TTP08-12

Asymptotic freedom: history and interpretation*

Andrey Grozin
Institut für Theoretische Teilchenphysik, Universität Karlsruhe

Abstract

In this lecture, the early history of asymptotic freedom is discussed. The first completely correct derivation of β_0 in non-abelian gauge theory (Khriplovich, 1969) was done in the Coulomb gauge; this derivation is reproduced (in modernized terms) in Sect. 2. A qualitative physical explanation of asymptotic freedom via chromomagnetic properties of vacuum (Nielsen, 1981) is discussed in Sect. 3.

1 Introduction

Evolution of the QED coupling $\alpha(\mu)$ is determined by the renormalization group equation

$$\mu \frac{d\alpha(\mu)}{d\mu} = -2\beta(\alpha(\mu))\,\alpha(\mu)\,, \qquad \beta(\alpha) = \beta_0 \frac{\alpha}{4\pi} + \beta_1 \left(\frac{\alpha}{4\pi}\right)^2 + \cdots \tag{1.1}$$

where the 1-loop β-function coefficient is

$$\beta_0 = -\frac{4}{3} \tag{1.2}$$

(we use the $\overline{\mathrm{MS}}$ renormalization scheme). The sign $\beta_0 < 0$ corresponds to the natural picture of charge screening — when the distance grows (μ reduces), $\alpha(\mu)$ becomes smaller.

Non-abelian gauge theories have $\beta_0 > 0$. This corresponds to the opposite behaviour called *asymptotic freedom* — the coupling becomes small at large μ (small distances). Discovery of this fact came as a great surprise, and completely changed ideas about applicability of quantum field theory for describing the nature. In fact, it was done several times, by several researchers, independently.

The first paper in which $\beta_0 > 0$ was obtained is by Vanyashin, Terentev [1]. They considered non-abelian gauge theory with the gauge boson mass term introduced by hand.

*Lecture given at VI Dubna International Advanced School of Theoretical Physics, Dubna, 2008

Of course, this theory is non-renormalizable. It contains longitudinal gauge bosons whose interaction grows with energy. They obtained

$$\beta_0 = \left(\frac{11}{3} - \frac{1}{6} \right) C_A \,. \tag{1.3}$$

The contribution $-1/6$ comes from longitudinal gauge bosons. In the massless gauge theory it is canceled by the ghost loop (which was not known in 1965). Now we know that renormalizable massive gauge theory can be constructed using the Higgs mechanism. In this theory, the result (1.3) is correct; the contribution $-1/6$ comes from the Higgs loop.

The first completely correct derivation of β_0 in non-abelian gauge theory has been published by Khriplovich [2]. He used the Coulomb gauge which is ghost-free. The Ward identities in this gauge are simple (as in QED), and it is sufficient to renormalize the gluon propagator in order to renormalize the charge. This derivation clearly shows why screening is the natural behaviour, and how non-abelian gauge theories manage to violate these general arguments. It is also very simple. We shall discuss this derivation of β_0 in Sect. 2, following [2], but using a modernized language.

’t Hooft discovered asymptotic freedom in 1971–72 while studying renormalization of various field theories in dimensional regularization [3]. He reported his result during a question session after Symanzik’s talk at a small conference in Marseilles in June 1972.

And finally, Gross, Wilczek [4] and Politzer [5] discovered it again in 1973. They were first who suggested to apply asymptotically free gauge theory (QCD) to strong interactions, in particular, to deep inelastic scattering.

The early history of asymptotic freedom is discussed in several papers, e.g., [6, 7, 8], and in the Nobel lectures [9].

Now β_0 is derived in every quantum field theory textbook. In the standard approach, the covariant gauge is used (with Faddeev–Popov ghosts, of course). The coupling constant renormalization can be obtained from renormalizing any vertex in the theory together with all propagators attached to this vertex. Usually, the light-quark – gluon vertex or the ghost – gluon vertex is used (calculations are slightly shorter in the later case). The 3-gluon vertex or even the 4-gluon one [10] can also be used. All these standard derivations can be found, e.g., in [11]. The infinitely-heavy-quark – gluon vertex can also be used [12]; this derivation is, perhaps, as easy as the ghost – gluon one.

If the Ward identities are QED-like, it is sufficient to renormalize the gluon propagator; no vertex calculations are required. One such case is the Coulomb gauge [2]; it has a slight disadvantage of being not Lorentz invariant (this makes gluon loop calculations more difficult). Another example is the background field formalism [13]. It is Lorentz invariant, and provides, probably, the shortest derivation of β_0 in QCD.

However, all these derivations don’t explain asymptotic freedom (antiscreening) in simple physical terms. Probably, the simplest explanation of this kind is presented in [14] (it is also discussed in [8]). We shall discuss in in Sect. 3.

2 Coulomb gauge (Khriplovich 1969)

2.1 Feynman rules

This gauge has several advantages. It is ghost-free, and Ward identities in it are simple. It also has one significant disadvantage: it is not Lorentz-invariant. Currently, it is widely used in non-relativistic QCD. In this effective theory, there is no Lorentz invariance from the very beginning, and the disadvantage does not matter. On the other hand, separating Coulomb and transverse (chromomagnetic) gluons helps to estimate various contributions when considering non-relativistic bound states.

The gluon propagators in this gauge are

 $= -i\delta^{ab} D_0(q)\,,$ $= -i\delta^{ab} D_0^{ij}(q)\,.$ (2.1)

The Coulomb gluon propagator

$$D_0(q) = -\frac{1}{\vec{q}^{\,2}} \tag{2.2}$$

does not depend on q_0. This means that the Coulomb gluon propagates instantaneously. The transverse gluon propagator

$$D_0^{ij}(q) = \frac{1}{q^2 + i0}\left(\delta^{ij} - \frac{q^i q^j}{\vec{q}^{\,2}}\right) \tag{2.3}$$

describes a normal propagating massless particle (the usual denominator $q^2 + i0$). In accordance with the gauge-fixing condition $\vec{\nabla} \cdot \vec{A}^a = 0$, this 3-dimensional tensor is transverse to $\vec{q}$.

The three-gluon vertices look as usual. There is no vertex with three Coulomb legs (if we contract the usual three-gluon vertex with $v^\mu = (1, \vec{0})$ in all three indices, it vanishes).

We are going to calculate the potential between an infinitely heavy quark and an infinitely heavy antiquark in the colour-singlet state.

The infinitely-heavy quark propagator

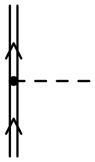
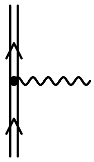
 $= \dfrac{i}{p_0 + i0}$ (2.4)

does not depend on $\vec{p}$ (if the mass M were large but finite, it would be

$$\frac{i}{p_0 - \frac{\vec{p}^{\,2}}{2M} + i0}\,;$$

the kinetic energy disappears at $M \to \infty$). This means that the coordinate-space propagator is $\sim \delta(\vec{r})\theta(t)$: the infinitely heavy quark does not move in space, and propagates forward in time. The energy of an on-shell infinitely heavy quark is 0, and does not depend on $\vec{p}$ (it would be $\vec{p}^{\,2}/(2M)$ if M were large but finite).

An infinitely heavy quark interacts with Coulomb gluons but not with transverse ones:

 $= ig_0 t^a\,,$  $= 0\,.$ (2.5)

The infinitely-heavy antiquark – Coulomb gluon vertex differs from (2.5) by a minus sign.

2.2 Quark–antiquark potential

We calculate the scattering amplitude of an on-shell infinitely heavy quark and an on-shell infinitely heavy antiquark in the colour-singlet state with the momentum transfer $\vec{q}$ in quantum field theory (Fig. 1), and equate it to the same amplitude in quantum mechanics. In the Born approximation, it is

$$iU_{\vec{q}}\,. \tag{2.6}$$

Two-particle-reducible diagrams in quantum field theory (which can be cut into two disconnected parts by cutting a quark line and an antiquark one) correspond to higher Born approximations in quantum mechanics, and we don't have to consider them.

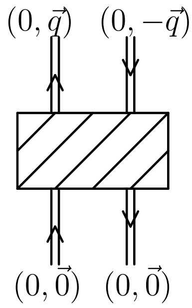
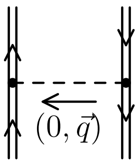


Figure 1: On-shell scattering amplitude

To the lowest order in g_0^2, the scattering amplitude is



$$= -iC_F g_0^2 D(0,\vec{q}) = iC_F \frac{g_0^2}{\vec{q}^{\,2}}\,. \tag{2.7}$$

Therefore,

$$U_{\vec{q}} = C_F g_0^2 D(0,\vec{q}) = -C_F \frac{g_0^2}{\vec{q}^{\,2}}\,, \tag{2.8}$$

and we obtain the Coulomb attraction potential

$$U(r) = -C_F \frac{\alpha_s}{r} \,. \tag{2.9}$$

Some people prefer to discuss the quark–antiquark potential in terms of the vacuum average of a Wilson loop (Fig. 2) with $T \gg r$. Of course, this is exactly the same thing, because the infinitely-heavy quark propagator *is* a straight Wilson line along the 4-velocity v.

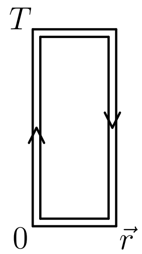


Figure 2: Wilson loop

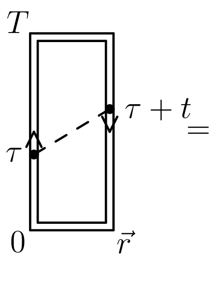
The energy of a colour-singlet quark–antiquark pair separated by $\vec{r}$ is $U(\vec{r})$. Therefore, neglecting boundary effects near time 0 and T, the Wilson loop is

$$e^{-i\,U(\vec{r})\,T} \,, \tag{2.10}$$

or, to the first order in α_s,

$$1 - i\,U(\vec{r})\,T \,.$$

This order-α_s term is



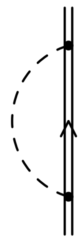
$$= -i\,C_F\,g_0^2\,T \int D(t,\vec{r})\,dt = -i\,C_F\,g_0^2\,T \int \frac{d^{d-1}\vec{q}}{(2\pi)^{d-1}}\,D(0,\vec{q})\,e^{i\,\vec{q}\vec{r}} \tag{2.11}$$

(integration in τ gives T), and hence we obtain (2.9).

2.3 Corrections to the scattering amplitude

Now we want to calculate the first correction to the scattering amplitude in Fig. 1.

First of all, there are external-leg renormalization factors. They are given by the derivative of the infinitely-heavy quark self-energy at its mass shell. The infinitely heavy quark only interacts with Coulomb gluons, and its self-energy is



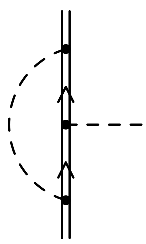
$$\sim \int \frac{d^{d-1}\vec{k}}{\vec{k}^{\,2}} = 0 \tag{2.12}$$

(because the infinitely-heavy quark propagator does not depend on $\vec{k}$). In the coordinate space, the infinitely heavy quark propagates along time, and the Coulomb gluon — along space. Therefore, the two vertices are at the same space–time point, and the self-energy is

$$\sim D(t=0, \vec{r}=0) \sim \int \frac{d^{d-1}\vec{k}}{\vec{k}^{\,2}} e^{i\vec{q}\cdot\vec{r}} \bigg|_{\vec{r}=0} \sim U(\vec{r}=0) \Rightarrow 0\,. \tag{2.13}$$

This is the classical self-energy of a point charge, and it is linearly divergent. In dimensional regularization, it is 0[1].

Transverse gluons don't interact with the infinitely heavy quark, and hence there is only one vertex correction. It vanishes for the same reason:

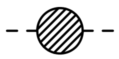


$$= 0\,. \tag{2.14}$$

Therefore, we only need to consider vacuum polarization corrections (Fig. 3):

$$U_{\vec{q}} = C_F g_0^2 D(0, \vec{q})\,, \tag{2.15}$$

where the Coulomb-gluon self-energy is



$$= -i\vec{q}^{\,2}\Pi(q)\,, \tag{2.16}$$

and its propagator is

$$D(q) = -\frac{1}{\vec{q}^{\,2}} \frac{1}{1-\Pi(q)} = -\frac{1}{\vec{q}^{\,2}} (1+\Pi(q)) \tag{2.17}$$

(up to the first correction).

[1]This linear ultraviolet divergence leads to an ultraviolet renormalon singularity at Borel parameter $u = 1/2$, and hence to an ambiguity of the on-shell heavy-quark mass proportional to Λ_{QCD} [15] (see also [12], Chapter 8).

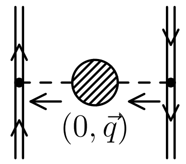


Figure 3: Vacuum polarization corrections

2.4 Quark and transverse-gluon loops

There are several contributions to the Coulomb-gluon self-energy at one loop. We begin with the quark-loop contribution Π_q. It is Lorentz-invariant, and is given by the spectral representation

$$\Pi_q(q^2) = \int\limits_0^\infty \frac{\rho_q(s)\,ds}{q^2 - s + i0} \tag{2.18}$$

with a positive spectral density

$$\rho_q(s) \geqslant 0\,. \tag{2.19}$$

The propagator (and $U_{\vec{q}}$) is a superposition of the massless propagator and massive ones having various masses, with positive weights:

$$\begin{aligned} U_{\vec{q}} &= -C_F \frac{g_0^2}{\vec{q}^{\,2}} \left[1 - \int \frac{\rho_q(s)\,ds}{\vec{q}^{\,2} + s} + \cdots \right] \\ &= -C_F g_0^2 \left[\left(1 - \int \frac{\rho_q(s)\,ds}{s} \right) \frac{1}{\vec{q}^{\,2}} + \int \frac{\rho_q(s)\,ds}{\vec{q}^{\,2} + s} + \cdots \right]. \end{aligned} \tag{2.20}$$

Therefore, the potential $U(r)$ is a superposition of the Coulomb potential and Yukawa ones having various radii, with positive weights:

$$U(r) = -C_F \frac{g_0^2}{4\pi r} \left[1 - \int \frac{\rho_q(s)\,ds}{s} + \int \rho_q(s)\, e^{-\sqrt{s}\,r}\, ds + \cdots \right]. \tag{2.21}$$

The farther we are from the source, the more Yukawa potentials switch off, and the weaker is the interaction. We have screening.

In QED, this is the only effect. It seems that screening follows from very general principles: causality (it allows one to express (2.18) $\Pi(q^2)$ via its imaginary part $\rho(s)$) and unitarity (it says that this imaginary part (2.19) is a sum of modulus squared of transition amplitudes to intermediate states). The only chance *not* to get screening is to find some contribution which is not given by the spectral representation.

Let's complete our calculation. Using Cutkosky rule (Fig. 4) in the q rest frame, we see that two integrations are eliminated by two δ-functions, and

$$\rho_q(s) = T_F n_f \frac{g_0^2 s^{-\varepsilon}}{(4\pi)^{d/2}} \left(\frac{4}{3} + \mathcal{O}(\varepsilon) \right) \tag{2.22}$$

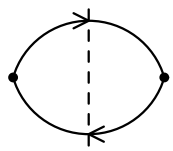


Figure 4: Quark loop discontinuity at the cut

The ultraviolet divergence of Π_q is

$$\int \frac{\rho_q(s)\,ds}{s+\vec{q}^{\,2}}\bigg|_{\mathrm{UV}} = \frac{4}{3} T_F n_f \frac{g_0^2}{(4\pi)^{d/2}} \int\limits_{\sim\vec{q}^{\,2}}^{\infty} s^{-1-\varepsilon} ds = \frac{4}{3} T_F n_f \frac{\alpha_s}{4\pi\varepsilon}\,.$$

Keeping only this divergent part, we obtain

$$U_{\vec{q}} = -C_F \frac{g_0^2}{\vec{q}^{\,2}} \left[1 + \frac{4}{3} T_F n_f \frac{\alpha_s}{4\pi\varepsilon} + \cdots \right], \tag{2.23}$$

where dots are contributions of other diagrams.

Qualitatively, the transverse-gluon loop is just like the quark one. Its contribution in the Coulomb gauge $\Pi_t(q)$ is not Lorentz-invariant, and the spectral representation is a little more complicated:

$$\Pi_t(q_0^2, \vec{q}^{\,2}) = \int \frac{\rho_t(s, \vec{q}^{\,2})\,ds}{q^2 - s + i0}\,. \tag{2.24}$$

When calculating the discontinuity by the Cutkosky rule (Fig. 5), we cannot use the q rest frame, and there is one extra integration not eliminated by δ-functions. The general expression for $\rho_t(s, \vec{q}^{\,2})$ is rather complicated (this is the only point in which the Coulomb-gauge derivation is more complicated than the usual covariant one).

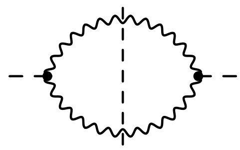


Figure 5: Transverse-gluon loop discontinuity at the cut

Fortunately, we don't need this general expression. In order to obtain the ultraviolet divergence of $\Pi_t(q)$, its limiting form at $s \gg \vec{q}^{\,2}$ is sufficient. We can simplify the integrand of our single integral, and obtain

$$\rho_t(s, \vec{q}^{\,2}) = C_A \frac{g_0^2 s^{-\varepsilon}}{(4\pi)^{d/2}} \left(\frac{1}{3} + \mathcal{O}(\vec{q}^{\,2}/s, \varepsilon) \right)\,. \tag{2.25}$$

The ultraviolet-divergent contribution of this diagram to $U_{\vec{q}}$ is

$$U_{\vec{q}} = -C_F \frac{g_0^2}{\vec{q}^{\,2}} \left[1 + \frac{1}{3} C_A \frac{\alpha_s}{4\pi\varepsilon} + \cdots \right] . \tag{2.26}$$

2.5 Coulomb gluon

There is one more contribution: the loop with a Coulomb gluon and a transverse one. We can understand its sign qualitatively.

Let's consider infinitely heavy quark and antiquark at a distance r (in the colour-singlet state) and the transverse-gluon field (Fig. 6). If we neglect the interaction, the ground-state energy is just

$$E_0 = U(r) , \tag{2.27}$$

because the vacuum energy of the transverse-gluon field is 0.

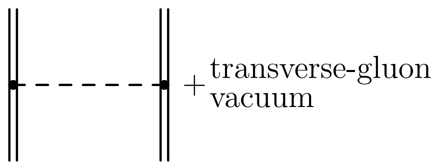


Figure 6: Q, $\bar{Q}$, and the transverse-gluon field

Now let's take into account their interaction in the second order of perturbation theory. Transverse gluons don't interact with infinitely heavy quarks, they can only couple to Coulomb gluons exchanged between the quark and the antiquark (Fig. 7). The ground-state energy decreases in the second order of perturbation theory. This means that the Coulomb attraction becomes stronger — antiscreening [16]!

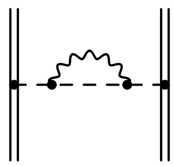


Figure 7: Interaction in the second order of perturbation theory

This loop (Fig. 8) depends on $\vec{q}$ but not on q_0, because we can always route the external energy q_0 via the Coulomb propagator, and it does not depend on energy. Therefore, this loop has no cut in the q_0 complex plane, and is not given by the spectral representation. Speaking more formally, we can say that the spectral density is 0, and the whole result is given by the subtraction term, which does not depend on q_0 (but depends on $\vec{q}$).

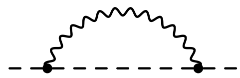


Figure 8: Loop with Coulomb and transverse gluons

This diagram can be easily calculated. Only the transverse-gluon propagator depends on k_0:

$$\Pi_c(\vec{q}^{\,2}) = \int \frac{d^d k}{(2\pi)^d} \frac{f(\vec{k},\vec{q})}{k^2 + i0} \,.$$

The integral in k_0 can be taken first:

$$\int \frac{dk_0}{2\pi} \frac{1}{k_0^2 - \vec{k}^{\,2} + i0} = - \frac{i}{2\left(\vec{k}^{\,2}\right)^{1/2}} \,.$$

We are left with a $(d-1)$-dimensional integral similar to the usual d-dimensional massless loop. It can be reduced to Γ-functions via Feynman parametrization. The result is

$$\Pi_c(\vec{q}^{\,2}) = C_A \frac{g_0^2 \left(\vec{q}^{\,2}\right)^{-\varepsilon}}{(4\pi)^{d/2}} \left(\frac{4}{\varepsilon} + \mathcal{O}(1) \right) \,. \tag{2.28}$$

At last, let's assemble all our findings. Keeping only ultraviolet-divergent terms in the corrections, we have the momentum-space potential

$$U_{\vec{q}} = -C_F \frac{g_0^2}{\vec{q}^{\,2}} \left\{ 1 + \frac{g_0^2 (\vec{q}^{\,2})^{-\varepsilon}}{(4\pi)^{d/2}} \left[\left(\left(4 - \frac{1}{3} \right) C_A - \frac{4}{3} T_F n_f \right) \frac{1}{\varepsilon} + \cdots \right] \right\} . \tag{2.29}$$

When expressed via the renormalized $\alpha_s(\mu)$:

$$\frac{g_0^2}{(4\pi)^{d/2}} = \mu^{2\varepsilon} \frac{\alpha_s(\mu)}{4\pi} Z_\alpha e^{\gamma\varepsilon} \,, \qquad Z_\alpha = 1 - \beta_0 \frac{\alpha_s}{4\pi\varepsilon} \,, \tag{2.30}$$

this potential must be finite. Therefore,

$$\beta_0 = \left(4 - \frac{1}{3} \right) C_A - \frac{4}{3} T_F n_f \,. \tag{2.31}$$

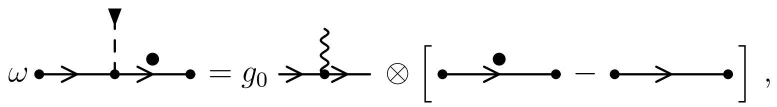
The antiscreening term 4 comes from the Coulomb-gluon loop; it overweights the screening term $-1/3$ from the transverse-gluon loop.

2.6 Ward identity

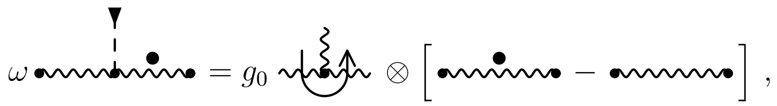
Until now, we used the potential between infinitely heavy quark and antiquark. Or, if we cut our picture in two halves, the infinitely-heavy quark – Coulomb gluon vertex.

This vertex is convenient, because both the external-leg renormalization and the vertex corrections vanish (Sect. 2.3). However, we can use some other vertex, for example, the finite-mass quark – Coulomb gluon vertex, equally easily. The tool needed to this end is the Ward identity. The Coulomb gauge is ghost-free. Therefore, it is natural to expect that the Ward identities in this gauge are simple, like in QED, in contrast to Ward–Slavnov–Taylor identities in covariant gauges, which are complicated by extra ghost terms.

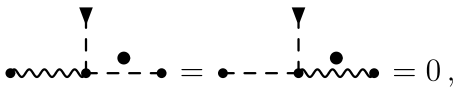
We shall proceed exactly as in QED. There, an external photon leg insertion with its polarization 4-vector parallel to its 4-momentum gives a difference of two propagators; most terms cancel pairwise. Now we have a Coulomb gluon, which is polarized along time. Therefore, let's set its incoming momentum to $q = \omega v$, where $v = (1, \vec{0})$ is the 4-velocity of the reference frame. We denote such an external Coulomb gluon by a leg with a black triangle. A dot near a propagator means that its momentum is shifted by q. It is easy to check the identities



(2.32)



(2.33)

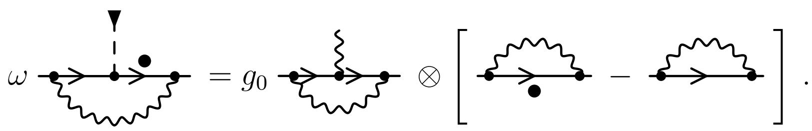


(2.34)

where the right-hand sides are written as (colour structure) $\otimes$ (Lorentz structure), and the curved arrow in (2.33) shows the order of indices in the colour structure if^{abc}. The last equality (2.34) is obvious: let's consider one of these diagrams with the transverse-gluon line removed; this object is a vector (has one index), and depends on two vectors, v and the transverse-gluon momentum p; and the transverse-gluon propagator is transverse to both v and p.

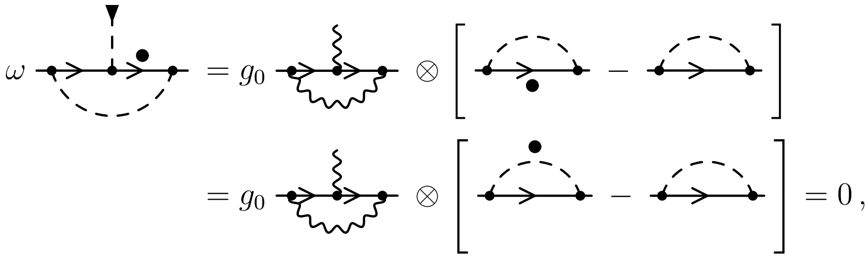
In covariant gauges, the right-hand side of (2.33) contains extra terms, where one of the gluon lines becomes longitudinally polarized. These terms produce ghost propagators, thus transforming simple Ward identities to more complicated Ward–Slavnov–Taylor ones.

Now we are ready to apply these identities to the one-loop vertex. Let's consider the QED-like diagram with a transverse gluon:



(2.35)

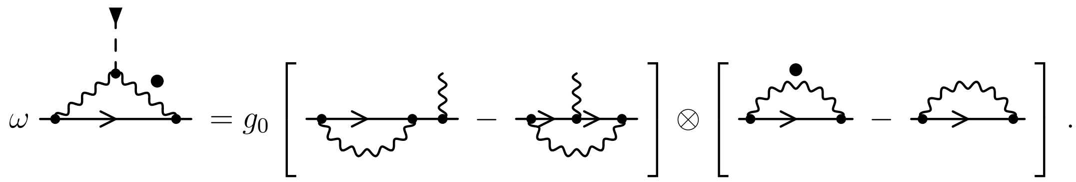
With a Coulomb gluon,



$$(2.36)$$

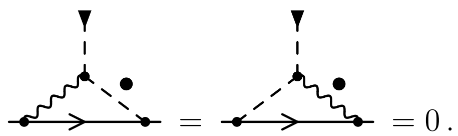
because we can route the extra momentum q along the Coulomb-gluon line instead of the quark one, and the Coulomb propagator does not depend on an additional momentum $q = \omega v$. In QED, the only contribution is (2.35) (and there is no colour factor in its right-hand side): the difference of the fermion self-energies with the momenta $p+q$ and p.

In QCD, we also have



$$(2.37)$$

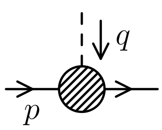
Here we have used the definition of if^{abc} via the commutator in order to re-write the colour structure as a difference: one term is the same as in (2.35) (and they cancel); the second one is the colour structure C_F of the quark self-energy times that of the elementary vertex it^a. Two remaining contributions vanish, due to (2.34):



$$(2.38)$$

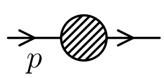
Qualitatively, we can say that the non-abelian charge flows along both lines in the quark self-energy diagram; the longitudinal gluon insertion into one of these lines "measures" the charge flowing along this line; in total, the whole quark charge flows, thus giving the full colour structure of the self-energy C_F.

The Ward identity provides a relation between the quark – Coulomb gluon vertex



$$= ig_0 t^a \Gamma(p,q)\,, \qquad \Gamma(p,q) = \gamma_0 + \Lambda(p,q) \tag{2.39}$$

at $q = \omega v$ and the quark self-energy



$$= -i\Sigma(p) \tag{2.40}$$

which is related to the propagator:

$$S(p) = \frac{1}{\not{p} - m_0 - \Sigma(p)}\,.$$

The Ward identity can be written as

$$\omega\Lambda(p,\omega v) = \Sigma(p) - \Sigma(p+\omega v) \tag{2.41}$$

or

$$\omega\Gamma(p,\omega v) = S^{-1}(p+\omega v) - S^{-1}(p)\,. \tag{2.42}$$

We have proved it at one loop, but it stays correct at higher loops, too.

Now let's recall how the coupling constant renormalization

$$g_0 = Z_\alpha^{1/2} g$$

is derived. When expressed via the renormalized coupling g, the vertex and the propagator should be

$$\Gamma = Z_\Gamma \Gamma_r\,, \qquad S = Z_\psi S_r\,,$$

where the renormalized vertex and propagator are finite at $\varepsilon \to 0$. The scattering amplitude is obtained by multiplying the proper vertex by the external-leg renormalization factors:

$$g_0 \Gamma Z_\psi Z_A^{1/2} = g \Gamma_r Z_\alpha^{1/2} Z_\Gamma Z_\psi Z_A^{1/2}\,,$$

and it must be finite. Therefore, $Z_\alpha^{1/2} Z_\Gamma Z_\psi Z_A^{1/2}$ must be finite at $\varepsilon \to 0$. But the only minimal renormalization constant finite at $\varepsilon \to 0$ is 1:

$$Z_\alpha = \left(Z_\Gamma Z_\psi\right)^{-2} Z_A^{-1}\,. \tag{2.43}$$

I. e., in order to find the coupling-constant renormalization, one needs the vertex renormalization factor and all external-leg renormalization factors for this vertex.

The Ward identity makes things simpler. From (2.42), $Z_\Gamma Z_\psi$ must be finite at $\varepsilon \to 0$, and hence

$$Z_\Gamma Z_\psi = 1\,. \tag{2.44}$$

Therefore,

$$Z_\alpha = Z_A^{-1}\,. \tag{2.45}$$

The coupling constant renormalization is determined by the Coulomb-gluon propagator renormalization only. And this is exactly what we studied in Sect. 2.4–2.5.

3 Chromomagnetic properties of vacuum (Nielsen 1981)

3.1 Dielectric and magnetic properties of vacuum

We shall start from QED, and then consider non-abelian theory.

We shall use a popular approach to renormalization due to Wilson. Suppose we have the path integral in momentum space. We integrate out fields with momenta $p > \Lambda$:

$$\int \prod_p d\phi_p e^{iS} = \int \prod_{p<\Lambda} d\phi_p e^{iS_\Lambda}\,, \tag{3.1}$$

where the effective action is defined by

$$e^{iS_\Lambda} = \int \prod_{p>\Lambda} d\phi_p e^{iS}\,. \tag{3.2}$$

When we are interested in processes with small momenta $p_i \ll \Lambda$, S_Λ can be expressed via a local Lagrangian. At the leading order in $1/\Lambda$, it is the standard dimension-4 Lagrangian; it contains renormalized fields (at the scale Λ) and the coupling $g(\Lambda)$.

Now we want to integrate out also fields with momenta between Λ' and Λ (where $\Lambda' \ll \Lambda$):

$$e^{iS_{\Lambda'}} = \int \prod_{\Lambda'<p<\Lambda} d\phi_p e^{iS_\Lambda}\,, \tag{3.3}$$

and to obtain $g(\Lambda')$. If we are interested in processes with characteristic momentum p, we should use Λ' not too far separated from p, in order to avoid large logarithms.

For example, if we consider interaction between a quark and an antiquark at a distance r, we should use $\Lambda' \sim 1/r$. Then the Coulomb potential will be simply $-e^2(\Lambda')/r$. But if we start from the theory with a high cut-off Λ, then vacuum modes with momenta between Λ' and Λ will act as a dielectric medium, and the potential will be $-e^2(\Lambda)/(\varepsilon r)$:

$$e^2(\Lambda') = \frac{e^2(\Lambda)}{\varepsilon}\,, \tag{3.4}$$

where only modes with momenta between Λ' and Λ contribute to ε. If $\varepsilon > 1$, we have screening; if $\varepsilon < 1$ — antiscreening (asymptotic freedom).

In the case of an ordinary matter, its dielectric and magnetic properties are independent. But vacuum should be Lorentz-invariant. Signals should propagate with velocity 1:

$$\varepsilon\mu = 1\,. \tag{3.5}$$

Therefore, diamagnetic vacuum ($\mu < 1$) means screening, and paramagnetic one ($\mu > 1$) — asymptotic freedom.

When we switch magnetic field B on, the vacuum energy changes by

$$\Delta E_{\mathrm{vac}} = \left(\mu^{-1} - 1\right) \frac{B^2}{2} V \,. \tag{3.6}$$

We shall show that

$$\Delta E_{\mathrm{vac}} = -\beta_0 \frac{g^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \cdot \frac{B^2}{2} V \,. \tag{3.7}$$

This means

$$\mu = \varepsilon^{-1} = 1 + \beta_0 \frac{g^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \,, \tag{3.8}$$

and therefore

$$e^2(\Lambda') = \left[1 + \beta_0 \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2}\right] e^2(\Lambda) \,, \tag{3.9}$$

i.e., β_0 is indeed the 1-loop β-function coefficient.

The vacuum energy of a charged scalar field (describing particles and antiparticles) is

$$E_{\mathrm{vac}} = 2 \sum_i \frac{\omega_i}{2} = \sum_i \omega_i \,. \tag{3.10}$$

For a charged fermion field, it is the energy of the Dirac sea

$$E_{\mathrm{vac}} = - \sum_i \omega_i \,. \tag{3.11}$$

Therefore, in general we can use (3.10) multiplied by $(-1)^{2s}$.

3.2 Pauli paramagnetism

How the vacuum energy changes when we switch the magnetic field B on? There are two effects: spin and orbital, and they can be considered separately.

First we discuss interaction of the spin magnetic moment with the magnetic field. Without the field, a massless particle has the energy $\omega = k$. When the magnetic field is switched on, the energy becomes

$$\omega = \sqrt{k^2 - g_s s_z e B} \,, \tag{3.12}$$

where g_s is the gyromagnetic ratio for our spin-s particle (we shall discuss this in Sect. 3.6 in more detail).

Suppose the magnetic field is along the z axis. Massless particles only have 2 spin projections $s_z = \pm s$. The vacuum energy change at the order B^2 is

$$\begin{aligned}
\Delta E_{\mathrm{Pauli}} &= (-1)^{2s} \int \frac{V\, d^3k}{(2\pi)^3} \left[\sqrt{k^2 + g_s s e B} + \sqrt{k^2 - g_s s e B} - 2k\right] \\
&= -(-1)^{2s} V \frac{(g_s s e B)^2}{4} \int \frac{d^3k}{(2\pi)^3} \frac{1}{k^3} = -(-1)^{2s} V \frac{(g_s s e B)^2}{8\pi^2} \int \frac{dk}{k} \\
&= -2(-1)^{2s} (g_s s)^2 \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \cdot \frac{B^2}{2} V \,,
\end{aligned} \tag{3.13}$$

where only modes with momenta between Λ' and Λ are included. Let's stress once more that what we are calculating is the *vacuum* energy: there are no particles, only empty modes.

3.3 Landau levels

Now let's discuss the effect of magnetic field B on the orbital motion. In order not to have complications related to spin, we consider a massless charged scalar field φ. Its Lagrangian is

$$L = (D_\mu \varphi)^+ D^\mu \varphi \,. \tag{3.14}$$

For magnetic field B along the z axis, we can choose the vector potential as $A_y = Bx$, $A_x = A_z = 0$. Then the equation of motion is

$$\left[\nabla^2 - e^2 B^2 x^2 - 2ieBx\frac{\partial}{\partial y} + E^2\right]\varphi = 0 \,. \tag{3.15}$$

Its solutions have the form

$$\varphi = e^{i(k_y y + k_z z)}\varphi(x) \,. \tag{3.16}$$

The equation for $\varphi(x)$ has the same form as the Schrödinger equation for harmonic oscillator:

$$\left[-\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{\omega^2}{2}x^2 - E_n\right]\psi_n = 0 \,. \tag{3.17}$$

The oscillator energies are

$$E_n = \omega\left(n + \tfrac{1}{2}\right) \,. \tag{3.18}$$

Comparing (3.15) with (3.17), we see that the energies of our massless particle in magnetic field B are

$$E^2 = k_z^2 + 2eB\left(n + \tfrac{1}{2}\right) \,, \tag{3.19}$$

and the corresponding wave functions are

$$\varphi = e^{i(k_y y + k_z z)}\psi_n\left(x - \frac{k_y}{eB}\right) \,. \tag{3.20}$$

In other words, E^2 consists of discrete Landau levels of transverse motion plus free motion along the magnetic field.

Each Landau level has a high degree of degeneracy. In order to find it, let's put our particle into a large box $V = L_x \times L_y \times L_z$. Then the allowed longitudinal momenta are

$$k_z = \frac{2\pi}{L_z}n_z \,;$$

therefore, the number of allowed modes in the interval dk_z is

$$dn_z = \frac{L_z\, dk_z}{2\pi} \,.$$

Similarly, the allowed values of k_y are

$$k_y = \frac{2\pi}{L_y} n_y \,.$$

As we see from (3.20), k_y is related to the x coordinate of the center of the Larmor orbit. It must be inside our box:

$$\frac{k_y}{eB} \in [0, L_x] \,.$$

Therefore,

$$n_y \in \left[0, \frac{eB\, L_x L_y}{2\pi}\right] \,.$$

Energy does not depend on k_y. Hence the degeneracy of each Landau level is

$$\frac{eB\, L_x L_y}{2\pi} \,. \tag{3.21}$$

It is equal to the magnetic flux through our box ($B\, L_x L_y$) measured in flux quanta $2\pi/e$.

The spectrum of $E_\perp^2 = E^2 - k_z^2$ at $B = 0$ is continuous. The number of states in the interval $dE_\perp^2$ is

$$\frac{L_x L_y}{4\pi} dE_\perp^2 \,.$$

When the magnetic field B is switched on, each interval $\Delta E_\perp^2 = 2eB$ is contracted into a single Landau level (Fig. 9). The number of states in each interval is the Landau level degeneracy (3.21).

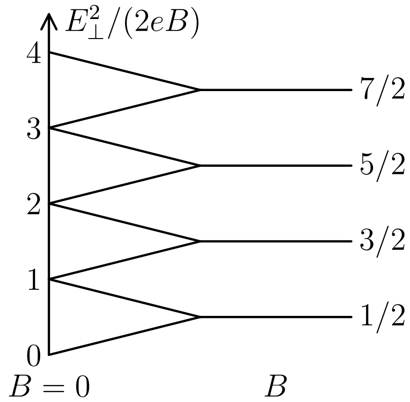


Figure 9: Continuous spectrum at $B = 0$ and Landau levels

The vacuum energy of our massless charged scalar field in the magnetic field B is

$$E_{\text{vac}} = \sum_{n=0}^{\infty} f\left(n + \tfrac{1}{2}\right) \,, \tag{3.22}$$

where

$$f(x) = \frac{eBV}{(2\pi)^2} \int\limits_{-\infty}^{+\infty} \sqrt{k_z^2 + 2eBx}\, dk_z\,. \tag{3.23}$$

3.4 Euler summation formula

Integrals are more convenient than sums. If we have a smooth function $f(x)$ (i.e., its characteristic length is $L \gg 1$), then, obviously,

$$\sum_{n=0}^{N} f\left(n + \tfrac{1}{2}\right) \approx \int\limits_{0}^{N+1} f(x)\, dx\,. \tag{3.24}$$

But how to find a correction to this formula?

Let's re-write this integral as a sum of integrals over unit intervals:

$$I = \int\limits_{0}^{N+1} f(x)\, dx = \sum_{n=0}^{N} \int\limits_{-1/2}^{1/2} f\left(n + \tfrac{1}{2} + x\right)\, dx\,.$$

The smooth function $f(x)$ can be expanded in Taylor series in each interval:

$$I = \sum_{n=0}^{N} \int\limits_{-1/2}^{1/2} \left[f\left(n + \tfrac{1}{2}\right) + \frac{1}{2} f''\left(n + \tfrac{1}{2}\right) x^2 + \cdots \right] dx$$

(terms with odd powers of x don't contribute). Calculating the integrals, we get

$$I = \sum_{n=0}^{N} f\left(n + \tfrac{1}{2}\right) + \frac{1}{24} \sum_{n=0}^{N} f''\left(n + \tfrac{1}{2}\right) + \cdots$$

The second sum in the right-hand side is a small correction (because $f'' \sim f/L^2$); therefore, we can replace it by an integral (see (3.24)):

$$I = \sum_{n=0}^{N} f\left(n + \tfrac{1}{2}\right) + \frac{1}{24} \int\limits_{0}^{N+1} f''(x)\, dx + \cdots = \sum_{n=0}^{N} f\left(n + \tfrac{1}{2}\right) + \frac{1}{24} f'(x) \Big|_{0}^{N+1} + \cdots$$

Finally, we arrive at the Euler summation formula

$$\sum_{n=0}^{N} f\left(n + \tfrac{1}{2}\right) = \int\limits_{0}^{N+1} f(x)\, dx - \frac{1}{24} f'(x) \Big|_{0}^{N+1} + \cdots \tag{3.25}$$

The correction here is of order $1/L^2$. It is easy to find a few more corrections, if desired.

3.5 Landau diamagnetism

The vacuum energy in magnetic field B (3.22) can be re-written using the Euler formula (3.25) as

$$E_{\rm vac} = \int_0^\infty f(x) - \frac{1}{24} f'(x)\Big|_0^\infty . \tag{3.26}$$

The integral here is the vacuum energy at $B = 0$ (when the spectrum of $E_\perp^2$ is continuous, Fig. 9). The shift of the vacuum energy due to the magnetic field is

$$\Delta E_{\rm Landau} = \frac{1}{24} f'(0) = \frac{e^2 B^2 V}{48\pi^2} \int \frac{dk}{k} = \frac{1}{3} \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \cdot \frac{B^2}{2} V , \tag{3.27}$$

where only modes with momenta between Λ' and Λ are included[2].

3.6 The QED result

The full $\Delta E_{\rm vac}$ in QED is the sum of the spin contribution (3.13) and the orbital one (3.27):

$$\Delta E_{\rm vac} = -\beta_0 \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \cdot \frac{B^2}{2} V , \tag{3.28}$$

where

$$\beta_0 = 2 \sum_s (-1)^{2s} \left[(g_s s)^2 - \frac{n_s}{6} \right] \tag{3.29}$$

is the sum over all charged fields, and n_s is the number of polarization states: $n_0 = 1$, $n_{s\neq 0} = 2$.

Let's demonstrate that a Dirac particle (e.g., electron) has gyromagnetic ratio $g_{1/2} = 2$. For a massless particle, the Dirac equation is

$$\not{D}\psi = 0 , \qquad D_\mu = \partial_\mu - ieA_\mu . \tag{3.30}$$

Let's multiply it by $\not{D}$:

$$\not{D}^2\psi = 0 , \qquad \not{D}^2 = \partial^2 - e^2 A^2 - 2ieA^\mu \partial_\mu - ie\gamma^\mu\gamma^\nu \partial_\mu A_\nu$$

(in the last term, ∂_μ acts only on A_ν). If we suppose that $\partial \cdot A = 0$,

$$\not{D}^2 = D^2 - \frac{ie}{4} F_{\mu\nu} [\gamma^\mu, \gamma^\nu] .$$

We choose $A^\mu = (0, 0, Bx^1, 0)$, then $F_{12} = -F_{21} = -B$ and

$$\not{D}^2 = D^2 + ieB\gamma^1\gamma^2 = D^2 + 2eBs_z .$$

[2]Many subtleties have been swept under the carpet in this derivation sketch. A somewhat more accurate derivation [14] yields the same result.

The equation of motion in magnetic field B (directed along z) becomes

$$\left[\nabla^2 - e^2B^2x^2 - 2ieBx\frac{\partial}{\partial y} + 2eBs_z + E^2\right]\psi = 0\,. \tag{3.31}$$

Its energy spectrum is

$$E^2 = k_z^2 + 2eB\left(n + \tfrac{1}{2}\right) - 2eBs_z \tag{3.32}$$

(see Sect. 3.3)[3]. Comparing this with (3.12), we see that

$$g_{1/2} = 2\,. \tag{3.33}$$

So, the electron contribution to β_0 (3.29) is $-4/3$; if a charged scalar particle exists, it contributes $-1/3$. Note that for $s = 1/2$ the spin effect overweights the orbital one. However, due to the factor $(-1)^{2s}$, the spin effect leads to diamagnetism, and the orbital one to paramagnetism. This is because here we are interested in the Dirac see. In physics of metals, we are interested in positive-energy electrons (below the Fermi surface), and the spin effect gives Pauli paramagnetism, while the orbital one — Landau diamagnetism.

3.7 The QCD result

Now it is easy to obtain β_0 in QCD. We have chromomagnetic field instead of magnetic. Let's choose its colour orientation along an axis a_0 such that t^{a_0} is diagonal (for the $SU(3)$ colour group with the standard choice of the generators t^a, t^8 is diagonal, and we choose $a_0 = 8$).

The quark contribution follows from the electron one in QED. The contribution to β_0 is proportional to the charge squared. The sum of squares of colour "charges" of a quark is $\operatorname{Tr} t^{a_0}t^{a_0}$ (no summation). Recalling

$$\operatorname{Tr} t^a t^b = T_F\delta^{ab}\,,$$

we arrive at the contribution

$$-\left(1 - \tfrac{1}{3}\right) 2T_F n_f \tag{3.34}$$

of n_f quark flavours. If there were scalar quarks, each flavour would contribute

$$-\frac{1}{3}2T_F\,.$$

What about gluons? First of all, we have to find their gyromagnetic ratio g_1. We shall do this for the $SU(2)$ colour group, because calculations are simpler in this case; the result will be valid for other colour groups, too. Let's consider the $SU(2)$ Yang–Mills equation

$$D^\nu G^a_{\mu\nu} = \left(\partial^\nu\delta^{ab} + g\varepsilon^{acb}A^{c\nu}\right) G^b_{\mu\nu} = 0\,. \tag{3.35}$$

[3] Strictly speaking, we should consider spin and orbital effects together, using (3.32). It is easy to see that in the order B^2 they can be treated separately.

The external field is A^3_μ. We linearize in the small components $A^{1,2}_\mu$. For $A^-_\mu = A^1_\mu - iA^2_\mu$ we get

$$D^\nu G^-_{\mu\nu} + igG^3_{\mu\nu} A^{-\nu} = 0\,,$$

where

$$G^-_{\mu\nu} = D_\mu A^-_\nu - D_\nu A^-_\mu\,, \qquad D_\mu = \partial_\mu - igA^3_\mu\,.$$

In the $D^\mu A^-_\mu = 0$ gauge, the equation of motion becomes

$$D^2 A^-_\mu - 2igG^3_{\mu\nu} A^{-\nu} = 0$$

(we have used $[D_\mu, D_\nu] = -igG^3_{\mu\nu}$). Our external field is oriented along z in space and along 3 in colour: $G^3_{12} = -G^3_{21} = -B$. Using $A^{-2} = is_z A^{-1}$ ($s_z = \pm 1$), we finally obtain

$$\left[D^2 + 2igBs_z\right] A^{-1} = 0\,. \tag{3.36}$$

This equation looks exactly the same as (3.31), and hence $g_1 = 2$.

Gluons with colour a_1 such that t^{a_1} is diagonal don't interact with our chromomagnetic field (for the standard $SU(3)$ generators, $a_1 = 3$). All other gluons can be arranged into pairs with positive and negative "colour charges" (particles and antiparticles). The sum of their squares (both signs!) is C_A: in the adjoint representation

$$\mathrm{Tr}\, t^a t^b = C_A \delta^{ab}\,,$$

i.e. we have to replace $2e^2 \to C_A g^2$ in QED results.

Finally, we arrive at

$$\beta_0 = \left(4 - \tfrac{1}{3}\right) C_A - \left(1 - \tfrac{1}{3}\right) 2T_F n_f\,. \tag{3.37}$$

Pauli paramagnetism of the gluon vacuum $(g_1 \cdot 1)^2 = 4$ is stronger than its Landau diamagnetism $-1/3$. This leads to antiscreening (asymptotic freedom).

Acknowledgments

I am grateful to I.B. Khriplovich for discussions of [2].

References

[1] V.S. Vanyashin, M.V. Terentev, ZhETF **48** (1965) 565
[Sov. Phys. JETP **21** (1965) 375]

[2] I.B. Khriplovich, Yad. Phys. **10** (1969) 409 [Sov. J. Nucl. Phys. **10** (1970) 235]

[3] G. 't Hooft, Marseilles meeting on field theory (1972), unpublished;
Nucl. Phys. **B62** (1973) 444

[4] D.J. Gross, F. Wilczek, Phys. Rev. Lett. **30** (1973) 1343

[5] H.D. Politzer, Phys. Rev. Lett. **30** (1973) 1346

[6] M. Shifman, in *At the frontiers of particle physics*, World Scientific (2001), v. 1, p. 126

[7] G. 't Hooft, in *The creation of quantum chromodynamics and the effective energy*, World Scientific (2001), p. 9

[8] F. Wilczek, hep-th/9609099

[9] D.J. Gross, Rev. Mod. Phys. **77** (2005) 837;
H.D. Politzer, Rev. Mod. Phys. **77** (2005) 851;
F. Wilczek, Rev. Mod. Phys. **77** (2005) 857

[10] S. Weinberg, *The quantum theory of fields*, v. 2, Cambridge University Press (1996)

[11] A.G. Grozin, *Lectures on QED and QCD: practical calculation and renormalization of one- and multi-loop Feynman diagrams*, World Scientific (2007)

[12] A.G. Grozin, *Heavy quark effective theory*, Springer (2004)

[13] L.F. Abbott, Nucl. Phys. **B185** (1981) 189; Acta Phys. Polon. **B13** (1982) 33

[14] N.K. Nielsen, Amer. J. Phys. **49** (1981) 1171

[15] M. Beneke, V.M. Braun, Nucl. Phys. **B426** (1994) 301

[16] V.N. Gribov, in *12 Winter School of the Leningrad Nuclear Physics Inst.* (1977); translation in *The Gribov theory of quark confinement*, ed. J. Nyiri, World Scientific (2001), p. 24