QUANTIZATION OF NON-ABELIAN GAUGE THEORIES

V.N. GRIBOV

_Leningrad Nuclear Physics Institute, Gatchina, Leningrad 188350, USSR_

Received 13 January 1978

It is shown that the fixing of the divergence of the potential in non-Abelian theories does not fix its gauge. The ambiguity in the definition of the potential leads to the fact that, when integrating over the fields in the functional integral, it is apparently enough for us to restrict ourselves to the potentials for which the Faddeev-Popov determinants are positive. This limitation on the integration range over the potentials cancels the infrared singularity of perturbation theory and results in a linear increase of the charge interaction at large distances.

1. Introduction

The quantization problem for non-Abelian gauge theories within the framework of perturbation theory was solved by Feynman [1], DeWitt [2] and Faddeev and Popov [3]. A subsequent analysis of perturbation theory in such theories (Politzer [4], Gross and Wilczek [5], Khriplovich [6]) has shown that they possess a remarkable property called asymptotic freedom. This property resides in the fact that zero-point field oscillations increase the effective charge not in the high-momentum region as in QED [7], but in the low-momentum region, i.e. at large distances between the charges. This gave hope that such theories may incorporate the phenomenon of color confinement which is fundamental to present day ideas concerning the structure of hadrons.

Answering the question as to whether color confinement occurs in non-Abelian theories proved to be a very difficult problem since the non-Abelian fields possessing charges ("color") strongly interact in the large-wavelength region.

Strong interaction between vacuum fluctuations in the region of large wavelengths means that at these wavelengths a significant role is played by field oscillations with large amplitudes, for which the substantially non-linear character of non-Abelian theories is decisive. Thus, the problem of color confinement is closely connected with that of the quantization of large non-linear oscillations. In this paper we show that in the region of large field amplitudes the method of quantization by Faddeev and Popov is to be improved.

As will be demonstrated, it is very likely that this improvement reduces simply
to an additional limitation on the integration range in the functional space of non-Abelian fields, which consists in integrating only over the fields for which the Faddeev-Popov determinant is positive. This additional limitation is not significant for high-frequency oscillations, but substantially reduces the effective oscillation amplitudes in the low-frequency region. This in turn results in the fact that the "effective" charge interaction does not tend to infinity at finite distances as occurs in perturbation theory, but goes to infinity at infinitely large distances between charges, if at all.

2. Non-uniqueness of gauge conditions

The difficulties in the quantization of gauge fields are caused by the fact that the gauge field Lagrangian

\[ \mathcal{L} = -\frac{1}{4g^2} \text{Sp} F_{\mu\nu} F^{\mu\nu}, \]  

\[ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu A_\nu], \]  

where \( A_\mu \) are antihermitian matrices, \( \text{Sp} A_\mu = 0 \), being invariant with respect to the transformation

\[ A_\mu = S^+ A'_\mu S + S^+ \partial_\mu S, \quad S^+ = S^{-1}, \]  

contains non-physical variables which must be eliminated before quantization. A conventional method of relativistic invariant quantization [3] is as follows. Let us consider a functional integral

\[ W = \int e^{-\int \mathcal{L} d^4x} \prod dA'_\mu \]  

in Euclidean space-time and imagine the functional space \( A_\mu \) in the form shown in

![Diagram](image-url)
fig. 1, where the transverse and longitudinal components of the field $A_\mu$ are plotted along the horizontal and vertical axes, respectively. Then for fixed $A_\mu$, eq. (3) defines the line $L$ (as a function of $S$) on which $\mathcal{L}$ is constant. The Faddeev and Popov idea is that, instead of integrating over $A_\mu$, one should integrate over matrices $S$ and fields $A_\mu$ which have a certain divergence $f = \partial_\mu A_\mu$. Then $W$ is written in the form

$$W = e^{-\int \mathcal{L} d^4x} \prod \int dA' \frac{1}{\Delta(A')} \times \int dS \cdot S^+ \delta \left[ f - S^+ (\partial_\mu A'_\mu + [\nabla_\mu (A') S \partial_\mu S^+]) \right] S,$$

where

$$\Delta(A') = \int dS \cdot S^+ \delta \left[ f - S^+ (\partial_\mu A'_\mu + [\nabla_\mu (A') S \partial_\mu S^+]) \right] S,$$

$$\nabla_\mu (A') = \partial_\mu + A'_\mu.$$

Since the variation with respect to $S$ of the expression under the sign of the $\delta$ function is $\partial_\mu [\nabla_\mu (A), S^+ \partial S]$ then

$$\frac{1}{\Delta(A')} = ||\hat{\Delta}(A)||,$$

where the operator $\hat{\Delta}(A)$ is defined by the equation

$$\hat{\Delta}(A) \psi = \partial^2 \psi + \partial_\mu [A_\mu \psi] \equiv \partial_\mu [\nabla_\mu (A) \psi].$$

Replacing in (4) the variables

$$A'_\mu = S A_\mu S^+ + S \partial_\mu S^+,$$

we obtain

$$W = e^{-\int \mathcal{L} d^4x} \delta (f - \partial_\mu A_\mu) \prod ||\hat{\Delta}(A)|| \int dS \cdot S^+.$$

Since (10) does not depend upon $f$, we may integrate over $f$ with any weight function, $\exp \{((1/2\alpha g^2) \text{Sp} \int f^2 dx)\}$ being commonly used for this purpose. In so doing, with the integration over $S$ omitted, $W$ is obtained in the form

$$W = \int \exp \left( - \int \mathcal{L} d^4x + \frac{1}{2\alpha g^2} \text{Sp} \int (\partial_\mu A_\mu)^2 d^4x \right) ||\hat{\Delta}(A)|| dA.$$
gauge-equivalent fields $A_\mu$ with a given divergence correspond to a given field $A'_\mu$ is typical in non-Abelian theories. Indeed, in order for two gauge-equivalent fields $A_{1\mu}$ and $A_{2\mu}$ with the same divergence to exist, there should be a unitary matrix $S$

\[ A_{2\mu} = S^+ A_{1\mu} S + S^+ \partial_\mu S , \]  

and satisfying the equation

\[ \partial_\mu S^+ [V_\mu (A_{1}) , S] = 0 , \]  

or obtained from it through the substitution of $A_{2\mu}$ for $A_{1\mu}$ and $S^+$ for $S$. In an Abelian theory, where $S = e^{i\phi}$ is a unit matrix, eq. (13) reduces to the Laplace equation

\[ \partial^2 \phi = 0 , \]  

and to eliminate non-uniqueness it is sufficient for us to confine ourselves to the fields which vanish at infinity. In a non-Abelian case, the non-linear equation (13) cannot have growing solutions and hence even for $A_{1\mu} = 0$ it has solutions for $S$ leading to a decreasing $A_{2\mu}$. In the appendix we consider examples of the solution to eq. (13) for $A_{1\mu} = 0$, from which it will be evident that a set of these solutions, i.e. of the transverse potentials equivalent to the vacuum, are in order of magnitude similar to a set of solutions to the Laplace equation, which grow at infinity, but that all of these, though corresponding to such $S$ that do not tend to unity at infinity, result in the potentials $A_{2\mu}$ decreasing as $1/r$.

In the appendix we shall also show that, with values of $A_{1\mu}$ large enough, (13) has solutions for $S$ which tend to unity at $r \to \infty$ and result in rapidly decreasing $A_{2\mu}$. Under these circumstances, to calculate correctly the functional integral in a non-Abelian theory, we must either replace eq. (11) by the expression

\[ \mathcal{W} = \int e^{-\int \mathcal{L} d^4 x} \frac{||\Box (A)||}{1 + N(A)} dA , \]  

where $N$ is the number of fields gauge-equivalent to a given field $A$ and having the same divergence, or restrict the integration range in the functional space so as to have no repetitions.

An intermediate case, when both things are required, is possible. For instance, when integrating only over $A_\mu$ vanishing at infinity faster than $1/r$, we eliminate the fields gauge-equivalent to “small” fields, but for large enough $A_\mu$ the gauge-equivalent fields will remain and hence $N(A)$ in (15) will be needed. In this case, the problem of calculating $N(A)$ reduces to the analysis of solutions of eq. (13) tending to unity under $r \to \infty$ which depend on the character and the magnitude of the field $A_\mu$. This problem seems to be almost hopeless, but we shall demonstrate below that there exists a possibility of a sufficiently universal solution leading to physically interesting results.
3. A limitation on the integration range in the functional space

In order to gain some insight into the nature of non-uniqueness in the functional space $A_\mu$, let us see for what fields $A_\mu$ there exist gauge-equivalent fields close to the former and having the same divergence, i.e. what are the conditions for solving eq. (13) with $S$ close to unity. Substituting into (13)

$$S = 1 + \alpha, \quad \alpha^+ = -\alpha, \quad (16)$$

we get

$$\square(A)\alpha = \delta_\mu \left[ \nabla_\mu(A), \alpha \right] = 0. \quad (17)$$

Since $\square(A)$ is the operator whose determinant enters into the functional integral, and eq. (17) is simply an equation for the eigenfunction of this operator with a zero eigenvalue, we draw the conclusion that the field $A_\mu$ can only have a close equivalent field when the Faddeev-Popov determinant for this field turns into zero, or (which is the same) if the field is such that the Faddeev-Popov ghost has a zero-mass bound state. Clearly, if the field $A_\mu$ is sufficiently small in the sense that the product of the width of the region where $A_\mu$ differs from zero with its amplitude over the same region is small, then there are no bound states in such a field, i.e. the equation

$$-\square(A)\psi = \epsilon\psi \quad (18)$$

is solvable for positive $\epsilon$ only. For a sufficiently large magnitude and a particular sign of the field there appears a solution with $\epsilon = 0$, which becomes one with a negative $\epsilon$ as the field increases further. For a particular still greater magnitude of the field, the level with a zero $\epsilon$ reappears, etc. Thus, one can imagine the fields for which eq. (17) is solvable as dividing the functional space into regions over each of which eq. (18) has a given number of eigenvalues, i.e. there exist a given number of bound states for the Faddeev-Popov ghosts. Fig. 2 shows this division of the field space into the regions $C_0, C_1, ..., C_n$, over which the ghosts have $0, 1, 2, ..., n$ bound states, by the lines $\xi_1, \xi_2, ..., \xi_n$ on which the ghosts have zero-mass levels.

Hence, if we imagine the space of the fields $A_\mu$ in this way, it may be asked whether two near equivalent fields that can exist close to the line, say, $\xi_1$, are located on different sides of this line, i.e. one field within the region $C_0$, another in $C_1$, or may be arbitrarily situated. We shall demonstrate below that, indeed, if there are only two near equivalent fields, they will always lie on different sides of the corresponding curve $\xi_1$. Moreover, we shall show that for any field located within the region $C_1$ close to the curve $\xi_1$ there is an equivalent field within the region $C_0$ close to the same curve.

If we could prove that not only for small neighbourhoods close to the curves $\xi_n$, but also for any field in the region $C_n$ there is an equivalent field in the region $C_{n-1}$, we would prove that instead of integrating over the entire space of the fields $A_\mu$, it would be sufficient for us to confine ourselves to the region $C_0$, i.e. to integrating only with respect to the fields $A_\mu$ which create no bound states for the ghosts (up
to the first zero of the Faddeev-Popov determinant). However, even at the level of the things we can prove, there is a significant statement that for the functional integral, integration can be cut off on the boundary of the region $C_0$, and if there exist fields $A_\mu$ not equivalent to those over the region $C_0$, they are separated by a finite region from the boundary $\xi_1$ and are in no way connected with the region of small fields $A_\mu$ lying within $C_0$ for which perturbation theory holds.

We shall assume below (until the contrary is established) that these additions are either non-existent or insignificant and that the integral (15) is determined over the region $C_0$. Generally speaking, we must retain $N(A)$ in (15) because we have not proved that there are no equivalent fields over the region inside $C_0$. We shall return to this subject below.

4. Proof of the field equivalence over the regions $C_0$ and $C_1$ close to their boundary

We shall first of all show that if the field $A_\mu$ is close to $\xi_1$, then there is always a similar field equivalent to the former, i.e. for such a field eq. (13) always has a solution with $S$ little different from unity and tending to unity at $r \to \infty$. The condition for $S \to 1$ as $r \to \infty$ is required because a solution with $S \to 0$ yields equivalent fields greatly different from the initial field. As shown in the appendix, these fields are located within the region $C_\infty$. We write the field $A_\mu$ in the form

$$A_\mu = C_\mu + a_\mu ,$$

where $C_\mu$ lies on $\xi_1$, i.e. there exists $\varphi_0$ decreasing at infinity and satisfying the equation

$$\partial_\mu [\nabla_\mu (C_\mu \varphi_0) = 0 ,$$

and $a_\mu$ is small compared to $C_\mu$ in the sense defined below.
Substituting the matrix $S$ in the form $\exp \alpha$ into (13) and confining ourselves to the terms quadratic in $\alpha$, we obtain

$$\partial_\mu [\nabla_\mu (A), \alpha] - \frac{1}{2} \partial_\mu [\alpha [\nabla_\mu, \alpha]] = 0 .$$

(21)

The solution to (21) may be sought for in the form

$$\alpha = \varphi_0 + \tilde{\alpha}, \quad \tilde{\alpha} \ll \varphi_0 .$$

(22)

Substituting (19), (22) into (21) and taking into account (20), we get

$$\partial_\mu [\nabla_\mu (C), \tilde{\alpha}] = -\partial_\mu [a_\mu, \varphi_0] + \frac{1}{2} \partial_\mu [\varphi_0 [\nabla_\mu (C), \varphi_0]] .$$

(23)

Since the right-hand side of (23) vanishes at infinity, for a vanishing solution it is sufficient for the r.h.s. to be orthogonal with respect to $\varphi_0$, i.e.

$$\text{Sp} \int d^4x \{\varphi_0 \partial_\mu [a_\mu, \varphi_0] - \frac{1}{2} \varphi_0 \partial_\mu [\varphi_0 [\nabla_\mu (C), \varphi_0]]\} = 0 .$$

(24)

This equation defines the normalization of $\varphi_0$, i.e. the difference between $S$ and unity.

Hence, we have found $S$, and thus may now find the field $A'_\mu = C_\mu + a'_\mu$, equivalent to the field $A_\mu = C_\mu + a_\mu$,

$$a'_\mu = a_\mu + [\nabla_\mu (C), \varphi_0] ,$$

(25)

and clear up the question of whether $A'_\mu$ and $A_\mu$ lie on both sides of $\xi$, i.e. whether it is true that there exists a bound ghost state in one of the fields $A_\mu, A'_\mu$ and that there is no such state in the other. For this purpose, it is sufficient to calculate the shift of the level from zero due to the fields $a_\mu$ and $a'_\mu$:

$$\epsilon(a) = -\frac{1}{N^2} \text{Sp} \int d^4x \varphi_0 \partial_\mu [a_\mu, \varphi_0] ,$$

$$\epsilon(a') = -\frac{1}{N^2} \text{Sp} \int d^4x \{\varphi_0 \partial_\mu [a_\mu, \varphi_0] + \varphi_0 \partial_\mu [[\nabla_\mu (C), \varphi_0] \varphi_0]\} ,$$

$$N^2 = \text{Sp} \int \varphi_0^2 d^4x .$$

(26)

Using (24), which defines the normalization $\varphi_0$, we obtain

$$\epsilon(a) = -\epsilon(a') ,$$

(27)

which was to be proved.

Non-strictness of the derivation due to the fact that we ignored the continuous character of the spectrum at $\varepsilon > 0$ can be easily avoided, and we shall not dwell on that. The derivation can just as readily be repeated for the fields close to any $\xi_n$, imposing the orthogonality condition of the eigenfunction $\varphi_0$ on the eigenfunctions of other bound states together with eq. (24).

In essence, one can gain some insight into the cause of the field equivalence over
the regions $C_0$ and $C_1$ when considering the effective Lagrangian in (15),
\[ \tilde{\mathcal{L}} = -\frac{1}{4g^2} \text{Sp} F_{\mu\nu}F_{\mu\nu} - \frac{1}{2ag^2} \text{Sp}(\partial_{\mu}A_{\mu})^2, \] (28)
close to $\ell_1$. The property of the fields $C_{\mu}$ lying on the line $\ell_1$, is that there exists a solution to eq. (20). Consider now the fields of the form
\[ A_{\mu} = C_{\mu} + [\nabla_{\mu}(C), \varphi_0]. \] (29)
Then the Lagrangian $\tilde{\mathcal{L}}$, to second-order in $\varphi_0$, takes the form
\[ \tilde{\mathcal{L}}(A) = \tilde{\mathcal{L}}(C) - \frac{1}{4} \text{Sp} \{ [F_{\mu\nu}(C), \varphi_0] [F_{\mu\nu}(C), \varphi_0] \]
\[ + 2F_{\mu\nu}(C) \{ [\nabla_{\mu}(C), \varphi_0], \nabla_{\nu}, \varphi_0] \}, \] (30)
i.e. $\tilde{\mathcal{L}}$ has an extremum along $[\nabla_{\mu}(C), \varphi_0]$ and does not change upon replacing $\varphi_0$ by $-\varphi_0$. From this it follows that the fields $A_{\mu}^\pm = C_{\mu} \pm [\nabla_{\mu}(C), \varphi_0]$ are equivalent. Clearly $\epsilon(A_{\mu}^+) = -\epsilon(A_{\mu}^-)$.

Since the direction $[\nabla_{\mu}(C), \varphi_0]$ is generally not orthogonal to the curve $\ell_1$, the distribution of the equivalent fields can correspond to that shown in fig. 3, where the equivalent fields are located on dashed lines in the opposite directions to $\ell_1$. If structural lines can intersect as shown in fig. 3, this will result in the existence of at least pairs of equivalent fields within the region $C_0$. As will be demonstrated in the appendix, such intersections actually occur, and in this sense the equivalent fields do exist in $C_0$, but in the examples considered, these equivalent fields turn out to be mirror-reflected ones that are always equivalent. Hence, the field doubling thus obtained is independent of the field magnitude and insignificant in the functional integral.

5. Gauge non-uniqueness and limitation on the integration range over the fields in physical space-time

So far we have discussed the functional integration over non-Abelian fields defined in four-dimensional Euclidean space. This somewhat simplifies the mathematics, but makes it more difficult to understand the physical content of the
theory and leaves a feeling of dissatisfaction related to the need for analytical continuation of the results.

Certainly, the general statement that the integration should only be performed over non-equivalent fields is independent of the nature of the space, and formula (15) holds. The difference is in the real form of eq. (13) defining the number of equivalent fields. Now, this equation is a hyperbolic one having non-zero solutions even in an Abelian case ($S = \exp \varphi$; $\varphi$ is an arbitrary solution to a wave equation). One of the ways for eliminating this non-uniqueness is a change-over to Euclidean space. In normal space-time this non-uniqueness does not show itself because, according to Feynman, we integrate over fields $A_\mu$ which have only positive frequencies at $t \to -\infty$ and negative frequencies at $t \to +\infty$. In this case, if we want to have two equivalent $A_\mu$ and $A'_\mu$ under the same boundary conditions, then $A'_\mu - A_\mu$ as $t \to -\infty$ and $+\infty$ should contain only negative and positive frequencies, respectively. This indicates that equivalent trajectories will only occur through those solutions of eq. (13) which have only negative and positive frequencies as $t \to -\infty$ and $+\infty$, respectively. Clearly, these conditions play the same part as those at infinity in the Euclidean case, and the linear equations will have no solution at $A_\mu = 0$ because of frequency conservation. Such solutions will exist for non-linear equations or for sufficiently large fields $A_\mu$. For instance, eq. (18) at $\epsilon = 0$ is one for the ghost wave function in the external field $A_\mu$. If the field $A_\mu$ is situated on the line $\mathcal{L}_1$, such an equation has a solution under the boundary condition specified above, and defines the ghost transition from the state with negative energy to that with positive energy. Since the ghosts are quantized in the same way as fermions, the process is apparently interpreted as the classic formation of ghost pairs in the external field. In a similar manner it can be said that solutions of eq. (13) result in the fields $A_\mu$ which differ from $A_\mu$ in pairs of the gauge quanta produced.

The restriction of the integration in the functional integral to the region $C_0$ implies the restriction to the fields in which no classical ghost formation occurs because the formation of ghosts merely redefines the fields $A_\mu$.

6. The effect of the field magnitude restriction on the zero-point oscillations and interaction in the low-momenta region

In this section we shall try to analyze how a limitation on the integration range over the field in the functional integral affects the physical properties of non-Abelian theories.

We shall proceed from eq. (15) for the action, disregarding the possibility for the equivalent fields to exist in $C_0$:

$$ W = \int e^{-\int \mathcal{L} d^4x} \|\mathcal{O}(A)\| \mathcal{V}(\square) dA, $$

where $\mathcal{V}(\square)$ indicates that the integration is performed only over the region $C_0$. 

First of all, let us see whether the restriction $\mathcal{V}(\square)$ is significant from the standpoint of what we know from the perturbation theory analysis. For this purpose, consider the Green function of the Faddeev-Popov ghost

$$G(k) = -\frac{1}{W} \int e^{-\int \mathcal{L} d^4x} \langle k | \prod_{\square(\lambda)} \mathcal{V}(\square) | 0 \rangle \frac{1}{\square(\lambda)} d\lambda.$$

(32)

It is well known that, if we calculate $G(k)$ in perturbation theory, i.e., perform the integration over $A$ in (32), omitting $\mathcal{V}(\square)$ and expanding over the coupling constant, we get

$$G(k) = \frac{1}{k^2} \left( 1 - \frac{11g^2C_2}{48\pi^2} \ln \frac{\Lambda^2}{k^2} \right)^{(3/22)(3/2-\alpha/2)}.$$

(33)

where $\Lambda$ is the ultraviolet cutoff, $\alpha$ is the gauge parameter in $\mathcal{D}$. From this it is obvious that $G(k)$ becomes large at $\alpha < 3$ and physical $k^*$ (in the Euclidean space) such that

$$1 = \frac{11g^2C_2}{48\pi^2} \ln \frac{\Lambda^2}{k^2} \sim g^2,$$

where (33) still holds. From the standpoint of (32), $G(k)$ can be large only due to the integration range for the fields where $\square$ is small, i.e. close to the lines $\mathcal{L}_n$.

It is interesting that transverse fields (low $\alpha$) act on the ghosts as attractive fields and longitudinal fields as repulsive ones. Since the influence of longitudinal fields cancels in the calculations of gauge-invariant quantities, we may say that we study the contribution to the functional integral close to the curves $\mathcal{L}_n$ when calculating $G(k)$ near the "infrared pole", and hence $\mathcal{V}(\square)$ is definitely significant at momenta below or of the order of the "infrared pole" position, whereas at large $k$ we are within $C_0$ (low $A$), where $\mathcal{V}(\square)$ is insignificant and perturbation theory works.

Furthermore, $\mathcal{V}(\square)$ makes it impossible for a singularity of $G(k)$ to exist at finite $k^2$ because, with $k^2$ below the singularity position, $G(k)$ would either reverse its sign or become complex. Both things would indicate that $\square$ has ceased to be a positively defined quantity, i.e. we have left the region $C_0$ when integrating over $A_\mu$.

The only possibility that now remains is that $k^2G(k)$ has a singularity at $k^2 = 0$. Such a possibility would indicate that at $k^2 = 0$ we feel the fields on the line $\mathcal{L}_1$.

Up to now, all attempts at finding the mechanism for removal of the "infrared pole" have been not successful. Higher corrections [8--10] and instantons [11,12] only bring it nearer. If no other causes are found, $\mathcal{V}(\square)$ will be the cause. The fact that there are no other causes for the interaction cutoff is equivalent to the statement that without $\mathcal{V}(\square)$ zero fluctuations of the fields tend to leave the region $C_0$. Hence it appears quite natural that the fields closest to the boundary of the region $C_0$, i.e. connected with the singularity of $k^2G(k)$ at $k^2 = 0$, will correspond to the real vacuum if $\mathcal{V}(\square)$ is taken into account.
For checking the above by a specific calculation, one must write $\mathcal{V}(\overline{A})$ in a constructive way, but unfortunately we have not yet succeeded in doing this. All we could have done was to write this criterion to second order in perturbation theory and then calculate the functional integral taking no account of the interaction except for $\mathcal{V}(\overline{A})$. In this case it turns out that there appears a characteristic scale $\kappa^2$ defined by the condition $g^2 \Lambda^2 / \kappa^2 \sim 1$, so that at $k^2 \gg \kappa^2$ the gluon and ghost Green functions remain free. The gluon Green function $D(k)$ has complex singularities and is non-singular at $k^2 \rightarrow 0$. The ghost Green function under $k^2 \rightarrow 0$ is $G(k) \sim C/k^4$.

If it were not for the roughness of the calculations and difficulties with complex singularities of $D(k)$, this would be the right thing for the colour confinement theory.

Let us show the way this is obtained. We write $G_{aa}(k, A) = -(k, a | 1/\overline{A} | k, a)$ in the form of an expansion in perturbation theory (where $a$ is the isotopic index)

$$G_{aa}(k, A) = \frac{-1}{k^2} + \frac{1}{k^2} + \frac{1}{k^2} + \cdots .$$

The first-order term gives no contribution to the diagonal element. The second-order term is

$$\mathcal{V} = \frac{4 \int d^4q}{k^4} \left[ A_\mu^a(q) A_\nu^a(-q) k_\mu (k-q)_\nu \right] \equiv \frac{V}{k^2} \sigma(k, A).$$

$A_\mu^a(q)$ is the Fourier component of the potential $A_\mu$, $V$ the volume of the system, $\sigma(k, A)$ defines positions of the poles $G(k, A)$, if any, to a second Born approximation since

$$G(k, A) \approx \frac{1}{k^2} \frac{V}{1 - \sigma(k, A)} .$$

In this case we assume, of course, that $k$ is conserved in a typical field of zero-point fluctuations ($\langle k | 1/\overline{A} | k' \rangle |_{k'=k}$ is proportional to the volume of the system which is replaced by $\delta(k - k')$ after averaging). The no-pole condition at a given $k$ is $\sigma(k, A) < 1$.

For simplicity, we choose a transverse gauge ($\alpha = 0$). On averaging over the gluon polarization directions $\lambda$, we have

$$\sigma(k, A) = 4 \int \frac{d^4q}{(2\pi)^4} \frac{|A^{a,\lambda}(q)|^2}{(k-q)^2} \left( 1 - \frac{(kq)^2}{k^2 q^2} \right) .$$

If $|A^{a,\lambda}(q)|^2$ over the main range of integration with respect to $q$ decreases monotonically with $q^2$, as will prove to be the case in what follows, then $\sigma(k, A)$ decreases as $k^2$ increases and hence as a no-level condition use can be made of

$$\sigma(0, A) = 3 \int \frac{d^4q}{(2\pi)^4} \frac{|A^{a,\lambda}(q)|^2}{q^2} < 1 .$$
Taking (38) as a condition for $\mathcal{V}(\square)$, replacing $\mathcal{L}$ by $\mathcal{L}_0$ in (31) and omitting $||\Box||$, instead of (31) we obtain a functional integral which is easy to calculate, if $\mathcal{V}(1 - o(k, A))$ is written in the form

$$\mathcal{V}(1 - o(0, A)) = \int \frac{d\beta}{2\pi i\beta} e^{\beta(1 - o(0, A))},$$

(39)

$$W = \int \frac{d\beta}{2\pi i\beta} e^{\beta} \int \prod dA^{a, \lambda} \times \exp \left\{ - \frac{1}{g^2} \sum q^2 |A^{a, \lambda}(q)|^2 - \frac{\beta}{V} \sum \frac{|A^{a, \lambda}|^2}{q^2} \right\},$$

(40)

where $V$ is the volume of the system. Calculating the integral over $A$, we get

$$W \sim \int \frac{d\beta}{2\pi i\beta} e^{\beta} \prod_q \left( \frac{1}{q^2 + \beta g^2/V} \right)^{3n/2},$$

(41)

$n$ being the number of isotopic states. The integral over $\beta$ can be obtained by the steepest-descent method, with the saddle-point value $\beta_0$ determined by

$$\frac{3n}{2} \frac{g^2}{V} \sum q^2 \frac{1}{q^4 + \beta_0 g^2/V} - 1 + \frac{1}{\beta_0} = 0.$$

(42)

Setting $\beta_0 g^2/V = \kappa^2$, with $V \to \infty$ we get

$$\frac{3}{2} ng^2 \int \frac{d^4q}{(2\pi)^4} \frac{1}{q^4 + \kappa^2} = 1$$

(43)

or

$$\frac{3n}{2\pi} g^2 \ln \frac{\Lambda^2}{\kappa^2} = 1.$$  

(44)

If the saddle-point value $\beta_0$ is known, we can return to the functional integral (40), substituting $\beta = \beta_0$ in it and omitting the integration over $\beta$, so as to obtain an effective functional integral for calculating the correlation functions of the fields $A$. In this case, $W$ is

$$W = \int dA \exp \left[ - \frac{1}{g^2} \sum_{k, \lambda, a} \left( k^2 + \frac{k^4}{\kappa^2} \right) |A^{\lambda, a}(k)|^2 \right].$$

(45)

Consequently, the gluon and ghost Green functions are

$$D^{ab}_{\mu\nu}(k) = A^a_{\mu}(k) A^b_{\nu}(k) = \delta_{ab} \left( \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right) \frac{g^2 k^2}{k^4 + k^2},$$

(46)

$$G(k) = \frac{1}{k^2 (1 - \sigma(k))} = \left\{ 1 - \frac{4g^2}{k^2} \int \frac{d^4q}{(2\pi)^4} \frac{k_\mu k_\nu D^{aa}_{\mu\nu}(q^2)}{(k - q)^2} \right\}^{-1}.$$
respectively, due to (43).

As \( k^2 \to 0 \)

\[ G(k) \approx \frac{32\pi k^2}{g^2 n k^4}, \quad (48) \]

in accordance with the above. The fact that the significant range of integration in the functional integral turns out to coincide with the boundary of the region \( \mathcal{I}_1 \), is evident without calculating \( G(k) \) because, when calculating the saddle-point value \( \beta_0 \), the last term in (42) has no effect at \( V \to \infty \) and hence \( \mathcal{V}(1 - \sigma) \) is equivalent to \( \delta(1 - \sigma) \). We would obtain the same result when calculating with the function \( (1 - \sigma)^2 \mathcal{V}(1 - \sigma) \), which is equivalent to an attempt at taking into account the effect of the determinant in (31).

7. Coulomb gauge

In sect. 6 we discussed the effect of limiting the integration over the fields on the properties of vacuum fluctuations in the invariant Euclidean formulation of the theory. In so doing, we adduced arguments for singularity of the ghost Green function as \( k^2 \to 0 \) (for example, \( 1/k^4 \)). This certainly is an indication of a substantial long-range effect in the theory that may result in colour confinement, but the ghost Green function in an arbitrary gauge is not connected directly with the Coulomb interaction at large distances. Hence, in this section we shall rewrite the foregoing analysis for the Coulomb gauge \([13]\) where the Green function of the ghost determines directly the Coulomb interaction. We shall show that the situation which involves a restriction on the integration range over fields and a cutoff of the infrared singularity found in perturbation theory is exactly the same as in invariant gauges. The arguments for singularity of the ghost Green function hold here as well. In this case, however, a singularity of the ghost Green function as \( k^2 \to 0 \) of the type \( 1/k^4 \) is indicative of a linear increase in the Coulomb interaction with distance.

The most natural way of formulating the Coulomb gauge is the Hamiltonian form which shows up vividly the unitarity of the theory because of the lack of ghosts. To this end, the functional integral \( W \) incorporates the fields which satisfy the three-dimensional transversality condition

\[ \frac{\partial A_i}{\partial x_i} = 0, \quad (49) \]

and momenta \( \pi_i \) which are canonically conjugated with them and stand for the transverse part of the electric field

\[ \pi_i = E_i^t = \left( \frac{\partial A_i}{\partial t} - [\nabla_i A_0] \right) \perp. \quad (50) \]
The integral over $A_0$ is calculated for fixed $A_i$ and cancels the Faddeev-Popov determinant.

As a result, the functional integral takes the form

$$W = \int \exp \left[ \frac{i}{g^2} \int d^4x \left\{ \pi_i \dot{A}_i - \mathcal{H}(\pi_i, A_i) \right\} \right] dA d\pi,$$

(51)

$$\mathcal{H}(\pi_i, A_i) = -\frac{i}{2} \left\{ \pi_i^2 + H^2(A_i) + \partial_i \varphi \partial_i \varphi \right\},$$

$$H^2 = \frac{1}{2} F_{ij} E_{ij},$$

(52)

$$\tilde{\Delta}(A) \varphi = -\rho, \quad \tilde{\Delta}(A) \varphi \equiv [\nabla_i(A), \partial_i \varphi], \quad \rho = [A_i, \pi_i].$$

(53)

In this case, the integration should be performed over gauge-invariant transverse fields, as for invariant gauges in (51).

The number of fields equivalent to a given field $A_i$ is determined by the number of solutions of the equation similar to eq. (13)

$$[\nabla_i(A), S \partial_i S^+] = 0.$$

(54)

The condition for the existence of two equivalent fields is the existence of a solution of the equation

$$[\nabla_i(A), \partial_\rho \varphi] = 0,$$

(55)

a zero eigenvalue of the operator $\tilde{\Delta}(A)$ defining the Coulomb potential according to eq. (53). Repeating the arguments given above for the four-dimensional case (see also the appendix), we draw a conclusion that the integration range in (51) should be restricted to the region $C_0$ where the operator $\tilde{\Delta}(A)$ has no eigenvalues.

This region coincides with the one where the Coulomb energy density

$$V_{\text{Coul}} = \frac{1}{2} \rho = \frac{1}{\Delta} \tilde{\Delta} \frac{1}{\Delta} \rho$$

(56)

does not go to infinity anywhere except for its boundary. In the Coulomb gauge, instead of the functional integral we can use the Schrödinger equation

$$\frac{1}{g^2} \int d^4x \mathcal{H} \left( -ig \frac{\partial}{\partial A_i}, A_i \right) \psi(A_i) = E \psi(A_i),$$

(57)

bearing in mind that $\mathcal{H}$ is determined only for the region $C_0$.

Let us discuss now how a limitation by the fields within $C_0$ affects the spectrum and zero oscillations defined by eq. (57). As before, for the no-level condition in the field $A_i$ we set

$$\frac{4}{3} \int \frac{d^3k}{(2\pi)^3} \sum_{\lambda, q} \frac{|A^{\lambda, q}(k)|^2}{k^2} < 1,$$

(58)
obtained in a similar manner from the Green function. Omitting in $\mathcal{H}(\pi, A)$ all terms except for $\mathcal{H}_0 = -\frac{1}{2}(\pi_i^2 - A_i \partial^2 A_i)$, we obtain, instead of (57), an equation for an oscillator system

$$\sum_{k, \lambda, q} \{ \pi_{\lambda, q}^2(k) + k^2 a_{\lambda, q}^2(k) \} \psi(a) = E \psi(a) ,$$  

(59)

provided that

$$\frac{4g^2}{3V} \sum_{k, \lambda, q} \frac{|a_{\lambda, q}(k)|^2}{k^2} < 1 .$$  

(60)

Taking no account of (60), for free oscillations $|a_{\lambda, q}(k)|^2 \sim 1/k$ and hence the left-hand side of (60) is infinite. This means that free oscillations correspond to the fields far outside the region $C_0$. Eqs. (59), (60) can be solved approximately by the variational method assuming that

$$\psi = \prod_{k, \lambda, q} f(a_{\lambda, q}(k)) .$$  

(61)

Calculating the energy minimum for the system with the fixed value of the left-hand side in (60), we get an oscillator equation for $f$ with $k^2 + k^4/k^2$ instead of $k^2$, where $k^4$ is the variational parameter. In this case, the ground-state energy and average squares of oscillation amplitudes will be

$$E = \frac{1}{2} \sum_{k, \lambda, q} \sqrt{k^2 + \frac{k^4}{k^2}} ,$$  

$$|a_{\lambda, q}(k)|^2 = \frac{k}{\sqrt{k^4 + k^4}} .$$  

(62)

The energy is at a minimum with $\kappa = 0$, but as this takes place, the left-hand side in (60) equals infinity. Therefore, $\kappa$ is determined from the condition

$$\frac{8}{3} g^2 \int \frac{d^3 k}{(2\pi)^3} \frac{1}{k^4} \approx \frac{2g^2}{3\pi^2} \ln \frac{\Lambda^2}{k^2} = 1 .$$  

(63)

The ghost Green function is

$$G(k) = \langle k | \frac{1}{\Delta} | k \rangle$$

$$= \frac{1}{k^2 4g^2 n \left\{ \int \frac{d^3 k'}{(2\pi)^3} \frac{(k^2 - 2kk')(1 - (kk')^2/k^2 k'^2)}{k'(k' - k)^2 \sqrt{k'^4 + k^4}} \right\}^{-1} ;$$  

(64)

as $k^2 \to 0$

$$G(k) = \frac{6}{5g^2 nk^4 \ln(k^2/k^2)} .$$  

(65)
Thus, this crude calculation shows that again there is a characteristic scale beyond which zero-point oscillations become small and the Coulomb potential increases linearly with distance.

In conclusion, it should be noted that since zero-point oscillations of the fields in vacuum turn out to be on the boundary of the region $C_0$, we have no right to ignore the Coulomb energy which may go to infinity on this boundary. However, considering that the number of zero oscillation modes pushing the system outside the region $C_0$ is infinitely large and that the boundary equation comprises one condition (60), we have good cause to think that the system will remain close to the boundary despite the Coulomb interaction. In this case, the condition (60) will be satisfied all the same, which is indicative of a decrease in the zero oscillation amplitudes for momenta below a particular value and of a linear increase in the Coulomb potential with distance.

Finally, I wish to express my sincere gratitude to A.A. Belavin, A.M. Polyakov, L.N. Lipatov and Yu.L. Dokshitser for numerous most helpful discussions.

Appendix

In this appendix we consider the properties of gauge-equivalent fields with equal divergence, giving the simplest examples. Let us begin with the case of the three-dimensional space (Coulomb gauge) and the group SU(2). We consider "spherically symmetric" fields $A_i$, i.e., the fields dependent on one unit vector $n_i = x_i/r$ ($r = \sqrt{x^2}$). The general expression for such a field has the form

$$A_i(x) = f_1(r) \frac{\partial \hat{n}}{\partial x_i} + f_2(r) \hat{n} \frac{\partial \hat{n}}{\partial x_i} + f_3(r) \hat{n} n_i,$$  \hspace{1cm} (A.1)

$\hat{n} = i n_a \sigma_a$, $\sigma_a$, are Pauli matrices, $\hat{n}^2 = -1$. Under the spherically symmetric gauge transformation of the form

$$S = \exp\left\{ \frac{1}{2} \alpha(r) \hat{n} \right\} = \cos(\frac{1}{2} \alpha) + \frac{\hat{n} \sin(\frac{1}{2} \alpha)}{\hat{n}}$$  \hspace{1cm} (A.2)

$A_i(x)$ goes into $\tilde{A}_i(x) = S^* A_i S + S^* \partial_i S$ so that

$$f_1' = f_1 \cos \alpha + (f_2 + \frac{1}{2}) \sin \alpha,$$

$$f_2' + \frac{1}{2} = -f_1 \sin \alpha + (f_2 + \frac{1}{2}) \cos \alpha,$$

$$f_3' = f_3 + \frac{\alpha'}{2},$$  \hspace{1cm} (A.3)

$$\partial A_i/\partial x_i = \hat{n} [f_2' + (2/r) f_3 - 2 f_1'/r^2].$$  \hspace{1cm} (A.4)

The condition

$$\partial \tilde{A}_i/\partial x_i = \partial A_i/\partial x_i$$  \hspace{1cm} (A.5)
is equivalent to the equation
\[ \alpha'' + \frac{(2/r) \alpha'}{r^2} \{ (f_2 + \frac{1}{2}) \sin \alpha + f_1 (\cos \alpha - 1) \} = 0. \] (A.6)

If we introduce the variable \( \tau = \ln r \), eq. (A.6) reduces to the equation for a pendulum with damping in the field of the vertical force \( 4f_2 + 2 \), horizontal force \( 4f_1 \), and the force perpendicular to the pendulum, \(-4f_1\)

\[ \ddot{\alpha} + \dot{\alpha} - (2 + 4f_2) \sin \alpha + 4f_1 (1 - \cos \alpha) = 0. \] (A.7)

From this analogy the general properties of solutions to the equations of the equivalence conditions (A.5), (A.6) are readily seen.

If the forces \( f_1 \) and \( f_2 \) are equal to zero as \( r \to 0 \) (\( \tau \to -\infty \); otherwise the field \( A_1 \) is singular as \( r \to 0 \)) and tend to zero as \( r \to \infty \) (\( \tau \to +\infty \)), then for a solution to exist at finite \( \tau \), as \( r \to -\infty \), the pendulum should be in the position of unstable equilibrium, \( \alpha = 0 \). In such an event, if its initial velocity as \( \tau \to -\infty \) is not specifically selected, upon executing a number of oscillations in the field, the pendulum starts damping and once only the vertical force remains, it comes to stable equilibrium. Such a solution corresponds to \( \mathcal{S} \to \hat{n} \), as \( r \to \infty \), and the equivalent field

\[ A_i = -\frac{\delta \hat{n}}{\delta x_i} \sim \frac{1}{r}. \] (A.8)

decreases slowly at infinity.

However, exceptional cases are possible. If sufficiently large, these forces can under specifically selected initial conditions restore the pendulum to its unstable equilibrium position. In this case, we obtain the equivalent field \( \vec{A}_i \) which decreases fairly rapidly at infinity. We consider several versions of such a possibility. Let the forces and initial conditions be such that throughout the whole \( -\infty < \tau < +\infty \), \( |\alpha(\tau)| \ll 1 \). Then, to a zero approximation the equation

\[ \ddot{\alpha} + \dot{\alpha} - 2\alpha (1 + 2f_2) = 0 \] (A.9)

should be satisfied. This equation is simply eq. (20) for the three-dimensional case. In order for eq. (A.9) to hold, \( -\int f_2 \, d\tau \) should have a particular and sufficiently large value. The field with a corresponding \( f_2 \) lies on the curve \( \xi_1 \) independently of
the values for $f_1$ and $f_3$. Taking into account the quadratic term in (A.7) enables us to get a solution with a somewhat larger or smaller $f_2$ (within or outside the region $C_0$) and to determine the amplitude of $\alpha$ as demonstrated in the text.

If $f_1 = 0$, there are no second-order terms and a solution exists only for $f_2$ larger than the value required for a zero energy level. The amplitude of $\alpha$ is determined by taking into account third-order terms. It is obvious that in this case there are two solutions that differ from one another by the sign of $\alpha$. It can easily be shown that they are in the region $C_0$. This situation corresponds to the phenomenon of intersection of the equivalent field lines discussed in the text and illustrated in fig. 3.

The field for which $f_1 = 0$ lies on the intersection of such lines. The occurrence of two equivalent fields in $C_0$ in this example does not point to the necessity of introducing $N(A)$ in the region $C_0$ because it shows the symmetry of the problem with respect to reflection. It is easily shown that the fields $f_1, f_3, f_2$ and $-f_1, -f_3, -f_2$ are gauge-equivalent and have equal divergence. Hence the functional integral should be divided by two, independently of the magnitudes of the fields $f_1, f_3, f_2$, which do not matter. Accordingly, eq. (A.6) always has two solutions $\alpha_1 = \alpha(f_1, f_2)$ and $\alpha_2 = -\alpha(-f_1, f_2)$. The second solution in the field $-f_1, -f_3, f_2, \alpha_2(-f_1, f_2)$, leads to an equivalent, reflected field obtained from $f_1, f_2, f_3$ using the solution $\alpha_1(f_1, f_2)$. With $f_1 = 0$, both reflected fields are obtained from one initial field. Two other solutions $\alpha_1(-f_1, f_2)$ and $\alpha_2(f_1, f_3)$ are outside $C_0$ anyway, for small $f_1$ and $f_2$ close to $\varphi_1$.

Another possibility for the pendulum to regain its unstable equilibrium position is to execute one complete revolution (or more). This possibility may exist even with $f_2 = 0$, i.e., deep inside the region $C_0$. Such solutions exist at large enough $f_1$, but the equivalent fields corresponding to them differ greatly from the initial field and are likely to lie outside $C_0$. For example, with $f_2 = 0, f_2$ is of the order of unity according to (A.3) because $\alpha$ in a complete revolution changes from 0 to $2\pi$.

Hence, the fields inside $C_0$ have their equivalents of two types, i.e. the fields which possess asymptotics at $\infty$ of the type (A.8) ($\alpha \to \pi$) and lying within the region $C_\infty$ (it is easy to see that with $f_2 = 1$ as $r \to \infty$, (A.9) has an infinite number of solutions) and the fields situated in $C_n$ with a finite $n$.

Finally, let us discuss the question as to whether for a particular field $A_i$ an equivalent field $A_i'$ with a specified difference $\delta A_i$ in their divergence can always be found. The equation for a corresponding $\alpha(\rho)$ will differ from (A.7) in the external force $2\Delta f e^{2\rho}$ on the right-hand side perpendicular to the pendulum. In this case, it is likely that there exists, "almost" without exception, a solution with $\alpha$ tending to $2n\pi$ as $r \to \infty$ because, as we have seen, if $f_1$ and $f_2$ are large, the solution comes into play through choosing the initial conditions; should $f_1, f_2$ and $\Delta f$ be small, we have an inhomogeneous linear equation for which the choice is made in a trivial way.

We now turn to the four-dimensional space. In this case, it is convenient to deal with the group O(4) from which SU(2) is trivially separated.

Instead of $\Gamma_{ij}$ as antihermitian matrices for infinitesimal transformations in the group O(4), one may choose $\sigma_{\mu\nu} = \frac{i}{2} (\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu)$. For constructing a scalar, we
need an antisymmetric tensor, i.e. at least two vectors are required. This indicates that the field cannot be spherically symmetric. It can be axially symmetric if we choose as antisymmetric tensor

$$F_{\mu\nu} = \frac{n_\mu l_\nu - n_\nu l_\mu}{\sqrt{1 - (n_\alpha l_\alpha)^2}},$$

(A.10)

where \(n_\mu = x_\mu/\sqrt{x^2}\) and \(l_\mu\) is a constant unit vector. The gauge transformation matrix between such axially symmetric fields can be written as

$$S = \exp \left\{ \frac{i}{2} n_l \psi \right\} = \cos \frac{1}{2} \beta + \psi \sin \frac{1}{2} \beta,$$

$$\psi = \frac{i}{2} \partial_\mu F_{\mu\nu}, \quad n_l = l_\mu n_\mu, \quad r = \sqrt{x^2_\mu}.$$  (A.11)

The field \(A_\mu\) which preserves its shape under this transformation has the form

$$A_\mu = f_1 \partial_\mu \psi + f_2 \psi \partial_\mu \psi + \psi \partial_\mu f_3.$$  (A.12)

The transformation formulae between \(f'_1\) and \(f'_2\) coincide with (A.3), if \(\alpha\) is replaced by \(\beta\).

The equivalence condition is

$$\frac{\partial^2 \beta}{r^2(1 - n_l)} \left( f_2 + \frac{1}{2} \right) \sin \beta - f_1 (1 - \cos \beta) = 0.$$  (A.13)

With \(f'_1 = f'_2 = 0\), there is a solution similar to (A.6) which is dependent on one variable \(\rho^2 = r^2(1 - n_l^2)\) and has the same asymptotics. Despite two variables, which make this equation more cumbersome, its structure is much the same as that of (A.6), and we do not see any reasons why the structure of its solutions should differ markedly from (A.6).

References