ISSN 1063-7761, Journal of Experimental and Theoretical Physics, 2011, Vol. 113, No. 4, pp. 619–627. © Pleiades Publishing, Inc., 2011.
Original Russian Text © I.M. Suslov, 2011, published in Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki, 2011, Vol. 140, No. 4, pp. 712–721.

On Wilson's Theory of Confinement[1]

I. M. Suslov
Kapitza Institute for Physical Problems, Moscow, 119337 Russia
e-mail: suslov@kapitza.ras.ru
Received February 22, 2011

Abstract—According to recent results, the Gell-Mann–Low function $\beta(g)$ of four-dimensional ϕ^4 theory is nonalternating and has a linear asymptotics at infinity. According to the Bogoliubov and Shirkov classification, it means the possibility of constructing a continuous theory with finite interaction at large distances. This conclusion is in visible contradiction to the lattice results indicating triviality of ϕ^4 theory. This contradiction is resolved by a special character of renormalizability in ϕ^4 theory: to obtain the continuous renormalized theory, there is no need to eliminate a lattice from the bare theory. In fact, such kind of renormalizability is not accidental and can be understood in the framework of Wilson's many-parameter renormalization group. Application of these ideas to QCD shows that Wilson's theory of confinement is not purely illustrative, but has a direct relation to a real situation. As a result, the problem of analytical proof of confinement and a mass gap can be considered solved, at least on the physical level of rigor.

DOI: 10.1134/S106377611109010X

1. INTRODUCTION

Recent investigations of the strong coupling regime in ϕ^4 theory revealed an unexpected feature in its renormalizability: the continual limit in the renormalized theory does not require the continual limit in the bare theory. We show below that such a kind of renormalizability has a general character and can be understood in the framework of Wilson's many-parameter renormalization group. These results make it possible to give a final solution to the problem of triviality or nontriviality of ϕ^4 theory. Application of these ideas to Wilson's theory of confinement shows that this theory is not purely illustrative, but has a direct relation to real QCD. As a result, the problem of analytical proof of confinement and a mass gap can be considered solved, at least on the physical level of rigor.

2. CHARACTER OF RENORMALIZABILITY IN ϕ^4 THEORY

According to recent results [1–4] (see also [5, 6]), the Gell-Mann–Low function $\beta(g)$ in four-dimensional ϕ^4 theory is nonalternating and has an asymptotic behavior $\beta(g) = 4g$ at $g \longrightarrow \infty$. According to the Bogoliubov and Shirkov classification [7] (see discussion in [3]), it means the possibility of constructing a continuous theory with finite interaction at large distances. This conclusion is in visible contradiction to the lattice results indicating triviality of ϕ^4 theory (see [8–12] and numerous references in [13]).

In fact, we should differentiate two definitions of triviality. According to Wilson [8], triviality means that integration of the Gell-Mann–Low equation

$$-\frac{dg}{d\ln L} = \beta(g) \tag{1}$$

in the direction of large distances L gives an effective charge g tending to zero (Fig. 1a); this definition implies massless theory, since in the opposite case the distance scale is saturated by the inverse mass. The definition of true triviality is different (Fig. 1b). In this case we consider the massive theory and suggest the finite interaction g_∞ for $L \gtrsim m^{-1}$; a theory is trivial if integration of the Gell-Mann–Low equation in the direction of small L gives a divergency at finite L_0 (the so-called Landau pole) and does not allow the $L \longrightarrow 0$ limit to be reached. Such a situation is internally inconsistent [7] and signifies incorrectness of the initial suggestion on finite interaction at large distances; in fact, $L_0 \longrightarrow 0$ if $g_\infty \longrightarrow 0$. Wilson triviality means that the β-function is nonnegative and has a zero only for $g = 0$. True triviality needs in addition its sufficiently quick growth at infinity, $\beta(g) \propto g^\alpha$ with $\alpha > 1$. According to [1–6], ϕ^4 theory and QED are trivial in the Wilson sense, but do not possess true triviality.

Two definitions of triviality were hopelessly mixed in literature [13]. The reasons for it are as follows:

(a) Bogoliubov and Shirkov's work is poorly known to the Western community.

[1] The article was translated by the author.

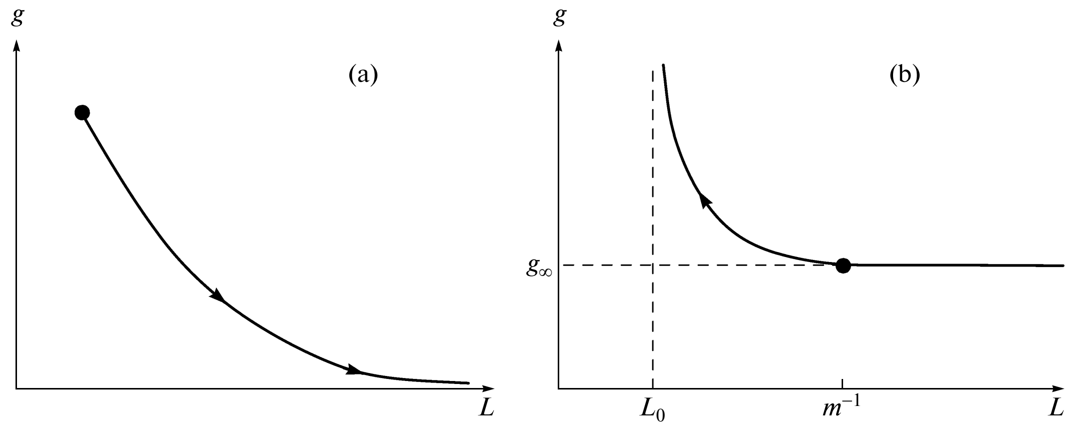

Fig. 1. Wilson triviality (a) in comparison to true triviality (b).

(b) It is rather difficult to test true triviality in the lattice approach.[2]

(c) There exist arguments that "prove" the equivalence of two definitions.

As an illustration to the latter point, consider the following reasoning. The only alternative to the perturbative approach is to express all quantities related to renormalized theory in terms of functional integrals. The latter depend on the bare charge g_0, bare mass m_0, and the ultraviolet cut-off Λ. Taking into account their dimensional character, we have the following relations for the renormalized charge g, renormalized mass m, and observable quantities A_i:

$$g = F_g(g_0, m_0/\Lambda), \quad m = \Lambda F_m(g_0, m_0/\Lambda), \tag{2}$$

$$A_i = \Lambda^{d_i} F_i(g_0, m_0/\Lambda), \tag{3}$$

where d_i is a physical dimensionality of A_i. Excluding g_0 and m_0/Λ in favor of g and m/Λ, we have

$$A_i = m^{d_i} \tilde{F}_i(g, m/\Lambda). \tag{4}$$

To eliminate the dependence on Λ we should take the limit $m/\Lambda \longrightarrow 0$. In the lattice approach, this limit corresponds to $\xi/a \longrightarrow \infty$ (ξ is the correlation length and a is lattice spacing), i.e., to the phase transition point. The latter is determined by a zero of the β-function, which gives $g = 0$ in four-dimensional ϕ^4 theory.

In this argumentation, Wilson triviality was considered as given, while true triviality was "derived" from it. Of course, it cannot be correct, because two definitions are surely not equivalent. This shortcoming originates from our assumption that a general-position situation takes place in Eq. (4): in this case we indeed should take the limit $m \longrightarrow \Lambda$. This limit is unnecessary if dependence on m/Λ is absent in Eq. (4). Such a special case fills the "gap" between the two definitions and renders the inequivalent.

Such a special case actually holds in ϕ^4 theory [3, 4]. Let us return to Eqs. (3) and impose the condition $m \ll \Lambda$, corresponding to the continuum limit of the renormalized theory. If this condition is imposed in the region $g_0 \gg 1$, then ϕ^4 theory reduces to the Ising model, containing the single parameter κ, which plays the role of inverse temperature [3, 4]; relations (3) take the form

$$\begin{gathered} g = F_g(\kappa), \quad m = \Lambda F_m(\kappa), \\ A_i = \Lambda^{d_i} F_i(\kappa). \end{gathered} \tag{5}$$

So far there is nothing unusual: the condition $m/\Lambda \longrightarrow 0$ gives a relation between g_0 and m_0/Λ, so all functions in Eq. (3) depend on the single parameter, which we denoted as κ. The nontrivial point consists in the following: the condition $m \ll \Lambda$ is sufficient for transformation to the Ising model, but not necessary for it. In fact, such transformation is possible under the weaker conditions, which are compatible with the arbitrary value of m/Λ [3, 4]. Excluding κ from Eqs. (5), we obtain the equations

$$A_i = m^{d_i} F_i(g), \tag{6}$$

which are analogous to (4), but do not contain the parameter m/Λ. As a result, the program of renormalization is completely fulfilled and no additional limiting transitions are necessary. It means that (a) we can

[2] A definition of true triviality in the lattice approach was given in mathematical papers [9, 10]. When the lattice spacing a tends to zero, the bare parameters g_0 and m_0 should be considered as functions of a. A theory is nontrivial if there exists some choice of functions $g_0(a)$ and $m_0(a)$ providing finite interaction at large distances; if such functions do not exist, then a theory is trivial. Of course, it is rather difficult to test "existence" or "nonexistence" in numerical simulations.

retain the lattice in the bare theory (as a convenient tool for representation of functional integrals) and (b) the relation between m and Λ (or ξ and a) can be arbitrary, so the arbitrary value of g becomes possible.

Usually, the lattice theory contains more parameters than the initial field theory. For example, discretization of the gradient term in d-dimensional ϕ^4 theory

$$\int d^d x[\nabla\phi(x)]^2 = -\int d^d x\phi(x)\nabla^2\phi(x) \longrightarrow \sum_{x, x'} J_{x-x'}\phi_x\phi_{x'}, \tag{7}$$

corresponds to the replacement of $-\nabla^2 = \hat{p}^2$ by $\epsilon(\hat{p})$, where $\epsilon(p)$ is a bare lattice spectrum

$$\epsilon(p) = \sum_x J_x e^{ipx} = p^2 + O(p^4), \tag{8}$$

while $\hat{p}$ is the momentum operator and $\exp\{i\hat{p}x\}$ is the operator of shift on the vector x. The overlap integrals J_x can be taken as arbitrary and are restricted only by the condition (8). The interesting question arises: if we can retain a lattice in the bare theory, then what lattice model should be chosen?

A solution can be found from Eq. (4). Since the dependence on m/Λ is absent, we can take $m/\Lambda \longrightarrow 0$. However, in this limit (when $\xi/a \longrightarrow \infty$) there are physical grounds for independence of functions F_i on the way of cut-off. If such independence takes place for $m/\Lambda \longrightarrow 0$, it is retained for arbitrary m/Λ due to independence of functions F_i on this parameter. In fact, this argumentation implies renormalizability of the theory (due to which the dependence on Λ can be excluded) and belonging of the lattice model to the proper universality class (inside which the dependence on the way of cut-off is absent).

The lattice theory is frequently considered a reasonable approximation to the true field theory. In this case we should accept the condition $\xi \gg a$, which signifies that there are a lot of lattice sites on the characteristic scale of variation of field. This condition can be stregthened up to $\xi/a \longrightarrow \infty$ or liberalized up to $\xi \gtrsim a$. The first case corresponds to the point of the phase transition and gives $g = 0$. In the second case, we obtain restriction $g \lesssim 1$ (for the proper charge normalization [4]), which can be used to obtain the upper bound on the Higgs mass [12, 14].

In fact, the lattice theory should not be considered as any approximation to field theory, though it is possible for $g_0 \longrightarrow 1$. The true field theory is continuous from the very beginning and does not contain any lattice. The lattice is present only in the bare theory, which is an auxiliary construction and is completely removed later. No physical requirements, like $\xi \gg a$, are relevant for it. If one removes the condition $\xi \gg a$, then any values of g become admissible.[3] In fact, a real designation of the bare theory is to represent the relations between physical quantities in the parametric form (3). Such representation has no deep sense already due to its ambiguity: it can be written in many different forms, replacing g_0 and m_0/Λ by any other pair of variables.

We see that contradiction between the continual and lattice approaches is resolved by a special character of renormalizability in ϕ^4 theory:

Correct relations (6) between physical quantities can be obtained for the arbitrary value of the parameter a/ξ, while the dependence on this parameter is absent. To obtain the continuous renormalized theory, there is no need to eliminate a lattice from the bare theory.

3. GENERAL SITUATION IN RENORMALIZABLE THEORIES

The interesting question arises: Is such kind of renormalizability related with the specific properties of ϕ^4 theory, or it is a manifestation of some general mechanism?

We shall see below that the second variant is correct. It can be understood in the framework of Wilson's many-parameter renormalization group (RG) [8]. According to it, the parameters p_i of some lattice Hamiltonian are considered as functions of the length scale l.[4] The flow of these parameters is determined by the RG equations, which can be written in the differential form

$$-\frac{dp_i}{d\ln(l/a)} = F_i\{p_k\}. \tag{9}$$

These equations can be linearized near the fixed point

$$p_i(l) = p_i^* \quad (\text{for all } l) \tag{10}$$

and investigated by the standard methods of linear algebra. The ordinary phase transitions are described by the saddle points of such equations. The simplest

[3] One can consider $\xi/a \gg 1$ as a technical condition providing a good approximation, but it is not actual due to the absence of ξ/a dependence (6). The stated point of view is in complete agreement with mathematical definitions [9, 10], according to which the limit $a \longrightarrow 0$ is taken for the arbitrarily chosen dependences $g_0(a)$ and $m_0(a)$ (see Footnote 2). We impose conditions $g_0 \longrightarrow \infty$, $g_0^{-1/2}m_0^2a^2 \longrightarrow -\infty$, $g_0^{-1}m_0^2a^2 = -\kappa$, necessary for transformation to the Ising model [3, 4].

[4] Physically it is explained by the well-known Kadanoff construction. In the description of magnetics, one begins with the microscopic Hamiltonian for elementary spins in the lattice sites. Then it is possible to introduce the macroscopic spin variables corresponding to the blocks of size l and write the effective exchange Hamiltonian for them. Since the blocks of size nl can be composed of n^d blocks of size l, then recalculation $p_i(l) \longrightarrow p_i(nl)$ is possible, i.e., $p_i(nl) = H_i(n, \{p_k(l)\})$. Taking n close to unity, we can obtain Eqs. (9).

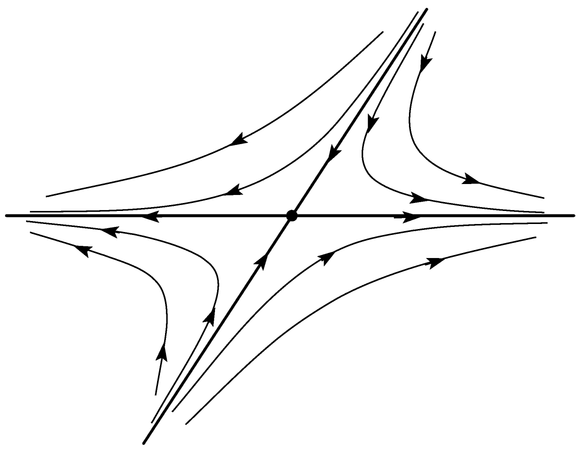

Fig. 2. Simplest variant of the saddle point.

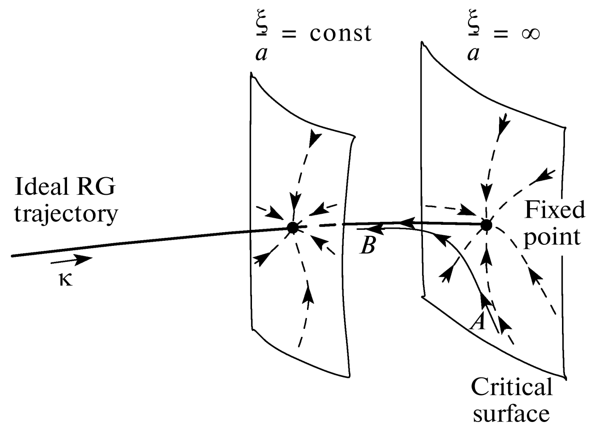

Fig. 3. Schematic of the Wilson many-parameter space.

saddle point in the two parameter space (Fig. 2) has straight-line trajectories in two main directions (one stable and one unstable), while the rest of the trajectories are hyperbolic. For the usual phase transitions, there are an infinite number of stable directions and one (in the simplest case) unstable direction. The latter is related to some controlling parameter like temperature, measuring the distance to the critical point.

Instead of increasing l for fixed a, we can decrease a for fixed l. The continuum limit $a \longrightarrow 0$ of field theory corresponds to the critical surface $\xi/a = \infty$ in the many-parameter space (Fig. 3). All trajectories at the critical surface tend to the fixed point. The unstable trajectory originating at the fixed point will be referred as an "ideal RG trajectory": along it one has exact one-parameter scaling, which is a pipe dream in many fields of physics (see, e.g., [15]). To define it rigorously, let us consider the limit $a \longrightarrow 0$ with fixed ξ/a; then all trajectories lying at the surface ξ/a = const (Fig. 3) tend to one point (analogously to the critical surface), while the locus of such points is the ideal RG trajectory.

Let parameter κ measure the distance along the ideal trajectory: then ξ/a (or Λ/m) is a function of κ. Analogously, all dimensionless quantities depend only on κ, while the dimensional quantities are measured in units of Λ. As a result, we come to the equations

$$g = F_g(\kappa), \quad m = \Lambda F_m(\kappa), \quad A_i = \Lambda^{d_i} F_i(\kappa), \quad (11)$$

which coincide with (5) and give the relations (6) with no dependence on m/Λ.

The above construction has the following sense. If the limit $a \longrightarrow 0$ is taken in the arbitrary manner, then the system will go to infinity along the unstable direction and appear far from the critical surface, which is our goal. Therefore, we suggest taking the continual limit in two steps:

(a) take a limit $a \longrightarrow 0$ for a/ξ = const;

(b) take a limit $a/\xi \longrightarrow 0$.

It appears that the dependence on a in renormalized theory disappears already at the first step. The second step becomes unnecessary and we need not take the continuum limit in the bare theory. The const appearing in (a) is one of the possible definitions of parameter κ.

These ideas are close to those of QCD specialists, and in fact the above consideration was partially taken from "Introduction to Lattice QCD" by R. Gupta [16]. This picture is discussed there in relation to improvement in the lattice action, and the author claims that simulations performed along the ideal RG trajectory will reproduce the continuum physics without discretization errors. It implies the absence of the a/ξ dependence, in accordance with our results. Only a final conclusion was not made, that the continuum limit is not necessary in the bare theory. In fact, this conclusion goes against the present-day practice in lattice simulations, which are made in the region of large ξ/a (typically $\xi/a = 5$–15) with accurate extrapolation to $a/\xi \longrightarrow 0$.

Any RG trajectory is a line of "constant physics," since the RG transformation is simply a mental construction, which does not affect the large-scale properties of the system. All trajectories belonging to the critical surface and meeting at the fixed point are physically equivalent, corresponding to the unique continuous field theory. The ideal RG trajectory originating at the fixed point gives the equivalent representation for field theory. Let us consider the trajectory AB, which begins near the critical surface, goes along it, and then tends to the ideal RG trajectory (Fig. 3). Introducing $\tilde{\kappa}$ as a distance along AB, we come to the parametric representation analogous to (11)

$$g = \tilde{F}_g(\tilde{\kappa}), \quad m = \Lambda \tilde{F}_m(\tilde{\kappa}), \quad A_i = \Lambda^{d_i} \tilde{F}_i(\tilde{\kappa}), \quad (12)$$

and relations (6), independent of ξ/a. The choice of small or large ξ/a values corresponds to the "ends" of trajectory AB which are arbitrarily close to the critical surface and the ideal trajectory; hence, the obtained relations (6) correspond to continual theory. However, the parametric representation (12) is essentially different from (11) and is not reduced to the change of variables $\kappa = f(\tilde{\kappa})$. To understand it, let us retain the definition of κ as a distance along the ideal trajectory and assign it to the point of AB, corresponding to the same value of ξ/a. Then the second relation (12) will be the same as (11), but the other two relations remain different:

$$g = \tilde{F}_g(\kappa), \quad m = \Lambda F_m(\kappa), \quad A_i = \Lambda^{d_i}\tilde{F}_i(\kappa). \quad (12')$$

Indeed, the charge g usually belongs to irrelevant parameters and we can introduce "the axis of charges" at the critical surface; the fixed point corresponds to $g = 0$. If the limit $a/\xi \longrightarrow 0$ is taken along the ideal trajectory, then $g \longrightarrow 0$. If this limit is taken along AB, then the arbitrary function $g = \tilde{F}_g(\kappa)$ in (12') is possible: it depends on the direction of AB relative to the axis of charges. The functional relation between g and κ becomes indeterminate and can be omitted.

As a result, the renormalized and bare sectors of theory become decoupled. The renormalized sector contains relations (6), where g and m are considered independent variables. The bare sector contains only relation $a/\xi = m/\Lambda = F_m(\kappa)$, which determines κ as a function of a and is irrelevant from the viewpoint of physics. Parameter a/ξ becomes absolutely free.

The set of different trajectories AB defines the universality class of the corresponding field theory. Such trajectories fill the whole space if the critical surface and the ideal RG trajectory are unbounded. In fact, the critical surface is certainly restricted in some directions, because there are a lot of such surfaces corresponding to different phase transitions. To obtain the correct relations (6), there is no need to construct the ideal RG trajectory: it suffices to find the arbitrary trajectory like AB, belonging to the same universality class.

As a result, we come to the following conclusion:

Renormalizable theory of the considered type allows representation in the form of lattice theory, which gives the correct relations between physical quantities and contains free parameter a/ξ, which does not enter these relations.

4. APPLICATION TO THEORY OF CONFINEMENT

QCD with one sort of quarks contains two parameters, interaction constant g and the quark mass m. Its renormalization properties are analogous to those of ϕ^4 theory or QED and are expressed by the relations (3), (4); in fact Section 3 is axiomatic for study of such theories. We restrict our discussion to a theory without quarks, i.e., the pure Yang–Mills theory; then the quark mass is not included as a parameter and the theory does not contain any natural mass scale. To avoid specific difficulties related to such a situation, we introduce the "extended version" of the Yang–Mills theory, where the role of the bare quark mass m_0 (more exactly, the ratio m_0/Λ) is played by some auxiliary parameter p characterizing the lattice theory; as a renormalized mass, we accept the mass m of the lightest glueball (the bound state of several gluons), while the correlation length ξ is defined as m^{-1}. Thereby, two bare parameters g_0 and p provide the observable values for renormalized g and m (g corresponds to the momentum scale m). In order to return to the standard variant of the theory, we should remove the introduced extra degree of freedom by fixing one relation between observable quantities. However, it can be done at the late stage (see the end of Section 4), while the main of analysis is produced for the "extended version," which is analogous to ϕ^4 theory.

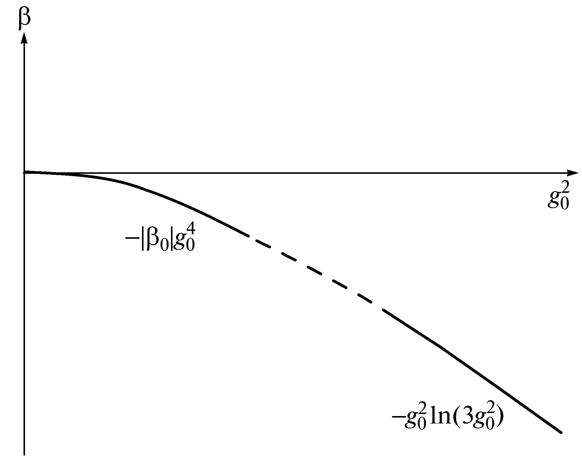

Fig. 4. Gell-Mann–Low function of the Yang–Mills theory in the cut-off scheme.

According to Wilson [17], confinement can be proved in the lattice version of the Yang–Mills theory for a large value of the bare charge g_0. The interaction energy for two probe quarks, separated by a distance R, is $V(R) = \sigma R$, while the string tension σ and the glueball mass m are given by expressions [16–18]

$$\sigma = \frac{\ln(3g_0^2)}{a^2}, \quad m = \frac{4\ln(3g_0^2)}{a}. \quad (13)$$

In spite of the evident success, the Wilson theory is considered purely illustrative and having no relation to real QCD. As was indicated by Wilson himself, his theory corresponds to the situation

$$\xi \ll a \quad \text{or} \quad m \gg \Lambda, \quad (14)$$

which is considered unphysical. An attempt to advance into the physical region inevitably destroys the strong coupling regime. Indeed, fixing σ to its observable value, we have $g_0^2(a) = (1/3)\exp(\sigma a^2)$ and substitution to the Gell-Mann–Low equation in the cut-off scheme [19]

$$-\frac{dg_0^2}{d\ln a^2} = \beta(g_0^2) = \beta_0 g_0^4 + \beta_1 g_0^6 + \ldots \qquad (15)$$

gives $\beta(g_0^2) = -g_0^2\ln(3g_0^2)$ for large g_0 [20]. Together with a negative sign of β_0 and β_1 it implies the negative β-function for all g_0 (Fig. 4); the lattice results confirm this conclusion (see also [21]).[5] In this case g_0 tends to zero in the continuum limit $a \longrightarrow 0$. It does not mean triviality of the theory, because the behavior of $g_0 \longrightarrow 0$ is compatible with the finite value of the renormalized charge g, as can be seen from the one-loop result for $L \longrightarrow \infty$.

$$g_0^2 = \frac{g^2}{1 + |\beta_0| g^2 \ln \Lambda^2/m^2} \longrightarrow \frac{1}{|\beta_0| \ln \Lambda^2/m^2} \longrightarrow 0. \qquad (16)$$

Triviality is avoided, but the strong coupling regime is inevitably destroyed and Wilson's theory becomes inapplicable. The standard point of view is based on the argument that the strong coupling and weak coupling regions are not separated by the phase transition point and the Wilson picture can be extended qualitatively to the region $\xi \gg a$. This argument is confirmed by numerical simulation but of course cannot be considered as a proof.

The situation changes drastically if we use representation (5), (6) introduced in the previous sections. In this case:

(1) We do not need to take the continual limit in the bare theory, so g_0 remains finite.

(2) Due to absence of the ξ/a dependence, this parameter can be taken as arbitrary: it eliminates objections against the unphysical regime in Wilson's theory.

(3) Experience of ϕ^4 theory shows that there is no direct relation between the bare and renormalized charge:[6] representation (5) is rigorously introduced in the limit $g_0 \longrightarrow \infty$ (see Footnote 3), while g remains to be a finite function of κ [4]. With some reservations, the same property is valid in the Yang–Mills theory. Rewriting the second expression (13) in the form

$$g_0^2 = \frac{1}{3}\exp\left(\frac{ma}{4}\right) = \frac{1}{3}\exp\left\{\frac{a}{4\xi}\right\}, \qquad (17)$$

we see that, independently of renormalized values of g and m, it is possible to choose the free parameter a/ξ so as to obtain a sufficiently large value for g_0. Then Wilson's theory becomes applicable and the first relation (13) gives a finite value for σ, i.e., confinement.[7]

We have used the relations (13), which are valid for the simplest Wilson action [16–18]. However, the latter is not suitable for our purposes due to a trivial fact that it does not contain the sufficient number of parameters. To obtain the observable values of σ and m

$$\sigma = a^{-2} f_\sigma(g_0), \quad m = a^{-1} f_m(g_0), \qquad (18)$$

one should fix both g_0 and a; but the fixed a means that it is impossible to introduce representation with free parameter a/ξ. Therefore, we should consider some generalizations.

The simplest Wilson action [16–18]

$$S = -\frac{1}{g_0^2}\sum_{\square} W_{\square}^{1\times 1} \qquad (19)$$

is the sum over all plaquettes $\square$ of size 1×1, where the plaquette contribution $W_{\square}^{1\times 1}$ is determined by a product of matrices attributed to the sides of a plaquette. In the contemporary investigations, more complicated forms of the action are used which contain contributions of $m \times n$ plaquettes [16]:

$$S = -\frac{1}{g_0^2}\sum_{\square}\sum_{m,n} C_{mn} W_{\square}^{m\times n}. \qquad (20)$$

The coefficients C_{mn} sufficiently quickly decrease with an increase in m and n.[8] If a contribution of the $n \times n$ plaquette is dominant in the sum, we obtain Eq. (13) with na instead a. It is clear that generally we will have the effective averaging of (13) over a in some finite lim-

[5] The difference between the Abelian and non-Abelian cases arises in this point. It is well-known that Wilson's proof of confinement is equally valid for Abelian theories like QED [16]. Therefore, the β-function in QED has the same strong coupling behavior as in QCD (see Fig. 4), but appears to be positive for small g_0. As a result, the existence of the fixed point is inevitable and the regions of small and large g_0 are separated by the phase transition point. The subsequent considerations are equally valid for QED but refer to its unphysical branch. Strictly speaking, we cannot prove that QCD does not have the unphysical branch without confinement, but in any case the variant with confinement is preferred on physical grounds.

[6] Usually it is accepted that g_0 coincides with the renormalized charge g taken at the scale Λ; it is valid only if $g \ll 1$ and $g_0 \ll 1$ simultaneously.

[7] As opposed to the standard point of view (see the text after Eq. (16)), we do not try to justify the smooth crossover between the strong coupling and weak coupling regions but establish crossover between the $a \ll \xi$ and $a \gg \xi$ regimes. The latter crossover is trivial due to the absence of dependence on a/ξ.

[8] To understand this point, let us return to Eq. (8). The exchange integrals J_x should fall with $|x|$ in an exponential manner, so that the bare spectrum $\epsilon(p)$ can be regularly expanded in p. Analogous arguments can be given for the Yang–Mills theory.

its from $a_{\min} = a$ till $a_{\max} = ka$. As a result, relations (13) will have the form

$$\sigma = \frac{\ln(3g_0^2)}{a_1^2}, \quad m = \frac{4\ln(3g_0^2)}{a_2}, \tag{21}$$

where $a_1 = k_1 a$, $a_2 = k_2 a$ simply by dimensional reasons. These modifications do not affect the qualitative conclusions made above.

Relation (6) for σ has the form

$$\sigma = m^2 F_\sigma(g), \tag{22}$$

so g is functionally related to σ/m^2. Equations (21) give

$$\frac{\sigma}{m^2} = \frac{a_2^2}{16a_1^2\ln(3g_0^2)}, \tag{23}$$

and for $a_1 \sim a_2$ the ratio σ/m^2 is small in the strong coupling region. It means that only a restricted range of g values can be reproduced. Such a restriction is natural due to the physical essence of the problem. Indeed, the linear confinement potential is expected only at large distances, where g is certainly not small; hence, small values of g are inaccessible in the Wilson regime. On the contrary, the restricted range of σ/m^2 values goes against the logic of the theory. Indeed, a/ξ is a free parameter and all physical results can be obtained (analytically or not) at its arbitrary value. In the case $a/\xi \gg 1$, the regime of confinement is controlled analytically and any physically accessible value of σ/m^2 should be possible in this limit. Probably, the range of σ/m^2 values can be extended if we use the models with essentially different a_1 and a_2.[9] Absence of restrictions on σ/m^2 in the presence of restrictions on g is possible only if σ/m^2 has a maximum as a function of g; fortunately, we can demonstrate that it is really the case.

Investigations of complicated lattice versions of the Yang–Mills theory [16] show the existence of phase transitions (lying in the region $g_0 \sim 1$), corresponding to vanishing of the lightest glueball mass m, with finite values of σ and other mass parameters. These transitions are considered as lattice artifacts, since they do not survive in the continuum limit $a \longrightarrow 0$, when $g_0 \longrightarrow 0$. In our approach the limit $a \longrightarrow 0$ is not necessary and such phase transitions acquire physical sense. Their existence means that the dependence $\sigma/m^2 = F_\sigma(g)$ is singular (Fig. 5) and provides accessibility of arbitrary σ/m^2 values, retaining the restriction on values of g.

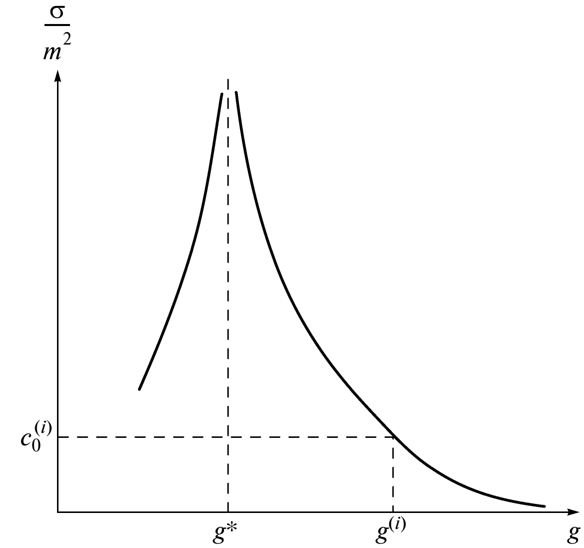

Fig. 5. Dependence of σ/m^2 against g. In order to obtain the special values $c_0^{(i)}$, corresponding to zero values of the mass gap, one should mark all stable fixed points $g^{(i)}$ on the horisontal axis and make a construction shown in the figure.

Existence of points with $m = 0$ in the parametric space means that the extended version of the Yang–Mills theory does not possess the mass gap. To eliminate this defect, we should return to the standard variant of the theory, fixing one relation between observable quantities. The character of such relations is well known and is determined by the so-called "dimensional transmutation" [16, Section 14.1], [22, Section IV.6]. If we have $A = a^\mu f(g_0)$ for the observable quantity A, then its independence of a means

$$\frac{dA}{da} = \mu a^{\mu-1} f(g_0) + a^\mu \frac{df(g_0)}{dg_0^2}\frac{dg_0^2}{da}$$
$$= a^{\mu-1}\left[\mu f(g_0) - 2\frac{df(g_0)}{dg_0^2}\beta(g_0^2)\right] = 0,$$

where Eq. (15) is taken into account. Integration of the obtained equation for $f(g_0)$ gives

$$A = \text{const}\, a^\mu \exp\left\{\frac{\mu}{2}B(g_0^2)\right\}, \quad B(g_0^2) = \int\frac{dg_0^2}{\beta(g_0^2)}$$

i.e. all quantities of the same dimensionality differ only by a constant factor, independent of g_0. For our purposes it is convenient to accept the condition

$$\sigma/m^2 = c, \tag{24}$$

which defines the one-parameter family of Yang–Mills theories with different values of the structural

[9] Such models certainly exist. If the contribution of plaquette $1 \times n$ dominates in the sum of (20), then usual tiling of the Wilson loop or correlational tube [16, 18] gives $a_1^2 = na^2$, $a_2 = na$, and the right-hand side of Eq. (23) is n times greater than for the Wilson action. To understand which values of σ/m^2 are actually accessible, it is necessary to investigate, does the strong coupling regime still correspond to condition $g_0 \gg 1$ or it is replaced by the more general n-dependent condition.

constant c.[10] Under condition (24), the points with $m = 0$, $\sigma = \text{const}$ become inaccessible.

It does not yet prove the existence of a mass gap, since σ and m can vanish simultaneously. In order to analyze such situations, consider the Gell-Mann–Low equation for the renormalized charge g attributed to the scale m:

$$\frac{dg^2}{d\ln m^2} = \beta(g^2) = \beta_0 g^4 + \beta_1 g^6 + \ldots, \quad (25)$$

where the β-function does not coincide with (15), but has the same first coefficients β_0 and β_1. It is clear that value g^* (Fig. 5) is a root of $\beta(g^2)$; generally, it has several roots determining the RG fixed points. In the limit $m \longrightarrow 0$, the charge g tends to one of these fixed points, while the following variants are possible for σ/m^2: (a) $\sigma/m^2 \longrightarrow \infty$, (b) $\sigma/m^2 \longrightarrow 0$, (c) $\sigma/m^2 \longrightarrow c_0$. The first two variants are incompatible with Eq. (24), while the third variant is possible in the case $c = c_0$. If there are several stable fixed points $g^{(i)}$, then there are several special values $c_0^{(i)}$ (see Fig. 5) for which the mass gap vanishes; for all other values of c the mass gap is finite.

Physically, it looks most probable[11] that only one fixed point g^* with $\sigma/m^2 \longrightarrow \infty$ is present, so no special values $c_0^{(i)}$ arise. Mathematically, an infinite number of fixed points can be suggested, which form a sequence $c_0^{(i)}$ everywhere dense in the interval $(0, \infty)$. However, small values of σ/m^2 correspond to the Wilson regime where finiteness of σ and m is verified immediately. As a result, the proof of the mass gap is complete for small values of the structural constant c.[12] The real perspective to strengthen this statement is outlined in Footnote 9.

It is worthwhile to indicate the papers [6, 23], which deal with β-functions close to (25). Paper [23] considers $\beta(g^2)$ defined in the MS scheme, where differentiation in (25) is performed over an arbitrary momentum scale μ; the behavior of $\beta(g^2) = \beta_\infty g^{2\alpha}$ with $\alpha \approx -13$ is obtained for large g, while the sign of β_∞ remains indefinite, so the existence of fixed point is one of the possible variants. An alternative definition of $\beta(g^2)$ can be obtained in QCD if g is attributed to the scale of the quark mass m; if g is defined through the quark-gluon vertex, then calculation of the asymptotics for the β-function can be performed in complete analogy to QED [6], giving the result $\beta(g^2) = g^2$ with necessary existence of a fixed point. We have seen above the existence of fixed point when the glueball mass m was making the scale. The listed definitions of $\beta(g^2)$ are technically different, but they physically correspond to the same dependence of renormalized charge on the length scale.[13] The physical sense of existence of the fixed point was clarified above.

If quarks with zero mass[14] are introduced, then the regime of dimensional transmutation is conserved and the trick with extension of theory remains actual; it seems, that the general structure of theory is also retained.

Final conclusions are as follows:

Whatever the properties of the continuous Yang–Mills theory, there exists a lattice theory that reproduces them. The bare charge g_0 in this lattice theory can be taken as arbitrary and, in particular, infinitely large. Any reasonable lattice version of the Yang–Mills theory gives finite values of a and m in the strong coupling limit. Vanishing of σ and m is possible for exceptional configurations in the many-parameter space, which are avoided in the general situation. As a result, the problem of analytical proof of confinement and the mass gap can be considered solved, at least at the physical level of rigor.

[10] The extended theory corresponds to the set of all "standard" theories with different c values.

[11] Calculation of β functions in different theories [5, 21] shows that they usually have the simple behavior interpolating between strong coupling and weak coupling regime.

[12] In fact, we have suggested that the extended Yang–Mills theory belongs to the type considered in Section 3. Motivation for this is as follows. The bare Yang-Mills theory contains the single parameter g_0, immediately related to the unstable direction. We can extend the theory along the stable directions in the many-parameter space; if there are unstable directions, we can artificially forbid extension along them. Indeed, additional essential parameters correspond to a theory that is more general than the Yang–Mills theory; such theories certainly exist, but they are not a subject of our consideration. We see that belonging of the extended Yang-Mills theory to the type considered in Section 3 can be accepted axiomatically.

[13] According to [24], existence of the root of the β-function is an invariant property, valid in all physical renormalization schemes.

[14] In the case of fermions, the renormalization of mass has a multiplicative character and the choice of the zero bare mass provides the zero renormalized mass.

REFERENCES

1. I. M. Suslov, JETP **93** (1), 1 (2001).
2. I. M. Suslov, JETP **107** (3), 413 (2008).
3. I. M. Suslov, JETP **111** (3), 450 (2010).
4. I. M. Suslov, JETP **112** (2), 274 (2011).
5. I. M. Suslov, JETP **100** (6), 1188 (2005).
6. I. M. Suslov, JETP **108** (6), 980 (2009).
7. N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Nauka, Moscow, 1976; Wiley, New York, 1980).
8. K. Wilson and J. Kogut, *Renormalization Group and the Epsilon Expansion* (Wiley, New York, 1974; Mir, Moscow, 1975).
9. J. Frölich, Nucl. Phys. B **200** [FS4] (2), 281 (1982).
10. M. Aizenman, Commun. Math. Phys. **86**, 1 (1982).
11. B. Freedman, P. Smolensky, and D. Weingarten, Phys. Lett. B **113**, 481 (1982).

12. M. Lüscher and P. Weisz, Nucl. Phys. B **290** [FS20] (1), 25 (1987); M. Lüscher and P. Weisz, Nucl. Phys. B **295** [FS21] (1), 65 (1988); M. Lüscher and P. Weisz, Nucl. Phys. B **318** (2), 705 (1989).
13. I. M. Suslov, arXiv:0806.0789.
14. R. F. Dashen and H. Neuberger, Phys. Rev. Lett. **50**, 1897 (1983).
15. E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishman, Phys. Rev. Lett. **42**, 673 (1979).
16. R. Gupta, arXiv:hep-lat/9807028.
17. K. G. Wilson, Phys. Rev. D: Part. Fields **10**, 2445 (1974).
18. M. Creutz, *Quarks, Gluons and Lattices* (Cambridge University Press, Cambridge, 1983; Mir, Moscow, 1987).
19. A. A. Vladimirov and D. V. Shirkov, Sov. Phys.—Usp. **22** (10), 860 (1979).
20. C. Callan, R. Dashen, and D. Gross, Phys. Rev. D: Part. Fields **20**, 3279 (1979).
21. J. B. Kogut, R. B. Pearson, and J. Shigemitsu, Phys. Rev. Lett. **43**, 484 (1979).
22. A. A. Slavnov and L. D. Faddeev, *Gauge Fields, Introduction to Quantum Theory* (Nauka, Moscow, 1988; Addison-Wesley, Boston, Massachusetts, United States, 1991).
23. I. M. Suslov, JETP Lett. **76** (6), 327 (2002).
24. I. M. Suslov, arXiv:hep-ph/0605115.