only need the leading order, it is instructive to check that the order g^2 term in the expansion is indeed small. To this order we have

$$U \simeq \left(1 + \frac{i}{2} g\phi^a\tau^a - \frac{1}{8} g^2 \phi^a\phi^b\tau^a\tau^b - \frac{i}{48} g^3 \phi^a\phi^b\phi^c\tau^a\tau^b\tau^c\right). \quad (5.52)$$

So that the σ-model action becomes

$$\mathcal{S} = \frac{M}{2g^2} \operatorname{tr}(\partial U \partial U^\dagger) = \frac{M}{4}\, \partial\phi\partial\phi + \frac{M}{192}\, g^2 (\partial\phi^a)(\partial\phi^c)\phi^b\phi^d \operatorname{tr}\left[\tau^a\tau^b\tau^c\tau^d - \tau^a\tau^c\tau^b\tau^d\right]. \quad (5.53)$$

The propagator of the phase field ϕ is thus

$$\langle \phi^a \phi^b \rangle = \frac{2}{Mk^2}\, \delta^{ab}\,. \quad (5.54)$$

To get the idea of the quality of this perturbative expansion we can calculate for example $\langle \mathcal{S} \rangle$. In this calculation one has to take into account the fact that the measure in the path integral over the phase ϕ^a is not the simple $D\phi$, but rather the group invariant $U(N)$ measure μ. To first order in g^2 it is

$$\mu = D\phi^a \exp\left\{\frac{M^3 N}{144\pi^2}\, g^2 \int d^3x \phi^2(x)\right\}. \tag{5.55}$$

Taking this into account we find that $\langle \mathcal{S} \rangle$ gets no correction of order g^2. We thus feel confident that the use of the perturbation theory in the ordered phase of the σ-model is an admissible approximation. In the following we will only keep leading order expressions.

Calculating to leading order the entropy Eq. (5.49) and keeping only the $O(N^2)$ terms we find

$$\langle \mathsf{S} \rangle = -\frac{N^2}{6\pi^2}\, M^2 H \ln \frac{H}{M}\,. \tag{5.56}$$

Introducing the dimensionless quantity

$$h = \frac{H}{M}\,, \tag{5.57}$$

we can write the expression for the free energy as

$$\begin{aligned} \mathsf{F} &= \langle \mathsf{H} \rangle - \mathsf{T}\langle \mathsf{S} \rangle \\ &= \frac{N^2}{120\pi^2}\, M^4 + \mathsf{T}\, \frac{N^2}{6\pi^2}\, M^3 h \ln h\,. \end{aligned} \tag{5.58}$$

We now have to minimize this expression with respect to h and M. It is convenient to first perform the minimization with respect to h at fixed M. This obviously gives

$$\frac{\partial \mathsf{F}}{\partial h} = 0 \to h = \frac{1}{e}\;. \tag{5.59}$$

Thus as a function of M only, the free energy becomes

$$\mathsf{F} = \frac{N^2}{120\pi^2}\, M^4 - \frac{\mathsf{T}}{e}\, \frac{N^2}{6\pi^2}\, M^3\,. \tag{5.60}$$

Now minimizing with respect to M we find

$$\frac{\partial \mathsf{F}}{\partial M} = 0 \to M = \frac{15\mathsf{T}}{e}\;. \tag{5.61}$$

Thus for $M \geq M_c$ the free energy of the best variational density matrix as a function of temperature is

$$\mathsf{F}_{M\geq M_c} = -\frac{N^2}{360\pi^2}\left(\frac{15\mathsf{T}}{e}\right)^4 . \tag{5.62}$$

We now have to compare this value with the free energy for $M \leq M_c$. As we have discussed above, this is given by the expectation value of the Hamiltonian alone, and is minimized at $M = M_c$. Its value is

$$\mathsf{F}_{M\leq M_c} = -\frac{N^2}{30\pi^2}\, M_c^4 . \tag{5.63}$$

Comparing the two expressions we find

$$\mathsf{T_c} = \frac{12^{\frac{1}{4}}e}{15}\, M_c . \tag{5.64}$$

Using the value of M_c from Eq. (4.63) we have

$$\mathsf{T_c} = 450\,\mathrm{MeV} . \tag{5.65}$$

For $\mathsf{T} \leq \mathsf{T_c}$ the free energy is minimized in the variational state with $M = M_c$. In our approximation this state is the same as at zero temperature. Its entropy vanishes, and the effective σ-model is in the disordered phase. The Polyakov loop vanishes, $\langle U\rangle = 0$ and according to the standard wisdom this is a confining state.

For $\mathsf{T} \geq \mathsf{T_c}$ the best variational state is very different. The entropy of this state is non-zero,

$$\mathsf{S} = \frac{N^2}{6\pi^2 e}\left(\frac{15\mathsf{T}}{e}\right)^3 . \tag{5.66}$$

The Polyakov loop is non-zero $\langle U\rangle \neq 0$ and thus the high temperature density matrix describes a deconfined phase.

Finally, we note that in the deconfined phase our best variational density matrix has a non-vanishing "electric screening" or "Debye" mass. The Debye mass is conveniently defined as the "mass" of the phase of the Polyakov loop. This mass is non-vanishing in our calculation for the following reason. As long as $H = 0$, the effective σ-model action has a global $SU_L(N)\otimes SU_R(N)$ symmetry. Thus in the ordered phase of the σ model the phases ϕ^a are massless. However, as discussed above, the terms of order H in Eq. (5.21) break this symmetry explicitly down to the diagonal $SU_V(N)$. As a result the would be "Goldstone" phases ϕ^a acquire mass. To calculate this mass it is convenient first to note that to $O(g^2)$

$$\begin{aligned}
&\operatorname{tr}(U^\dagger\partial U - \partial U^\dagger U)(\partial U U^\dagger - U\partial U^\dagger) \\
&\quad = -4\operatorname{tr}(\partial U^\dagger \partial U) - \frac{g^4}{4}\phi^a\phi^c\partial\phi^b\partial\phi^d \operatorname{tr}\left(\tau^a\tau^b\tau^c\tau^d - \tau^a\tau^c\tau^b\tau^d\right). \quad (5.67)
\end{aligned}$$

The contribution of the $SU_L(N)\otimes SU_R(N)$ term to the mass cancels against the contribution of the measure Eq. (5.55). Using Eqs. (5.21, 5.67) we then find to $O(g^2)$, and to leading order in H,

$$M_D^2 = \frac{4}{3\pi}\,\alpha_s(M)NMH\,. \quad (5.68)$$

As a function of temperature we have

$$M_D^2 = \alpha_s\left(\frac{15}{e}T\right)N\,\frac{300}{\pi e^3}\,\mathsf{T}^2\,. \quad (5.69)$$

Let us summarize the results of our analysis of the deconfinement transition. We find the phase transition at a temperature of about $\mathsf{T_c} \simeq 450\,\mathrm{MeV}$. The transition is strongly first order at large N. The latent heat is $\Delta E = \frac{N^2}{90\pi^2}\left(\frac{15\mathsf{T}}{e}\right)^4$. Below the transition the entropy is zero, the best variational state is the same as at zero temperature, and the average value of the Polyakov loop is zero. Above the transition, the entropy is non-zero and proportional to the number of "colored" degrees of freedom, $\mathsf{S} \propto N^2$. The average value of the Polyakov loop is non-zero and the phase is deconfined.

It is quite interesting that at high temperature our formulae numerically are quite close to the predictions of free gluon plasma. In particular, our value for the free energy, Eq. (5.62), should be compared to the free gluon plasma expression

$$\mathsf{F}_{free} = -\frac{N^2\pi^2}{45}\,\mathsf{T}^4\,. \quad (5.70)$$

The ratio between the two is

$$\frac{\mathsf{F}_{free}}{\mathsf{F}_{var}} \simeq 0.85\,. \quad (5.71)$$

The ratio of the entropies is the same. Interestingly we get the same ratio comparing our value for the Debye mass Eq. (5.69) with the leading order perturbative one, $M_{pert}^2 = \frac{4\pi}{3}\alpha_s NT^2$,

$$\frac{M_{pert}^2}{M_D^2} \simeq 0.85\,. \quad (5.72)$$

The pressure approaches its asymptotic value according to the simple formula

$$\frac{\mathsf{P}(\mathsf{T})}{\mathsf{P}_{\text{asympt}}} = 1 - \frac{\mathsf{T_c}^4}{\mathsf{T}^4}\,. \tag{5.73}$$

Here the asymptotic value of the pressure $\mathsf{P}_{\text{asympt}}$ is given by Eq. (5.62). The pressure $\mathsf{P}(\mathsf{T})$ is given by the difference between Eq. (5.62) and the value of the free energy at zero temperature, which coincides with expression Eq. (5.63).

One has to take the comparison Eqs. (5.71, 5.72) with a grain of salt. As explained above, our calculations were performed assuming small H. A priori we expect that this restriction should confine us to not too large temperatures. On the other hand the minimization of the free energy resulted in the value $H/M = 1/e$ independently of temperature. Thus, we feel that the comparison Eq. (5.71) may be meaningful.

The main features of these results are indeed what we expect from the deconfinement phase transition on general grounds. It is nice that a simple minded calculation such as this does qualitatively so well in such a complicated problem. It therefore appears that the projection of the trial density matrix on the gauge invariant Hilbert space is, just as at zero temperature, the crucial feature that dictates most if not all the important aspects of the low energy and low temperature physics. In the context of the present calculation the most important effect of the gauge projection is obviously vanishing of the entropy in the low temperature phase. We stress that this feature was not at all built into our initial ansatz, but followed naturally and unavoidably in the disordered phase of the effective σ-model.

Quantitatively, this calculation of course should be taken for what it is – an approximate implementation of the variational principle. As with any variational calculation, the range of validity of this calculation is not sharply defined. Even within the variational framework we had to resort to additional approximations. The most severe simplifications that we had to impose are the perturbation theory in the ordered phase of the σ-model and the assumption of smallness of H. The projection over the gauge group, which as we saw is so physically important, is what makes the calculational task difficult and forces us to make these approximations.

The assumptions of smallness of g and of smallness of H affect different aspects of our result. In particular, in the leading order of perturbation theory the expectation value of the Polyakov loop U is equal to unity. The actual value of U on the ordered side of the transition according to [74] is close to one half. Thus our perturbative calculation is rather more reliable somewhat further away from the transition. The closer to the transition we get, the more important higher order corrections in g^2 become. Thus to properly describe the transition region of QCD we need to improve our calculational method in the vicinity of the transition in the σ-model. In line with this we expect that the estimate for the critical temperature we obtained here is somewhat higher than we would get, had we treated the σ-model more accurately in the transition region. This is consistent with the fact that our result for $\mathsf{T_c}$ is by about 50% higher than the lattice value of 270 MeV.

The smallness of h is quite important in a different way. The value of $h = 1/e$ that we obtain is in fact a reasonably small number, so omitting the corrections in powers of h is fairly safe. On the other hand, the terms linear in h but not enhanced by $\ln h$, which we have ignored in the present calculation, have to be accounted for more carefully. With the value of h that we obtain, these terms are not suppressed in any obvious way.

The obvious stumbling block to any improvement along these lines is the calculation of the entropy $\mathsf{S} = -\mathrm{tr}\varrho \ln \varrho$. However, if one opts for restricting, as before, the analysis to leading order in g^2 the entropy and, therefore, the free energy can be, without any additional approximations, calculated to all orders in h [75]. This improved analysis is carried out below.

5.5. *All-order in h analysis*

Let us ask ourselves what would happen if we did not restrict H to be small, and more generally did not restrict the functional forms of $G(k)$ and $H(k)$ in our variational ansatz. We could still carry on our calculation for a while. Namely we would be able to integrate over the vector potentials in all averages, and would reduce the calculation to a consideration of some non-linear σ-model of the U-field. This σ-model quite generally will have a symmetry breaking phase transition as the variational functions $G(k)$ and $H(k)$ are varied. Since at this transition the Polyakov loop U changes its behavior, the disordered phase of the σ-model corresponds to the confining phase of the Yang-Mills theory, while the ordered phase of the σ-model represents the deconfined phase. Thus, in order to study deconfinement in the $SU(N)$ Yang-Mills theory, we should analyze the physics of each σ-model phase as accurately as possible and calculate the transition scale M_c (or rather $G_c(k)$). We then calculate the free energy of the σ-model in each phase at temperature T and extract the minimal free energy. The deconfinement transition occurs at the temperature for which the free energies calculated in the ordered and disordered phases of the sigma model coincide.

In practice in the disordered phase no progress seems possible without restricting the arbitrary kernels and we adopt the forms Eqs. (5.5, 5.6). The resulting minimal free energy is thus independent of the temperature and is given by Eq. (5.63).

On the other hand, in the ordered phase we can relax the restriction on H and G if, as before, we work in the leading order in perturbation theory. In this case minimization with respect to arbitrary kernels $G^{-1}(k)$ and $H(k)$ is possible. We now describe this calculation following [75].

In this approximation, for the U matrices we use the parameterization Eq. (5.51). Hence at leading order one can take

$$\begin{aligned} U &\simeq 1\,, \\ \partial_i U &\simeq ig\partial_i\phi^a\,\frac{\tau^a}{2}\;. \end{aligned} \tag{5.74}$$

Thus, the gauge transformations Eq. (4.6) reduce to

$$A_i^a \rightarrow A_i^a - \partial_i\phi^a \tag{5.75}$$

and the Hamiltonian Eq. (4.1) reduces to

$$\mathsf{H} = \frac{1}{2}\left[E_i^{a2} + (\epsilon_{ijk}\partial_j A_k^a)^2\right]\;. \tag{5.76}$$

These last two equations describe the theory $U(1)^{N^2-1}$; in the leading order of the σ-model perturbation theory, the $SU(N)$ Yang–Mills theory reduces to the $U(1)^{N^2-1}$ free theory. The density matrix Eq. (5.8) becomes Gaussian again, because the gauge transformations are linear. One has

$$\begin{aligned} \varrho[A,A'] = \int D\phi\,\exp\Big\{-\frac{1}{2}\Big[&AG^{-1}A + (A'-\partial\phi)G^{-1}(A'-\partial\phi) \\ &- 2AH(A'-\partial\phi)\Big]\Big\}\,. \end{aligned} \tag{5.77}$$

Now the theory of N^2-1 $U(1)$ free fields in 3+1 dimensions is completely tractable. Both the energy and the entropy can be calculated explicitly [76]. The free energy in terms of the arbitrary kernels G^{-1} and H is

$$\begin{aligned} \mathsf{F} = \frac{N^2-1}{2}\int\frac{d^3p}{(2\pi)^3}\Bigg[&G^{-1}(1+GH) + p^2G(1-GH)^{-1} \\ &- 4\mathsf{T}\left(\ln\left[\frac{GH}{\xi}\right] - \ln\left[\frac{\eta}{GH}\right]\cdot\frac{\eta}{\xi}\right)\Bigg]\,, \end{aligned} \tag{5.78}$$

where $\eta = 1 - (1 - (GH)^2)^{1/2}$ and $\xi = (1 - (GH)^2)^{1/2} - (1 - GH)$.

It is minimized by

$$G^{-1} = p\left(\frac{1 + e^{-\frac{2p}{\mathsf{T}}}}{1 - e^{-\frac{2p}{\mathsf{T}}}}\right),$$
$$H = 2p\left(\frac{e^{-\frac{p}{\mathsf{T}}}}{1 - e^{-\frac{2p}{\mathsf{T}}}}\right), \tag{5.79}$$

and the minimal value of the free energy at temperature T is

$$\begin{aligned}\mathsf{F} &= \frac{N^2-1}{\pi^2}\int_0^\infty p^2 dp\ \left[\frac{p}{2} + \mathsf{T}\ln(1 - e^{-p/\mathsf{T}})\right] \\ &= -\frac{(N^2-1)\mathsf{T}^4}{3\pi^2}\int_0^\infty dx \frac{x^3}{e^x - 1} \\ &= -\frac{\pi^2(N^2-1)\mathsf{T}^4}{45},\end{aligned} \tag{5.80}$$

where, just like in Eq. (5.63), zero-point term has been discarded. This is, of course, just the free energy of a free photon gas.

Thus the free energy of $SU(N)$ Yang-Mills theory is minimized with $M = M_c$ in the disordered phase of the σ- model for temperatures below T_c, which is obtained by equating the free energies Eqs. (5.63) and (5.80),

$$-\frac{N^2 M_c^4}{30\pi^2} = -\frac{\pi^2 N^2 \mathsf{T}_c^4}{45}, \tag{5.81}$$

which yields

$$\mathsf{T}_c = \left(\frac{3}{2}\right)^{1/4}\frac{M_c}{\pi} \simeq 470\,\mathrm{MeV}. \tag{5.82}$$

The all order in h improvement discussed here allows us to take more seriously our results at high temperature. At high temperatures the kernel, which corresponds to the Boltzmann factor, is of order unity and thus our original assumption of smallness of H makes sense. This is indeed obvious from Eq. (5.79). As a result we now reproduce the expected asymptotic free gluon plasma result for the free energy.

On the other hand this improvement affected very little our previous results in the transition region. The transition temperature is shifted only by about 5%. The same is true for the the value of the parameter H at low momentum. As before, we find that the deconfinement phase transition is strongly first order with latent heat $\Delta E = \frac{4\pi^2 N^2}{45}\,\mathsf{T}_c^4$.

Although the actual value of the transition temperature is considerably larger than the lattice estimate, as explained earlier it makes more sense to look at dimensionless quantities. In particular, if we identify $2M_c$ with the mass of the lightest glueball (see however [43] and the discussion in the previous section), we find

$$\frac{\mathsf{T}_c}{2M_c} = \frac{1}{2\pi}\left(\frac{3}{2}\right)^{1/4} \simeq 0.18\,. \tag{5.83}$$

This is in excellent agreement with the lattice estimate for $SU(3)$ pure gauge theory [77]. We should however caution that given the uncertainties in our calculation this agreement may well be fortuitous.

6. Conclusions

We have tried here to review critically the application of the variational principle to Quantum Field Theories with gauge invariance, with the main focus on the approach developed by Ian Kogan and collaborators [14, 30, 42, 50, 70].

Although it is too early to decide whether this approach can be a useful calculational scheme for strongly interacting gauge theories, we can draw encouragement from its performance in the non-trivial toy models. In particular, in compact QED in 2+1 dimensions, we have been able to reproduce all known non-trivial characteristics of the non-perturbative vacuum state: dynamical mass generation, confining potential between external charges and area law behavior of the spatial Wilson loop with parameterically correct values for the string tension and mass. Although this is the only example that we have covered extensively in this review, the method has also been applied to other lower dimensional systems, and it works very well in all cases. Thus the deconfining phase transition in 2+1 compact QED at finite temperature is described correctly [43]. In the (exactly solvable) Schwinger model the variational approach reproduces the exact ground state wave functional [78]. In the compact 2+1 QED with Chern-Simons term [79] it predicts a Kosterlitz-Thouless phase transition in the value of the Chern-Simons parameter, in agreement with earlier analysis [80].

In 3+1 dimensional gluodynamics, this variational method gives results which on the qualitative level at least, conform with our intuition about the structure of the ground state, both at zero and finite temperature. We find dynamical mass generation, corresponding to an acceptable value of the gluon condensate. At finite temperature we find a first order phase transition which corresponds to the Polyakov loop acquiring a non-zero average.

Although we have not calculated the string tension directly, the behavior of the Polyakov loop is very much indicative that this is indeed the deconfining phase transition. The value of the critical temperature (in units of glueball mass) we find is in good agreement with lattice results. We also found that in the low temperature phase the entropy remains zero all the way up to the transition temperature. This is a rather striking result, which has not been built into our variational ansatz, but rather emerged as the result of the dynamical calculation.

An important lesson we learned from the lower dimensional models is that the projection of the Gaussian trial state onto the gauge invariant Hilbert subspace dictates most, if not all, of the important aspects of the non-perturbative physics. It was absolutely essential to perform the projection non-perturbatively, fully taking into account the contribution of the overlap between gauge rotated Gaussians into the variational energy prior to minimization.

The same conclusion carries over to the pure Yang-Mills theory. We have seen that from the point of view of the effective σ-model the energy is minimized in the disordered phase. In other words, the low momentum fluctuations of the field U are large, unlike in the perturbative regime, where U is close to a unit matrix. From the point of view of the trial wave functional, this means that the off-diagonal contributions, coming from the Gaussian WF gauge rotated by a slowly varying gauge transformation, are large. It is these "off diagonal" contributions to the energy that lowered the energy of the best trial state below the perturbative value. In the low temperature phase the vanishing of the entropy was also a direct consequence of the effective σ-model being in the disordered phase, and thus of the non-perturbative nature of the gauge projection. The accounting for these off diagonal terms non-perturbatively is the main distinction between this approach and other attempts [23–28] to implement the variational principle in gauge theories.

Many outstanding questions remain. Is the best variational state confining? How do we calculate the interaction potential between external sources? How do we understand better the relation between the variational parameter and the glueball masses? Can we extend the Ansatz to include (massless) fermions?

Both to be confident in our results and to be able to approach these questions we need first and foremost to have a better way of treating analytically the effective non-linear σ-model. The use of the mean field approximation in the effective σ-model was the main source of uncertainties in our calculations both at zero and finite temperature. We believe that it should be possible to treat the σ-model in a better way, perhaps along the lines of a continuum version of [74]. Such an improvement is crucial to clarify whether the qualitatively appealing results that we have described here are a kind of fluke due to an interplay of two bad approximations (variational and mean field) or are genuine predictions of a useful, workable variational approach. Personally we do believe that these results are genuine and that there is enough scope for further development of the approach which warrants continuing active investigations.

Acknowledgments

We want to thank our collaborators Will Brown, Juan Pedro Garrahan, Ben Gripaios and Ben Svetitsky for an enjoyable collaboration spanning several years.

One of us (J.G.M.) wishes to express his gratitude to the NA60 collaboration, especially to Carlos Lourenço and João Seixas, for their hospitality. The work of J.G.M. was supported by the Fundação para a Ciência e a Tecnologia (Portugal) under contract SFRH/BPD/12112/2003.

References

1. M. Creutz, *Yang-Mills fields and the lattice*, [hep-lat/0406007]
2. H. Leutwyler, in M. Shifman (ed.): *At the frontier of particle physics* (World Scientific, 2001) Vol. 1, p. 271 [hep-ph/0008124].
3. I. Kogan and A. Kovner, in M. Shifman (ed): *At the frontier of particle physics* (World Scientific, 2001) Vol. 4, p. 2335 [hep-th/0205026].
4. J. Greensite, Progr. Part. Nucl. Phys. **51**, 1 (2003); [hep-lat/0209138].
5. See A. Armoni, M. Shifman and G. Veneziano in this volume, [hep-th/0403071].
6. R. Feynman, in Wangerooge 1987, Proceedings, *Variational calculations in quantum field theory*, L. Polley and D. Pottinger, eds (World Scientific, Singapore, 1988).
7. P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).
8. P. A. M. Dirac, Phys. Rev. **114**, 924 (1959).
9. See for example J. Negele and H. Orland, *Quantum many particle systems* (Adddison - Wesley, 1987).
10. M. Moshe and J. Zinn-Justin; Phys.Rept.**385** 69 (2003) [hep-th/0306133].
11. A. Kovner, in M. Shifman, (ed.): *At the Frontier of Particle Physics* (World Scientific, Singapore, 2001) Vol. 3, p. 1777 [hep-ph/0009138].
12. *Selected Papers, with Commentary, of Tony Royle Hilton Skyrme*, ed. G. Brown (World Scientific, Singapore, 1994).
13. For review of the instanton physics of QCD see D. Dyakonov, *Chiral Symmetry Breaking by Instantons*, lectures at the Enrico Fermi School in Physics, 1995, hep-ph/9602375.
14. I. I. Kogan and A. Kovner, Phys. Rev. D **52**, 3719 (1995) [hep-th/9408081].
15. M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **147**, 385 (1979).
16. S. Nojiri, Z. Phys. C **22**, 245 (1984).
17. B. Rosenstein and A. Kovner, Phys. Lett. B **177**, 71 (1986).
18. V. N. Gribov, Nucl. Phys. B **139**, 1 (1978).
19. D. Schutte, Phys. Rev. D **31**, 810 (1985).
20. R. E. Cutkosky, Phys. Lett. B **194**, 91 (1987).
21. R. E. Cutkosky and K. C. Wang, Phys. Rev. D **37**, 3024 (1988).
22. J. P. Greensite, Nucl. Phys. B **158**, 469 (1979).
23. M. Consoli and G. Preparata, Phys. Lett. B **154**, 411 (1985).
24. G. Preparata, Nuovo Cim. A **96**, 366 (1986).
25. G. Preparata, Nuovo Cim. A **96**, 394 (1986).
26. A. K. Kerman and D. Vautherin, Annals Phys. **192**, 408 (1989).
27. C. Heinemann, E. Iancu, C. Martin and D. Vautherin, Phys. Rev. D **61**, 116008 (2000) [hep-ph/9911515].
28. O. Schroeder and H. Reinhardt, Ann. Phys. **307**, 452 (2003) [hep-ph/0306244].
29. J. Greensite and M. B. Halpern, Nucl. Phys. B **271**, 379 (1986).
30. I. I. Kogan and A. Kovner, Phys. Rev. D **51**, 1948 (1995) [hep-th/9410067].
31. S. R. Coleman, Phys. Rev. D **11**, 2088 (1975).
32. A. M. Polyakov, *Gauge fields and strings* (Chur, Switzerland: Harwood, 1987) 301 P. (Contemporary Concepts in Physics, 3).
33. S. Samuel, Phys. Rev. D **18**, 1916 (1978).
34. D. J. Amit, Y. Y. Goldschmidt and G. Grinstein, J. Phys. A **13**, 585 (1980).
35. D. Boyanovsky, J. Phys. A**22**, 2601 (1989).
36. A. M. Polyakov, Phys. Lett. B **59**, 82 (1975).
37. A. M. Polyakov, Nucl. Phys. B **120**, 429 (1977).
38. D. Diakonov, Surveys High Energ. Phys. **14**, 29 (1999) [hep-th/9805137].

39. K. Zarembo, Phys. Lett. B **421**, 325 (1998) [hep-th/9710235].
40. K. Zarembo, Mod. Phys. Lett. A **13**, 2317 (1998) [hep-th/9806150].
41. A. Kovner and B. Svetitsky, Phys. Rev. D **60**, 105032 (1999) [hep-lat/9811015].
42. W. E. Brown and I. I. Kogan, Int. J. Mod. Phys. A **14**, 799 (1999) [hep-th/9705136].
43. B. M. Gripaios, Int. J. Mod. Phys. A **18**, 85 (2003) [hep-ph/0204310].
44. R. D. Pisarski and F. Wilczek, Phys. Rev. D **29**, 338 (1984).
45. F. R. Brown *et al.*, Phys. Rev. Lett. **65**, 2491 (1990).
46. F. R. Brown *et al.*, Phys. Lett. B **251**, 181 (1990).
47. E. Brezin, J. C. Le Guillou and J. Zinn-Justin, Phys. Rev. D **8**, 2418 (1973).
48. M. F. Atiyah and N. S. Manton, Phys. Lett. B **222**, 438 (1989).
49. J. B. Kogut, Rev. Mod. Phys. **51**, 659 (1979).
50. W. Brown, J. P. Garrahan, I. I. Kogan and A. Kovner, Phys. Rev. D **59**, 034015 (1999) [hep-ph/9808216].
51. E. V. Shuryak, Nucl. Phys. B **203**, 93 (1982).
52. E. V. Shuryak, Nucl. Phys. B **203**, 116 (1982).
53. D. Diakonov and V. Y. Petrov, Nucl. Phys. B **245**, 259 (1984).
54. E. V. Shuryak, Nucl. Phys. B **328**, 102 (1989).
55. E. V. Shuryak and J. J. M. Verbaarschot, Nucl. Phys. B **341**, 1 (1990).
56. D. Diakonov and V. Y. Petrov, Phys. Lett. B **224**, 131 (1989).
57. A. M. Polyakov, Mod. Phys. Lett. A **3**, 325 (1988).
58. A. Y. Alekseev and S. L. Shatashvili, Mod. Phys. Lett. A **3**, 1551 (1988).
59. G. 't Hooft, Nucl. Phys. B **138**, 1 (1978).
60. A. M. Polyakov, Phys. Lett. B **72**, 477 (1978).
61. L. Susskind, Phys. Rev. D **20**, 2610 (1979).
62. P. Arnold and C.-x. Zhai, Phys. Rev. D **51**, 1906 (1995) [hep-ph/9410360].
63. E. Braaten and A. Nieto, Phys. Rev. D **51**, 6990 (1995) [hep-ph/9501375].
64. E. Braaten and A. Nieto, Phys. Rev. D **53**, 3421 (1996) [hep-ph/9510408].
65. K. Kajantie *et al.*, Phys. Rev. Lett. **79**, 3130 (1997) [hep-ph/9708207].
66. K. Kajantie, M. Laine, K. Rummukainen and Y. Schroder, Phys. Rev. Lett. **86**, 10 (2001) [hep-ph/0007109].
67. A. Hart and O. Philipsen, Nucl. Phys. B **572**, 243 (2000) [hep-lat/9908041].
68. G. Boyd *et al.*, Nucl. Phys. B **469**, 419 (1996) [hep-lat/9602007].
69. S. Datta and S. Gupta, Phys. Lett. B **471**, 382 (2000) [hep-lat/9906023].
70. I. I. Kogan, A. Kovner and J. G. Milhano, JHEP **12**, 017 (2002) [hep-ph/0208053].
71. O. J. P. Eboli, R. Jackiw and S.-Y. Pi, Phys. Rev. D **37**, 3557 (1988).
72. W. A. Bardeen and M. Moshe, Phys. Rev. D **28**, 1372 (1983).
73. C. P. Korthals Altes, Nucl. Phys. B **420**, 637 (1994) [hep-th/9310195].
74. J. B. Kogut, M. Snow and M. Stone, Nucl. Phys. B **200**, 211 (1982).
75. B. M. Gripaios and J. G. Milhano, Phys. Lett. B **564**, 104 (2003) [hep-ph/0302172].
76. B. M. Gripaios, Phys. Rev. D **67**, 025023 (2003) [hep-th/0211104].
77. M. J. Teper, *Glueball masses and other physical properties of SU(N) gauge theories in D=3+1: A review of lattice results for theorists*, hep-th/9812187.
78. W. Brown, J.P. Garrahan, I. Kogan and A. Kovner; Phys. Rev. D **62**, 016004 (2000) [hep-th/9912136].
79. I. Kogan and A. Kovner, Phys. Rev. D **53**, 4510 (1996) [hep-th/9507137].
80. I. Kogan, JETP Lett. **45**, 709 (1987).