QUANTUM CHROMODYNAMICS

M. Musakhanov
Uzbekistan National University

Lecture course for the second year magistrates of NUUZ

– Typeset by FoilTEX –

QUARKS AND COLOUR

A fast look into the Particle Data Tables [1] reveals the richness and variety of the hadronic spectrum. The large number of known mesonic and baryonic states clearly signals the existence of a deeper level of elementary constituents of matter: $quarks$ [2]. In fact, the messy hadronic world can be easily understood in terms of a few constituent spin-$\frac{1}{2}$ quark $flavours$:

$Q = +\frac{2}{3}$	u	c	t
$Q = -\frac{1}{3}$	d	s	b

Assuming that mesons are $M \equiv q\bar{q}$ states, while baryons have three quark constituents, $B \equiv qqq$, one can nicely classify the entire hadronic spectrum:

$$
\begin{array}{llll}
\pi^+ = u\bar{d}, & K^+ = u\bar{s}, & K^0 = d\bar{s}, & \pi^0 = (u\bar{u} - d\bar{d})/\sqrt{2} \ \ldots \\
D^+ = c\bar{d}, & D^0 = c\bar{u}, & D_s^+ = c\bar{s} & \ldots \\
B^+ = u\bar{b}, & B^0 = d\bar{b}, & B_s^0 = s\bar{b}, & B_c^+ = c\bar{b} \ \ldots \\
p = uud, & n = udd, & \Sigma^+ = uus, & \Sigma^0 = uds \ \ldots \\
\Sigma_c^+ = udc, & \Sigma_c^{++} = uuc, & \Xi_c^+ = usc, & \Xi_c^0 = dsc \ \ldots \\
\Xi_{cc}^+ = dcc, & \Xi_{cc}^{++} = ucc, & \Omega_{cc}^+ = scc & \ldots
\end{array}
$$

There is a one–to–one correspondence between the observed hadrons and the states predicted by this simple classification; thus, the *Quark Model* appears to be a very useful *Periodic Table of Hadrons*. However, the quark picture faces a problem concerning the Fermi–Dirac statistics of the constituents. Since the fundamental state of a composite system is expected to have $L = 0$, the Δ^{++} baryon $(J = \frac{3}{2})$ corresponds to $u^\uparrow u^\uparrow u^\uparrow$, with the three quark-spins aligned into the same direction $(s_3 = +\frac{1}{2})$ and all relative angular momenta equal to zero. The wave function is symmetric and, therefore, the Δ^{++} state obeys the wrong statistics.

The problem can be solved assuming [3] the existence of a new quantum number, $colour$, such that each species of quark may have $N_c = 3$ different colours: q^α, $\alpha = 1, 2, 3$ (red, yellow, violet). Then one can reinterpret the Δ^{++} as the antisymmetric state

$$\Delta^{++} = \frac{1}{\sqrt{6}} \epsilon^{\alpha\beta\gamma} \left| u_\alpha^\uparrow u_\beta^\uparrow u_\gamma^\uparrow \right\rangle \tag{1}$$

(notice that at least 3 colours are needed for making an antisymmetric state). In this picture, baryons and mesons are described by the colour-singlet combinations

$$B = \frac{1}{\sqrt{6}} \epsilon^{\alpha\beta\gamma} \left| q_\alpha q_\beta q_\gamma \right\rangle , \qquad M = \frac{1}{\sqrt{3}} \delta^{\alpha\beta} \left| q_\alpha \bar{q}_\beta \right\rangle . \tag{2}$$

In order to avoid the existence of non-observed extra states with non-zero colour, one needs to further postulate that all asymptotic states are colorless, i.e. singlets under rotations in colour space. This assumption is known as the *confinement hypothesis*, because it implies the non-observability of free quarks: since quarks carry colour they are confined within colour-singlet bound states. The quark picture is not only a nice mathematical scheme to classify the hadronic world. We have strong experimental evidence of the existence of quarks. Fig. 1 shows a typical $Z \to \mathrm{hadrons}$ event, obtained at LEP. Although there are many hadrons in the final state, they appear to be collimated in 2 *jets* of particles, as expected from a two-body decay $Z \to q\bar{q}$, where the $q\bar{q}$ pair has later *hadronized*.

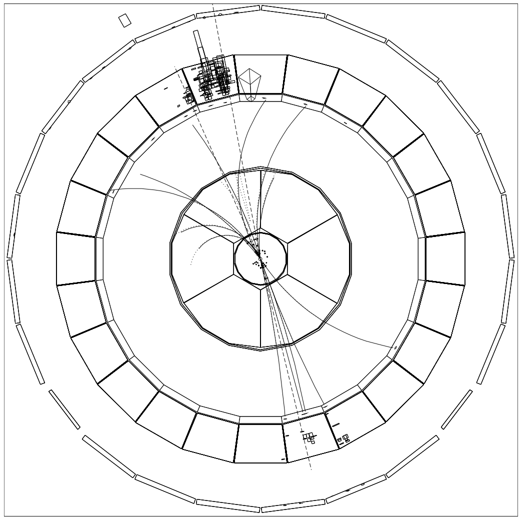


Figure 1: Two-jet event from the hadronic decay of a Z boson (DELPHI).

Evidence of colour

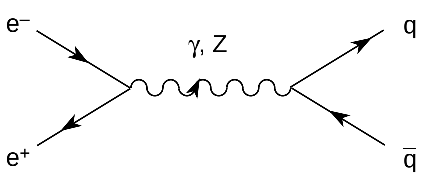


Figure 2: Feynman diagram for $e^+e^- \to \mathrm{hadrons}$.

A direct test of the colour quantum number can be obtained from the ratio

$$R_{e^+e^-} \equiv \frac{\sigma(e^+e^- \to \mathrm{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} \,. \tag{3}$$

The hadronic production occurs through $e^+e^- \to \gamma^*, Z^* \to q\bar{q} \to \mathrm{hadrons.}$ Since quarks are assumed to be confined, the probability to hadronize is just one; therefore, the sum over all possible quarks in the final state will give the total inclusive cross-section into hadrons. At energies well below the Z peak, the cross-section is dominated by the γ-exchange amplitude; the ratio $R_{e^+e^-}$ is then given by the sum of the quark electric charges squared:

$$R_{e^+e^-} \approx N_c \sum_{f=1}^{N_f} Q_f^2 \; = \left\{ \begin{array}{ll} \frac{2}{3} N_c = 2 \,, & (N_f = 3 \; : \; u,d,s) \\ \frac{10}{9} N_c = \frac{10}{3} \,, & (N_f = 4 \; : \; u,d,s,c) \\ \frac{11}{9} N_c = \frac{11}{3} \,, & (N_f = 5 \; : \; u,d,s,c,b) \end{array} \right. \,. \tag{4}$$

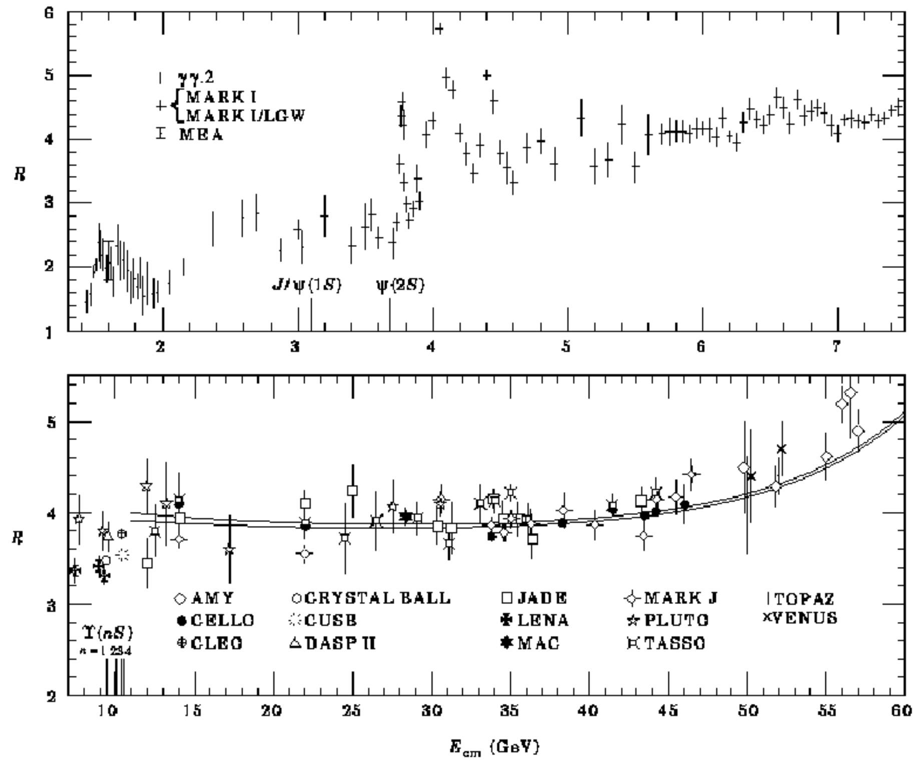


Figure 3: Measurements of $R_{e^+e^-}$ [1]. The two continuous curves are QCD fits.

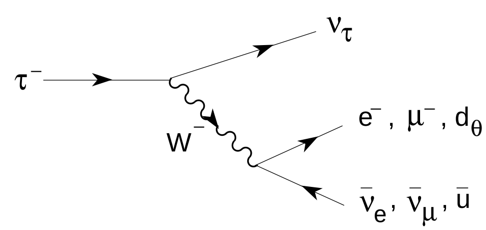


Figure 4: τ-decay diagram.

The measured ratio is shown in Fig. 3. Although the simple formula (4) cannot explain the complicated structure around the different quark thresholds, it gives the right average value of the cross-section (away from the thresholds), provided that N_c is taken to be three. The agreement is better at larger energies. Notice that strong interactions have not been taken into account; only the confinement hypothesis has been used.

The hadronic decay of the τ lepton provides additional evidence for $N_c = 3$. The decay proceeds through the W-emission diagram shown in Fig. 4. Since the W coupling to the charged current is of universal strength, there are $(2 + N_c)$

equal contributions (if final masses and strong interactions are neglected) to the τ-decay width. Two of them correspond to the leptonic decay modes $\tau^- \to \nu_\tau e^- \bar{\nu}_e$ and $\tau^- \to \nu_\tau \mu^- \bar{\nu}_\mu$, while the other N_c are associated with the possible colours of the quark–antiquark pair in the $\tau^- \to \nu_\tau d_\theta u$ decay mode ($d_\theta \equiv \cos\theta_C d + \sin\theta_C s$). Hence, the branching ratios for the different channels are expected to be approximately:

$$B_{\tau\to l} \equiv \quad \mathrm{Br}(\tau^- \to \nu_\tau l^- \bar{\nu}_l) \approx \frac{1}{2+N_c} = \frac{1}{5} = 20\% \,, \tag{5}$$

$$R_\tau \equiv \frac{\Gamma(\tau^- \to \nu_\tau + \mathrm{hadrons})}{\Gamma(\tau^- \to \nu_\tau e^- \bar{\nu}_e)} \approx N_c = 3 \,, \tag{6}$$

which should be compared with the experimental averages [1]:

$$B_{\tau\to e} = (18.01 \pm 0.18)\% \,, \qquad\qquad B_{\tau\to\mu} = (17.65 \pm 0.24)\% \,, \tag{7}$$

$$R_\tau = (1 - B_{\tau\to e} - B_{\tau\to\mu})/B_{\tau\to e} = 3.56 \pm 0.04 \,. \tag{8}$$

The agreement is fairly good. Taking $N_c = 3$, the naive predictions only deviate from the measured values by about 20%. Many other observables, such as the partial widths of the Z and $W^\pm$ bosons, can be analyzed in a similar way to conclude that $N_c = 3$.

A particularly strong test is obtained from the $\pi^0 \to \gamma\gamma$ decay, which occurs through the triangular quark loops in Fig. 5. The crossed vertex denotes the axial current $A^3_\mu \equiv (\bar{u}\gamma_\mu\gamma_5 u - \bar{d}\gamma_\mu\gamma_5 d)$. One gets:

$$\Gamma(\pi^0 \to \gamma\gamma) = \left(\frac{N_c}{3}\right)^2 \frac{\alpha^2 m_\pi^3}{64\pi^3 f_\pi^2} = 7.73\,\mathrm{eV}, \tag{9}$$

where the π^0 coupling to A^3_μ, $f_\pi = 92.4$ MeV, is known from the $\pi^- \to \mu^-\bar{\nu}_\mu$ decay rate (assuming isospin symmetry). The agreement with the measured value, $\Gamma = 7.7 \pm 0.6$ eV [1], is remarkable. With $N_c = 1$, the prediction would have failed by a factor of 9. The nice thing about this decay is that it is associated with an *anomaly*: a global flavour symmetry which is broken by quantum effects (the triangular loops). One can then proof that the decay amplitude (9) does not get corrected by strong interactions [4].

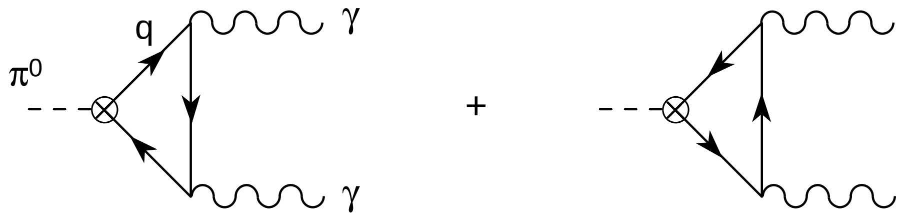


Figure 5: Triangular quark loops generating the decay $\pi^0 \to \gamma\gamma$.

Anomalies provide another compelling theoretical reason to adopt $N_c = \mathbf{3}$. The gauge symmetries of the Standard Model of electroweak interactions have also anomalies associated with triangular fermion loops (diagrams of the type shown in Fig. 5, but with arbitrary gauge bosons $-W^{\pm}, Z, \gamma-$ in the external legs and Standard Model fermions in the internal lines). These gauge anomalies are deathly because they destroy the renormalizability of the theory. Fortunately, the sum of all possible triangular loops cancels if $N_c = 3$. Thus, with three colours, anomalies are absent and the Standard Model is well-defined.

Asymptotic Freedom

The structure of the proton can be probed through the scattering $e^- p \to e^- p$. The cross-section is given by

$$\frac{d\sigma}{dQ^2} = \frac{\pi\alpha^2 \cos^2\frac{\theta}{2}}{4E^2 \sin^4\frac{\theta}{2} EE'} \left\{ \frac{|G_E(Q^2)|^2 + \frac{Q^2}{4M_p^2}|G_M(Q^2)|^2}{1 + \frac{Q^2}{4M_p^2}} + \frac{Q^2}{2M_p^2}|G_M(Q^2)|^2 \tan^2\frac{\theta}{2} \right\}, \tag{10}$$

where E and E' are the energies of the incident and scattered electrons, respectively, in the proton rest-frame, θ the scattering angle, M_p the proton mass and $Q^2 \equiv -q^2 = 4EE' \sin^2\frac{\theta}{2}$, with $q^\mu \equiv (k_e - k_e')^\mu$ the momentum transfer through the intermediate photon propagator.

G_E and G_M are the electric and magnetic form factors, respectively, describing the proton electromagnetic structure; they would be equal to one for a pointlike spin-$\frac{1}{2}$ target. Experimentally they are known to be very well approximated by the dipole form

$$G_M(Q^2)/\mu_p \approx G_E(Q^2) \approx \left(1 + \frac{Q^2}{0.7\,\mathrm{GeV}^2}\right)^{-2}, \tag{11}$$

where $\mu_p = 2.79$ is the proton magnetic moment (in proton Bohr magneton units). Thus, the proton is actually an extended object with a size of the order of 1 fm. At very low energies ($Q^2 << 1\,\mathrm{GeV}^2$), the photon probe is unable to get information on the proton structure, $G_{M,E}(Q^2) \approx G_{M,E}(0) = 1$, and the proton behaves as a pointlike particle. At higher energies, the photon is sensitive to shorter distances; the proton finite size gives then rise to form factors, which suppress the elastic cross-section at large Q^2, i.e. at large angles.

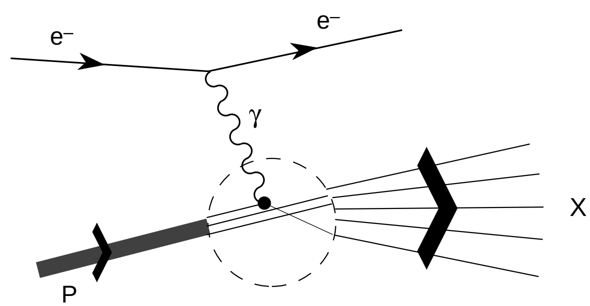


Figure 6: Inelastic $e^- p \to e^- X$ scattering.

One can try to further resolve the proton structure, by increasing the incident energy. The inelastic scattering $e^- p \to e^- X$ becomes then the dominant process. Making an inclusive sum over all hadrons produced, one has an additional kinematical variable corresponding to the final hadronic mass, $W^2 \equiv P_X^2$. The scattering is usually described in terms of Q^2 and

$$\nu \equiv \frac{(P \cdot q)}{M_p} = \frac{Q^2 + W^2 - M_p^2}{2M_p} = E - E' \,, \tag{12}$$

where P^μ is the proton cuadrimomentum; ν is the energy transfer in the proton rest-frame. In the one-photon approximation, the unpolarized differential cross-section is given by

$$\frac{d\sigma}{dQ^2\, d\nu} = \frac{\pi \alpha^2 \cos^2\frac{\theta}{2}}{4E^2 \sin^4\frac{\theta}{2} E E'} \left\{ W_2(Q^2, \nu) + 2\, W_1(Q^2, \nu)\, \tan^2\frac{\theta}{2} \right\}. \tag{13}$$

The proton structure is then characterized by two measurable *structure functions*.

For a pointlike proton, the elastic scattering (10) corresponds to

$$W_1(Q^2,\nu) = \frac{Q^2}{4M_p^2}\,\delta\!\left(\nu - \frac{Q^2}{2M_p}\right), \qquad W_2(Q^2,\nu) = \delta\!\left(\nu - \frac{Q^2}{2M_p}\right). \tag{14}$$

At low Q^2, the experimental data reveals prominent resonances; but this resonance structure quickly dies out as Q^2 increases. A much softer but sizeable continuum contribution persists at large Q^2, suggesting the existence of pointlike objects inside the proton.

To get an idea of the possible behaviour of the structure functions, one can make a very rough model of the proton, assuming that it consist of some number of pointlike spin-$\frac{1}{2}$ constituents (the so-called $partons$), each one carrying a given fraction ξ_i of the proton momenta, i.e. ${p_i}^\mu = \xi_i P^\mu$. That means that we are neglecting[1] the transverse parton momenta, and $m_i = \xi M_p$. The interaction of the photon-probe with the parton i generates a contribution to the structure functions given by:

$$W_1^{(i)}(Q^2,\nu) = \frac{e_i^2 Q^2}{4m_i^2}\,\delta\left(\nu - \frac{Q^2}{2m_i}\right) = \frac{e_i^2}{2M_p}\,\delta\left(\xi_i - x\right), \tag{15}$$

$$W_2^{(i)}(Q^2,\nu) = e_i^2\delta\left(\nu - \frac{Q^2}{2m_i}\right) = e_i^2\frac{x}{\nu}\,\delta\left(\xi_i - x\right), \tag{16}$$

where e_i is the parton electric charge and $x \equiv \frac{Q^2}{2M_p\nu} = \frac{Q^2}{Q^2+W^2-M_p^2}$. Thus, the parton structure functions only depend on the ratio x, which, moreover, fixes the momentum fractions ξ_i. We can go further, and assume that in the limit $Q^2 \to \infty$, $\nu \to \infty$, but keeping x fixed, the proton structure functions can

[1] These approximations can be made more precise going to the infinite momentum frame of the proton, where the transverse motion is negligible compared with the large longitudinal boost of the partons.

be estimated from an incoherent sum of the parton ones (neglecting any strong interactions among the partons). Denoting $f_i(\xi_i)$ the probability that the parton i has momentum fraction ξ_i, one then has:

$$W_1(Q^2,\nu) = \sum_i \int_0^1 d\xi_i \, f_i(\xi_i) \, W_1^{(i)}(Q^2,\nu) = \frac{1}{2M_p} \sum_i e_i^2 f_i(x) \equiv \frac{1}{M_p} \, F_1(x) \, , \tag{17}$$

$$W_2(Q^2,\nu) = \sum_i \int_0^1 d\xi_i \, f_i(\xi_i) \, W_2^{(i)}(Q^2,\nu) = \frac{x}{\nu} \sum_i e_i^2 f_i(x) \equiv \frac{1}{\nu} \, F_2(x) \, . \tag{18}$$

This simple parton description implies then the so-called Bjorken *scaling* [5]: the proton structure functions only depend on the kinematical variable x. Moreover, one gets the Callan–Gross relation [6] $F_2(x) = 2xF_1(x)$, which is a consequence of our assumption of spin-$\frac{1}{2}$ partons. It is easy to check that spin-0 partons would have lead to $F_1(x) = 0$.

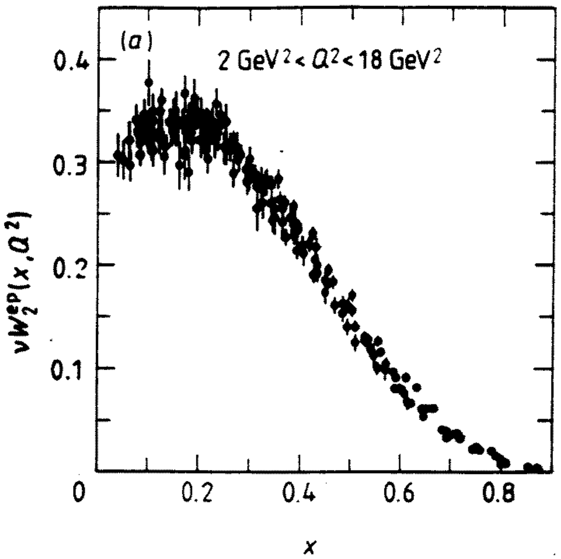


Figure 7: Experimental data on νW_2 as function of x, for different values of Q^2 [7] (taken from Ref. [8]).

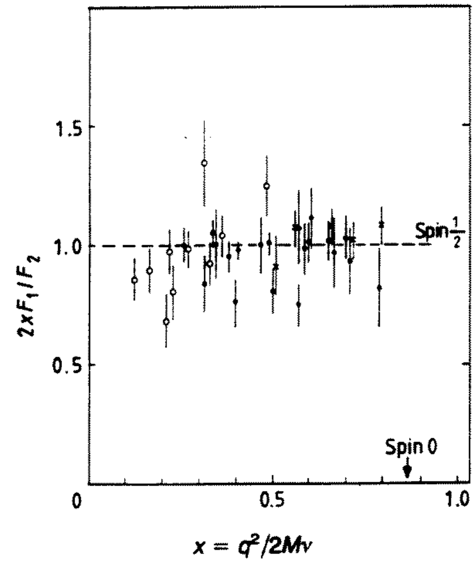


Figure 8: The ratio $2xF_1/F_2$ versus x, for different Q^2 values (1.5 GeV^2 $< Q^2 < 16$ GeV^2) [9] (taken from Ref. [8]).

The measured values of $\nu W_2(Q^2, \nu)$ are shown in Fig. 7 as function of x, for many different values of Q^2 between 2 and 18 GeV^2; the concentration of data points along a curve indicates that Bjorken scaling is correct, to a quite good approximation. Fig. 8 shows that the Callan–Gross relation is also reasonably well satisfied by the data, supporting the spin-$\frac{1}{2}$ assignment for the partons.

The surprising thing of this successful predictions is that we have assumed the existence of free independent pointlike partons inside the proton, in spite of the fact that quarks are supposed to be confined by very strong colour forces. Bjorken scaling suggests that the strong interactions must have the property of *asymptotic freedom*: they should become weaker at short distances, so that quarks behave as free particles for $Q^2 \to \infty$. This also agrees with the empirical observation in Fig. 3, that the free-quark description of the ratio $R_{e^+e^-}$ works better at higher energies.

Thus, the interaction between a $q\bar{q}$ pair looks like some kind of rubber band. If we try to separate the quark form the antiquark the force joining them increases. At some point, the energy on the elastic band is bigger than $2m_{q'}$, so that it becomes energetically favourable to create an additional $q'\bar{q}'$ pair; then the band breaks down into two mesonic systems, $q\bar{q}'$ and $q'\bar{q}$, each one with its corresponding half-band joining the quark pair. Increasing more and more the energy, we can only produce more and more mesons, but quarks remain always confined within colour-singlet bound states. Conversely, if one tries to approximate two quark constituents into a very short-distance region, the elastic band loses the energy and becomes very soft; quarks behave then as free particles.

Why $SU(3)$?

Flavour-changing transitions have a much weaker strength than processes mediated by the strong force. The quark-flavour quantum number is associated with the electroweak interactions, while strong forces appear to be flavour-conserving and flavour-independent. On the other side, the carriers of the electroweak interaction (γ, Z, $W^{\pm}$) do not couple to the quark colour. Thus, it seems natural to take colour as the charge associated with the strong forces and try to build a quantum field theory based on it [10]. The empirical evidence described so far puts a series of requirements that the fundamental theory of colour interactions should satisfy:

1. Colour is an exact symmetry G_C (hadrons do not show colour multiplicity).

2. $N_c = 3$. Thus, quarks belong to the triplet representation $\underline{3}$ of G_C.

3. Quarks and antiquarks are different states. Therefore, $\underline{3}^* \neq \underline{3}$, i.e. the triplet representation has to be complex.

4. Confinement hypothesis: hadronic states are colour singlets.

5. Asymptotic freedom.

Among all compact simple Lie groups there are only four having 3-dimensional irreducible representations; moreover, three of them are isomorphic to each other. Thus, we have only two choices: $SU(3)$ or $SO(3) \simeq SU(2) \simeq Sp(1)$. Since the triplet representation of $SO(3)$ is real, only the symmetry group $SU(3)$ survives the conditions 1, 2 and 3. The well-known $SU(3)$ decomposition of the products of $\underline{3}$ and $\underline{3}^*$ representations,

$$\begin{aligned}
q\bar{q}: &\qquad \underline{3} \otimes \underline{3}^* = \underline{1} \oplus \underline{8}\,, \\
qqq: &\qquad \underline{3} \otimes \underline{3} \otimes \underline{3} = \underline{1} \oplus \underline{8} \oplus \underline{8} \oplus \underline{10}\,, \\
qq: &\qquad \underline{3} \otimes \underline{3} = \underline{3}^* \oplus \underline{6}\,, \\
qqqq: &\qquad \underline{3} \otimes \underline{3} \otimes \underline{3} \otimes \underline{3} = \underline{3} \oplus \underline{3} \oplus \underline{3} \oplus \underline{6}^* \oplus \underline{15} \oplus \underline{15} \oplus \underline{15} \oplus \underline{15'}\,,
\end{aligned} \tag{19}$$

guarantees that there are colour-singlet configurations corresponding to meson ($q\bar{q}$) and baryon (qqq) states, as required by the confinement hypothesis. Other exotic combinations such as diquarks (qq, $\bar{q}\bar{q}$) or four-quark states ($qqqq$, $\bar{q}\bar{q}\bar{q}\bar{q}$) do not satisfy this requirement.

Clearly, the theory of colour interactions should be based on the $SU(3)_C$ group. It remains to be seen whether such a theory is able to explain confinement and asymptotic freedom as natural dynamical consequences of the colour forces.

GAUGE SYMMETRY: QED

Let us consider the Lagrangian describing a free Dirac fermion:

$$\mathcal{L}_0 = i\,\overline{\Psi}(x)\gamma^\mu\partial_\mu\Psi(x) - m\,\overline{\Psi}(x)\Psi(x)\,. \tag{20}$$

$\mathcal{L}_0$ is invariant under *global* $U(1)$ transformations

$$\Psi(x) \xrightarrow{\mathrm{U(1)}} \Psi'(x) \equiv \exp\{iQ\theta\}\,\Psi(x)\,, \tag{21}$$

where $Q\theta$ is an arbitrary real constant. The phase of $\Psi(x)$ is then a pure convention-dependent quantity without physical meaning. However, the free Lagrangian is no-longer invariant if one allows the phase transformation to depend on the space-time coordinate, i.e. under *local* phase redefinitions $\theta = \theta(x)$, because

$$\partial_\mu\Psi(x) \xrightarrow{\mathrm{U(1)}} \exp\{iQ\theta\}\,\left(\partial_\mu + iQ\partial_\mu\theta\right)\,\Psi(x)\,. \tag{22}$$

Thus, once an observer situated at the point x_0 has adopted a given phase-convention, the same convention must be taken at all space-time points. This looks very unnatural.

The "Gauge Principle" is the requirement that the $U(1)$ phase invariance should hold *locally*. This is only possible if one adds some additional piece to the Lagrangian, transforming in such a way as to cancel the $\partial_\mu\theta$ term in Eq. (22). The needed modification is completely fixed by the transformation (22): one introduces a new spin–1 (since $\partial_\mu\theta$ has a Lorentz index) field $A_\mu(x)$, transforming as

$$A_\mu(x) \xrightarrow{\mathrm{U(1)}} A'_\mu(x) \equiv A_\mu(x) + \frac{1}{e}\,\partial_\mu\theta\,, \tag{23}$$

and defines the covariant derivative

$$D_\mu\Psi(x) \equiv \left[\partial_\mu - ieQA_\mu(x)\right]\,\Psi(x)\,, \tag{24}$$

which has the required property of transforming like the field itself:

$$D_\mu \Psi(x) \stackrel{\mathrm{U(1)}}{\longrightarrow} (D_\mu \Psi)'(x) \equiv \exp\{iQ\theta\}\, D_\mu \Psi(x)\,. \tag{25}$$

The Lagrangian

$$\mathcal{L} \equiv i\,\overline{\Psi}(x)\gamma^\mu D_\mu \Psi(x) - m\,\overline{\Psi}(x)\Psi(x) = \mathcal{L}_0 + eQA_\mu(x)\,\overline{\Psi}(x)\gamma^\mu \Psi(x) \tag{26}$$

is then invariant under local $U(1)$ transformations.

The gauge principle has generated an interaction between the Dirac spinor and the gauge field A_μ, which is nothing else than the familiar QED vertex. Note that the corresponding electromagnetic charge eQ is completely arbitrary. If one wants A_μ to be a true propagating field, one needs to add a gauge-invariant kinetic term

$$\mathcal{L}_{\rm Kin} \equiv -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \, , \tag{27}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ is the usual electromagnetic field strength. A possible mass term for the gauge field, $\frac{1}{2} m^2 A^\mu A_\mu$, is forbidden because it would violate gauge invariance; therefore, the photon field is predicted to be massless. The total Lagrangian in (26) and (27) gives rise to the well-known Maxwell equations.

From a simple gauge-symmetry requirement, we have deduced the right QED Lagrangian, which leads to a very successful quantum field theory. Remember that QED predictions have been tested to a very high accuracy, as exemplified by the electron and muon anomalous magnetic moments [$a_l \equiv (g_l - 2)/2$, where

$\mu_l \equiv g_l\,(e\hbar/2m_l)]$ [11]:

$$a_e \;=\; \left\{ \begin{array}{ll} (115\,965\,214.0 \pm 2.8) \times 10^{-11} & \text{(Theory)} \\ (115\,965\,219.3 \pm 1.0) \times 10^{-11} & \text{(Experiment)} \end{array} \right. , \qquad (28)$$

$$a_\mu \;=\; \left\{ \begin{array}{ll} (1\,165\,919.2 \pm 1.9) \times 10^{-9} & \text{(Theory)} \\ (1\,165\,923.0 \pm 8.4) \times 10^{-9} & \text{(Experiment)} \end{array} \right. . \qquad (29)$$

THE QCD LAGRANGIAN

Let us denote ψ_f^α a quark field of colour α and flavour f. To simplify the equations, let us adopt a vector notation in colour space: $\psi_f \equiv \mathrm{column}(\psi_f^1, \psi_f^2, \psi_f^3)$. The free Lagrangian

$$\mathcal{L}_0 = \sum_f \bar{\psi}_f \, (i\gamma^\mu \partial_\mu - m_f) \, \psi_f \tag{30}$$

is invariant under arbitrary global $SU(3)_C$ transformations in colour space,

$$\psi_f^\alpha \longrightarrow (\psi_f^\alpha)' = U^\alpha{}_\beta \, \psi_f^\beta \, , \qquad UU^\dagger = U^\dagger U = 1 \, , \qquad \det U = 1 \, . \tag{31}$$

The $SU(3)_C$ matrices can be written in the form

$$U = \exp\left\{ -ig_s \frac{\lambda^a}{2} \theta_a \right\} , \tag{32}$$

where λ^a $(a = 1, 2, \ldots, 8)$ denote the generators of the fundamental representation of the $SU(3)_C$ algebra, and θ_a are arbitrary parameters. The matrices λ^a are traceless and satisfy the commutation relations

$$\left[\lambda^a, \lambda^b\right] = 2i f^{abc} \lambda^c , \tag{33}$$

with f^{abc} the $SU(3)_C$ structure constants, which are real and totally antisymmetric. Some useful properties of $SU(3)$ matrices are collected in Appendix A.

As in the QED case, we can now require the Lagrangian to be also invariant under *local* $SU(3)_C$ transformations, $\theta_a = \theta_a(x)$. To satisfy this requirement, we need to change the quark derivatives by covariant objects. Since we have now 8 independent gauge parameters, 8 different gauge bosons $G^\mu_a(x)$, the so-called $gluons$, are needed:

$$D^\mu \psi_f \;\equiv\; \left[\partial^\mu - i g_s \frac{\lambda^a}{2} A^\mu_a(x)\right] \psi_f \;\equiv\; \left[\partial^\mu - i g_s A^\mu(x)\right] \psi_f \,. \tag{34}$$

Notice that we have introduced the compact matrix notation

$$[A^\mu(x)]_{\alpha\beta} \;\equiv\; \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} A^\mu_a(x)\,. \tag{35}$$

We want $D^\mu \psi_f$ to transform in exactly the same way as the colour-vector ψ_f; this fixes the transformation properties of the gauge fields:

$$D^\mu \;\longrightarrow\; (D^\mu)' = U\, D^\mu\, U^\dagger\,; \qquad G^\mu \;\longrightarrow\; (G^\mu)' = U\, G^\mu\, U^\dagger - \frac{i}{g_s}\,(\partial^\mu U)\, U^\dagger\,.$$

$$(36)$$

Under an infinitesimal $SU(3)_C$ transformation,

$$\begin{aligned}
\psi_f^\alpha \quad &\longrightarrow \quad (\psi_f^\alpha)' = \psi_f^\alpha - i g_s \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} \delta\theta_a \, \psi_f^\beta \,, \\
A_a^\mu \quad &\longrightarrow \quad (A_a^\mu)' = A_a^\mu - \partial^\mu(\delta\theta_a) + g_s f^{abc} \delta\theta_b \, A_c^\mu \,.
\end{aligned} \tag{37}$$

The gauge transformation of the gluon fields is more complicated that the one obtained in QED for the photon. The non-commutativity of the $SU(3)_C$ matrices gives rise to an additional term involving the gluon fields themselves. For constant $\delta\theta_a$, the transformation rule for the gauge fields is expressed in terms of the structure constants f^{abc} only; thus, the gluon fields belong to the adjoint representation of the colour group (see Appendix A). Note also that there is a unique $SU(3)_C$ coupling g_s. In QED it was possible to assign arbitrary electromagnetic charges to the different fermions. Since the commutation relation

(33) is non-linear, this freedom does not exist for $SU(3)_C$.

To build a gauge-invariant kinetic term for the gluon fields, we introduce the corresponding field strengths:

$$\begin{aligned} F^{\mu\nu}(x) &\equiv \frac{i}{g_s}\,[D^\mu, D^\nu] \,=\, \partial^\mu A^\nu - \partial^\nu A^\mu - i g_s\,[A^\mu, A^\nu] \,\equiv\, \frac{\lambda^a}{2}\, F^{\mu\nu}_a(x)\,, \\ F^{\mu\nu}_a(x) &= \partial^\mu A^\nu_a - \partial^\nu A^\mu_a + g_s f^{abc} A^\mu_b A^\nu_c\,. \end{aligned} \tag{38}$$

Under a gauge transformation,

$$A^{\mu\nu} \,\longrightarrow\, (A^{\mu\nu})' = U\, A^{\mu\nu}\, U^\dagger\,, \tag{39}$$

and the colour trace $\mathrm{Tr}(G^{\mu\nu}G_{\mu\nu}) = \frac{1}{2}G^{\mu\nu}_a G^a_{\mu\nu}$ remains invariant.

Taking the proper normalization for the gluon kinetic term, we finally have the

$SU(3)_C$ invariant QCD Lagrangian:

$$\mathcal{L}_{\rm QCD} \equiv -\frac{1}{4} G F^{\mu\nu}_a F^a_{\mu\nu} + \sum_f \bar{\psi}_f \, (i\gamma^\mu D_\mu - m_f) \, \psi_f \, . \tag{40}$$

It is worth while to decompose the Lagrangian into its different pieces:

$$\begin{aligned} \mathcal{L}_{\rm QCD} &= -\frac{1}{4} (\partial^\mu A^\nu_a - \partial^\nu A^\mu_a)(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu) + \sum_f \bar{\psi}^\alpha_f \, (i\gamma^\mu \partial_\mu - m_f) \, \psi^\alpha_f \\ &\quad + g_s \, A^\mu_a \sum_f \bar{\psi}^\alpha_f \gamma_\mu \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} \psi^\beta_f \\ &\quad - \frac{g_s}{2} f^{abc} (\partial^\mu A^\nu_a - \partial^\nu A^\mu_a) \, A^b_\mu A^c_\nu \, - \frac{g_s^2}{4} f^{abc} f_{ade} \, A^\mu_b A^\nu_c A^d_\mu A^e_\nu \, . \end{aligned} \tag{41}$$

The first line contains the correct kinetic terms for the different fields, which give rise to the corresponding propagators. The colour interaction between quarks and gluons is given by the second line; it involves the $SU(3)_C$ matrices λ^a. Finally, owing to the non-abelian character of the colour group, the $G_a^{\mu\nu} G^a_{\mu\nu}$ term generates the cubic and quartic gluon self-interactions shown in the last line; the strength of these interactions is given by the same coupling g_s which appears in the fermionic piece of the Lagrangian. In spite of the rich physics contained in it, the Lagrangian (40) looks very simple, because of its colour-symmetry properties. All interactions are given in terms of a single universal coupling g_s, which is called the *strong coupling constant.* The existence of self-interactions among the gauge fields is a new feature that was not present in the QED case; it seems then reasonable to expect that these gauge self-interactions could explain properties like asymptotic freedom and confinement, which do not appear in QED.

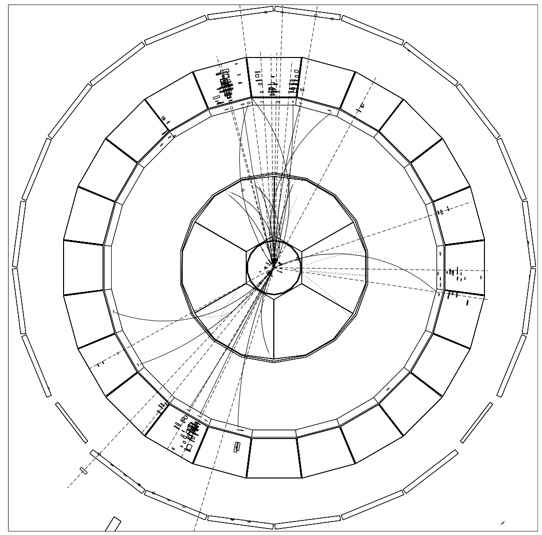


Figure 9: Three-jet event from the hadronic decay of a Z boson (DELPHI).

Without any detailed calculation, one can already extract qualitative physical consequences from $\mathcal{L}_{\mathrm{QCD}}$. Quarks can emit gluons. At lowest-order in g_s, the dominant process will be the emission of a single gauge boson. Thus, the hadronic decay of the Z should result in some $Z \to q\bar{q}G$ events, in addition to the dominant $Z \to q\bar{q}$ decays. Fig. 9 clearly shows that 3-jet events, with the required kinematics, indeed appear in the LEP data. Similar events show up in e^+e^- annihilation into hadrons, away from the Z peak.

In order to properly quantize the QCD Lagrangian, one needs to add to $\mathcal{L}_{\rm QCD}$ the so-called *Gauge-fixing* and *Faddeev–Popov* terms. Since this is a rather technical issue, its discussion is relegated to the following sections.

SUMMARY

Strong interactions are characterized by three basic properties: asymptotic freedom, confinement and dynamical chiral symmetry breaking.

Owing to the gluonic self-interactions, the QCD coupling becomes smaller at short distances, leading indeed to an asymptotically-free quantum field theory. Perturbation theory can then be applied at large momentum transfers. The resulting predictions have achieved a remarkable success, explaining a wide range of phenomena in terms of a single coupling. The running of α_s has been experimentally tested at different energy scales, confirming the predicted QCD behaviour.

The growing of the running coupling at low-energies makes very plausible that the QCD dynamics generates the required confinement of quarks and gluons into colour-singlet hadronic states. A rigorous proof of this property is, however, still lacking. At present, the dynamical details of hadronization are completely unknown.

Non-perturbative tools, such as QCD sum rules and lattice calculations, provide indirect evidence that QCD also implies the proper pattern of chiral symmetry breaking. The results obtained so far support the existence of a non-zero q-$\bar{q}$ condensate in the QCD vacuum, which dynamically breaks the chiral symmetry of the Lagrangian. However, a formal understanding of this phenomena has only been achieved in some approximate limits.

Thus, we have at present an overwhelming experimental and theoretical evidence that the $SU(3)_C$ gauge theory correctly describes the hadronic world. This makes QCD the established theory of the strong interactions. Nevertheless, the non-perturbative nature of its low-energy limit is still challenging our theoretical capabilities.

QUANTIZATION OF QCD

As was repeatedly emphasized in this notes, quantum field theory is nothing else as quantum mechanics with infinite number of degrees of freedom. We know three main methods of studying quantum mechanical systems: *i)* explicit solution of the Schrödinger equation, *ii)* quasiclassical approximation, and *iii)* perturbation theory. Speaking of the first "brute force" approach, it is in principle possible, but not very practical for a system of interacting fields: we cannot solve the Schrödinger equation analytically while solving it numerically is extremely difficult (though possible in principle with lattice methods). But the technique most widely used which allows one to obtain a lot of nontrivial results for physically observable effects is, of course, the perturbation theory. It is especially fruitful for the theories like QED where the coupling constant is small and the perturbative series converge rapidly. In many cases, perturbative expansion works well also for QCD. [2] For quantum field systems, the alias for perturbation theory is the

[2] We will see later that it is so for processes with large characteristic energy transfer.

Feynman graph technique. In this and in the following sections we will construct the diagram technique for QCD, will learn how to calculate the simplest Feynman graphs in QCD, and will understand what asymptotic freedom is.

The partition function for the Yang–Mills theory can be presented as

$$Z = \int \prod_{\vec{x},\tau} dA^a_\mu(\vec{x},\tau) \exp\left\{ -\frac{1}{2g^2} \int_0^\beta d\tau \int d\vec{x} \mathrm{Tr}\{F_{\mu\nu}F_{\mu\nu}\} \right\} \tag{1}$$

This expression is gauge invariant. Also Lorentz invariance is explicitly seen. The Minkowski path integral is obtained from Eq.(1) by the change $-1 \to i$ in the exponent:

$$Z_M = \int \prod_{\vec{x},t} dA^a_\mu(\vec{x},t) \exp\left\{ i\frac{1}{2g^2} \int d^4x \mathrm{Tr}\{F_{\mu\nu}F^{\mu\nu}\} \right\} \tag{2}$$

Feynman rules from path integral.

We assume that a reader is familiar with the standard operator way to derive the Feynman graph technique. Here we will give a brief sketch how it is done with path integrals. Consider the simplest field theory with non-trivial interaction, the $\lambda\phi^4$ theory. Its lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{m^2}{2}\phi^2 - \frac{\lambda}{24}\phi^4 \tag{3}$$

Our task is to find the elements of S–matrix — the matrix elements $< out|in >$. On the first step we make use of the *reduction formula* which relates the scattering amplitudes to the residues at the poles of the vacuum expectation value of the T–product of the Heisenberg field operators $\hat{\phi}(x) = e^{i\hat{H}t}\phi(x)e^{-i\hat{H}t}$. For example, for the scattering $p_1 p_2 \to p_3 p_4$ $(p_{i0} > 0)$, we may write

$$\int \prod_{i=1}^{4} d^4 x_i e^{-i \sum_{i=1}^{4} p_i x_i} < 0|T\{\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(x_3)\hat{\phi}(x_4)\}|0 >$$

$$= \left(\prod_{i=1,\ldots,4} \frac{i\sqrt{Z}}{p_i^2 - m_R^2 + i0} \right) \; iM_{12\to 34}\; (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4)$$

$$+ \;\; \text{less singular terms}\;, \tag{4}$$

where Z is the residue of the exact propagator at the pole:

$$\int d^4 x e^{ipx} < 0|T\{\hat{\phi}(x)\hat{\phi}(0)\}|0 > \sim \; \frac{iZ}{p^2 - m_R^2 + i0}$$

when $p^2 \sim m_R^2$, and m_R is the renormalized physical mass which does not generally coincide with the bare mass m entering the lagrangian (3). The result

(4) can be derived both in the operator and in the path integral language (see e.g. [12], Chapt. 7.2). We skip it.

Secondly, we must express the vacuum expectation value in the L.H.S. of Eq.(4) into a path integral. We already know the formula expressing a v.e.v. $< \hat{O} >$ via an Euclidean path integral. In our case, we need to find the average of the T–product of the Heisenberg operators depending explicitly on the real Minkowski time x_{i0}. Skipping the details again, we write the answer $< 0|T\{\hat{\phi}(x_1)\ldots\hat{\phi}(x_n)\}|0> =$

$$= \lim_{T\to\infty(1-i0)} \frac{\int \mathcal{D}\phi\ \phi(x_1)\ldots\phi(x_n)\exp\left\{i\int_{-T}^{T} dtd\vec{x}\mathcal{L}(x)\right\}}{\int \mathcal{D}\phi\ \exp\left\{i\int_{-T}^{T} dtd\vec{x}\mathcal{L}(x)\right\}} \tag{5}$$

where the infinitesimal imaginary shift $-i0$ ensures the dominance of the vacuum contribution and the boundary values of $\phi(-T,\vec{x})$ and $\phi(T,\vec{x})$ need not be specified: they do not affect the result.

On the third step we have to calculate actually the path integrals in Eq.(5). For theories with non-trivial interaction, that can be done only approximatively. We will do it perturbatively presenting the result as a series over the coupling constant λ. To this end, we first present $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\rm int}$ where $\mathcal{L}_0$ involves at most second powers of the fields and $\mathcal{L}_{\rm int}$ — all the rest. In our case, $\mathcal{L}_{\rm int} = -\lambda\phi^4/24$. We next expand

$$\exp\left\{-i\lambda\int\frac{\phi^4}{24}d^4x\right\} \tag{6}$$

in series in λ and express our vacuum average into the integrals

$$<\phi_1\ldots\phi_m>_0 = \frac{\int\mathcal{D}\phi\ \phi(x_1)\ldots\phi(x_m)\exp\left\{i\int d^4x\mathcal{L}_0(x)\right\}}{\int\mathcal{D}\phi\ \exp\left\{i\int dtd\vec{x}\mathcal{L}_0(x)\right\}} \tag{7}$$

($m = n + 4k$ if the $k^{\underline{\rm th}}$ term in the expansion of the exponential (6) is taken into account). The point is that the integrals (7) have Gaussian form and can be calculated analytically. [3] It is convenient to do it via the generating functional. In our case, we may write

$$< \phi_1 \ldots \phi_m >_0 = \frac{1}{Z[0]} \left(-i \frac{\delta}{\delta J(x_1)} \right) \cdots \left(-i \frac{\delta}{\delta J(x_m)} \right) Z[J] \Big|_{J=0} \tag{8}$$

where

$$\begin{aligned} Z[J] &= \int \mathcal{D}\phi \exp \left\{ i \int d^4x [\mathcal{L}_0(x) + J(x)\phi(x)] \right\} \\ &= \int \mathcal{D}\phi \exp \left\{ i \int d^4x [\frac{1}{2}\phi(x)(-\partial^2 - m^2 + i0)\phi(x) + J(x)\phi(x)] \right\} \end{aligned} \tag{9}$$

In the last line, we performed the integration by parts; the term $i0$ should be added to take into account the fact that, according to Eq.(5), the integral over

[3] The trick we use here is exactly equivalent to going into the interaction representation in the operator approach.

dt was originally done over a path somewhat below the real time axis (in the last line, we integrate just over the real axis). The shift $i0$ makes the integral (9) convergent. It can be explicitly done if defining

$$\phi \; = \; \phi' + i \int d^4 y D_F(x-y) J(y) \tag{10}$$

where $D_F(x-y)$ is the Feynman scalar Green's function

$$(\partial^2 + m^2 - i0) D_F(x-y) = -i\delta(x-y)$$

The change of variables (10) kills the linear term, the integral over $\mathcal{D}\phi'$ gives an irrelevant constant, and we obtain

$$Z[J] \; = \; Z[0] \exp\left\{ -\frac{1}{2} \int d^4 x d^4 y J(x) D_F(x-y) J(y) \right\} \tag{11}$$

Varying it according to Eq.(8), we easily obtain

$$
\begin{aligned}
& < \phi(x)\phi(0) >_0 = \ D_F(x) \\
& < \phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4) >_0 = D_F(x_1 - x_2)D_F(x_3 - x_4) \\
& + D_F(x_1 - x_3)D_F(x_2 - x_4) + D_F(x_1 - x_4)D_F(x_2 - x_3)
\end{aligned}
\tag{12}
$$

etc. i.e. the average of the product of $2p$ factors $\phi(x_i)$ presents the sum of $(2p-1)!!$ terms, each terms being the product of p Feynman Green's functions $< \phi(x_i)\phi(x_j) >_0$ with different arguments. The property (12) (the absence of non–trivial higher correlators) is a well-known property of Gaussian stochastic ensembles. When constructing the diagram technique, it plays the same role as the Wick contraction rules used in operator formalism.

To make it absolutely clear, consider an example. Let us find the one–loop correction to the propagator $< \phi(x)\phi(0) >$. In the first order in λ, the propagator is given by the ratio

$$\frac{\int \mathcal{D}\phi\phi(x)\left(1-\frac{i\lambda}{24}\int d^4y\phi^4(y)\right)\phi(0)\exp\left\{i\int d^4x\mathcal{L}_0(x)\right\}}{\int \mathcal{D}\phi\left(1-\frac{i\lambda}{24}\int d^4y\phi^4(y)\right)\exp\left\{i\int d^4x\mathcal{L}_0(x)\right\}} \tag{13}$$

Consider the numerator. The average of the product of 6 fields involves many terms corresponding to different pairing among the fields. Let first $\phi(x)$ be paired directly with $\phi(0)$. Then we have $D_F(x)$ as a common factor, and whatever it multiplies, the same factor appears also in the denominator and cancels exactly the corresponding λ – dependent terms in the numerator. This phenomenon is known as *cancellation of vacuum loops*. A non-trivial perturbative correction comes from the terms involving the products $<\phi(x)\phi(y)><\phi(y)\phi(y)><\phi(y)\phi(0)>$. There are 12 such terms, and we finally obtain

$$<\phi(x)\phi(0)>= D_F(x)-\frac{i\lambda}{2}\int d^4yD_F(x-y)D_F(0)D_F(y)+O(\lambda^2) \tag{14}$$

The second term here corresponds, of course, to the simplest diagram in Fig. 10.

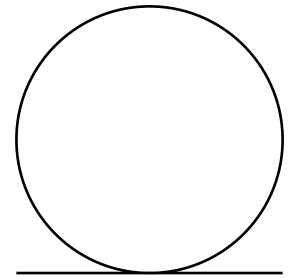


Figure 10: 1–loop correction to scalar propagator.

Fixing the gauge

To treat QCD perturbatively, we must ensure that characteristic field fluctuations are small. This is not so in the gauge fields path integral — a gauge transformation which can change the values of the fields substantially leaves the integrand invariant. All such gauge–transformed fields come on equal footing. To cope with this unwanted effect, we should somehow *fix the gauge*, i.e. impose a constraint which picks out only one representative from the whole gauge orbit, a set of field configurations differing by only a gauge transformation [4].

[4] Note that we need not be concerned here whether it is a topologically trivial or a large non–trivial gauge transformation. We are set now to study perturbative series in the coupling constant g whereas the effects due to non–trivial topology are non–analytic in g being exponentially suppressed $\propto \exp\{-8\pi^2/g^2\}$ when g is small.

In the toy oscillator model, the requirement $\phi = 0$ played the role of such gauge fixing constraint. The constrains imposed in QCD may have a lot of different forms. Different constrains (different gauges) are used for different purposes. We will discuss here Lorentz–invariant gauges and in the first place, the Landau (or Lorentz) gauge

$$\partial^\mu A_\mu^a = 0 \tag{15}$$

Let us try to write a modified path integral implementing the constraint (15) as a δ - function:

$$Z = \int \prod_{xa\mu} dA_\mu^a(x) \prod_{xa} \delta[\partial^\lambda A_\lambda^a(x)] \exp\left\{\frac{-i}{4g^2}\int d^4x F_{\mu\nu}^a F^{\mu\nu\ a}\right\} \tag{16}$$

(we may forget about θ in perturbation theory). This is what is usually done in QED. However, this is *wrong* in non–abelian case: the expression (16) is not gauge–invariant. Indeed, let us perform an infinitesimal gauge transformation.

The argument of the δ - functions in (16) is shifted by $M^{ab}(A)\chi^b$ where

$$M^{ab}(A) = \partial^2 \delta^{ab} + f^{acb} A^c_\mu \partial^\mu \equiv \partial^\mu \mathcal{D}_\mu \ , \tag{17}$$

$\mathcal{D}_\mu$ being the covariant derivative. The shift depends on A and this brings about a non-trivial Jacobian

$$J = \det \left\| \frac{\delta[\partial^\mu A^a_\mu + M^{ac}(A)\chi^c]_x}{\delta(\partial^\mu A^b_\mu)_y} \right\| \tag{18}$$

We have

$$\prod_{xa} \delta[\partial^\mu A^a_\mu + M^{ab}(A)\chi^b] \ = \ J^{-1} \prod_{xa} \delta(\partial^\mu A^a_\mu)$$

so that the path integral is not gauge–invariant, indeed. In the abelian case, the second term in $M^{ab}(A)$ was absent and the shift did not depend on A. In that case, $J = 1$ and the problem does not arise.

Faddeev and Popov were the first to realize how to write down a correct gauge–

invariant path integral with gauge–fixing condition imposed in the non–abelian case. They invented two ingenious tricks. The first trick consists in inserting the unity

$$1 \; = \; \prod_x \int d\Omega(x) \delta R(A^\Omega) \det \left\| \frac{\delta R(A^\Omega)}{\delta \Omega} \right\| \tag{19}$$

in the integrand (2) where $R(A)$ is a gauge fixing function. In our case, $R(A) = \prod_{ax} \partial_\mu A^a_\mu$. The integral in (19) is done over all gauge transformations with the Haar measure $d\Omega$. As the original expression (2) was gauge–invariant, the expression

$$\int \mathcal{D}\Omega \; \delta R(A^\Omega) \det \left\| \frac{\delta R(A^\Omega)}{\delta \Omega} \right\| \mathcal{D}A e^{iS[A]} \tag{20}$$

is $[\mathcal{D}\Omega \equiv \prod_x d\Omega(x), \quad \mathcal{D}A \equiv \prod_{xa\mu} dA^a_\mu(x)\ ]$. That allows us to change the variables $A^\Omega \to A$ in the above integral and rewrite it as

$$\int \mathcal{D}\Omega \int \mathcal{D}A\delta[R(A)] \det \left\| \frac{\delta R(A^\chi)}{\delta \chi} \right\|_{\chi=0} e^{iS[A]}$$
$$= \int \mathcal{D}A\delta[R(A)] \det \| M^{ab}(A) \| e^{iS[A]} \tag{21}$$

with $M^{ab}(A)$ from Eq. (17). The integral over $\mathcal{D}\Omega$ in the R.H.S. is lifted as the integrand does not depend on Ω anymore and $\int \mathcal{D}\Omega$ just brings about an irrelevant infinite constant.

The meaning of the transformations performed is the following: the integral (20) is done over *all* gauge fields as the integral (2) was. For each A, δ - function in Eq.(20) picks out a gauge transformation $\Omega(x)$ such that the gauge transformed field satisfies the gauge fixing constraint $R(A) = 0$. Thereby the whole range of integration in the path integral is splitted into a gauge orbits (or fibers in our fiber bundle). As the contribution of the each element of a given orbit in the path integral is the same (it is the determinant factor which takes care of it), we can suppress the integral over the orbit $d\Omega$ and write the integral in a way that only one element of the orbit satisfying the gauge fixing constraint $R(A^\Omega) = 0$ is picked out.

The expression (21) is correct, but still inconvenient. And here comes the second Faddeev and Popov trick. Let us present the determinant in Eq. (21) as a path integral over fictitious Grassmann variables $c^a(x)$ and $\bar{c}^a(x)$:

$$\det \|M^{ab}(A)\| \;=\; \int \mathcal{D}c\mathcal{D}\bar{c}\;\exp\left\{i\int d^4x\;\bar{c}^a[-\partial^2\delta^{ac} - f^{abc}A^b_\mu\partial^\mu]c^c\right\} \tag{22}$$

which is true up to an irrelevant constant factor (the choice of the sign in the exponent is a pure convention). The product of the determinant (22) and the exponential $e^{iS[A]}$ in the path integral can be written as $e^{i\tilde{S}[A]}$ where $\tilde{S}$ is a modified action involving on top of the gauge fields also fictitious scalar fermion fields $c^a(x)$:

$$\tilde{S} \;=\; \int d^4x \left[-\frac{1}{4g^2}(F^a_{\mu\nu})^2 - \bar{c}^a(\partial^\mu \mathcal{D}^{ac}_\mu)c^c\right] \tag{23}$$

we still have the δ - function $\delta(\partial^\mu A_\mu)$ in the integrand which is not too convenient. But it is handled in the same way as in QED. One can introduce a family of gauges

$$\partial^\mu A^a_\mu = \omega^a(x) \tag{24}$$

and integrate over all $\omega^a(x)$ with the weight

$$\exp\left\{-i\frac{1}{2\xi g^2}\int d^4x\ \omega^a(x)\omega^a(x)\right\}$$

. This procedure is equivalent to inserting in the path integral the weighing factor

$$\exp\left\{\frac{-i}{2\xi g^2}\int d^4x(\partial^\mu A^a_\mu)^2\right\} \tag{25}$$

and defines a general Lorentz – invariant ξ - gauge. The Landau gauge corresponds to the limit $\xi \to 0$. The determinant factor appears for each gauge in the family (24) by the same token as in the Landau gauge. It does not depend on $\omega^a(x)$. The exponent in (25) is added to the lagrangian, and we finally obtain

$$\mathcal{L}_{\rm FP} = -\frac{1}{4}(\partial^\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu)^2 - \frac{1}{2\xi}(\partial^\mu A^a_\mu)^2$$
$$-\bar{c}^a(\partial^2\delta^{ac} + gf^{abc}A^b_\mu\partial^\mu)c^c\ , \tag{26}$$

where we went into normalization $A \to gA$ convenient for perturbative analysis. The ghosts are not real physical particles. If they would be, the theory would have no sense: the hamiltonian involving scalar fermion fields would not have a ground state and unitarity would be broken. Therefore, one is not allowed to consider ghosts in the physical $|in>$ and $< out|$ asymptotic states. However, the ghosts appear with a vengeance in the loops. Loop integrals over ghosts fields produce the perturbative expansion of the Faddeev –Popov determinant. We will see a bit later that the ghosts loops are not a luxury but, in fact, it *is* necessary to take them into account to provide for unitarity of the amplitudes.

Before going further, let us make a final comment concerning the gauge invariance of the path integral. The integral

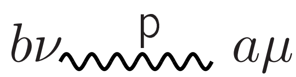
$$\int \mathcal{D}c\, \mathcal{D}\bar{c}\, \mathcal{D}A\ \exp\left\{ i \int d^4x \mathcal{L}_{\mathrm{FP}} \right\} \tag{27}$$

came from the gauge invariant path integral (2) and should be gauge invariant. This invariance is not seen explicitly: if transforming the gauge fields only, both the gauge fixing term and the ghost-ghost-gluon interaction term in Eq.(26) are varied non–trivially. One can be convinced, however, that the lagrangian (26) is invariant under a *global* symmetry transformation which acts both on the gauge and the ghost fields :

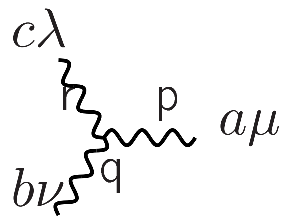
$$\begin{aligned} \delta A_\mu^a &= \epsilon(\mathcal{D}_\mu c)^a \\ \delta c^a &= -\frac{1}{2}\epsilon g f^{abc} c^b c^c \\ \delta \bar{c}^a &= -\epsilon \frac{1}{\xi}\partial^\mu A_\mu^a \ , \end{aligned} \tag{28}$$

where ϵ is an anticommuting Grassmann parameter. The invariance of (26) under (28) is a "remnant" of the gauge invariance after the gauge is fixed. It is called the BRST - symmetry. The symmetry (28) is useful when deriving generalized Ward identities (see e.g. [13]).

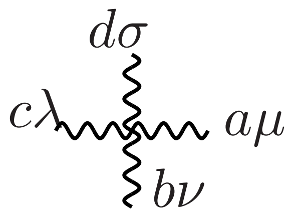
With the path integral (27) in hand, we can derive the diagram technique in the same way as was outlined for the $\lambda\phi^4$ theory. Alternatively, if one wishes, one could introduce the unphysical ghosts creation and annihilation operators and derive the Feynman rules in the operator language (one should only keep in mind that ghosts never appear as asymptotic states). The result is displayed in Eq. (29) where we also added the quark propagator and the quark–antiquark–gluon vertex coming from the term $i\bar{\psi}(\partial_\mu - igA_\mu)\psi$ in the lagrangian.



gluon propagator: $$D^{ab}_{\mu\nu}(p) \;=\; \frac{-i\delta^{ab}}{p^2+i0}\left[\eta_{\mu\nu} - \frac{(1-\xi)p_\mu p_\nu}{p^2+i0}\right]$$

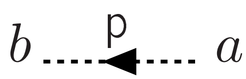


3–gluon vertex: $$\Gamma^{abc}_{\mu\nu\lambda}(p,q,r) \;=\; -gf^{abc}[(p-q)_\lambda\eta_{\mu\nu} + (q-r)_\mu\eta_{\nu\lambda} + (r-p)_\nu\eta_{\mu\lambda}]$$

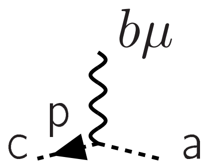


(29)

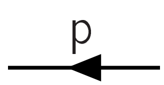
4–gluon vertex: $$\Gamma^{abcd}_{\mu\nu\lambda\sigma} \;=\; -ig^2 f^{abe}f^{cde}(\eta_{\mu\lambda}\eta_{\nu\sigma} - \eta_{\mu\sigma}\eta_{\nu\lambda}) -ig^2 f^{ace}f^{bde}(\eta_{\mu\nu}\eta_{\sigma\lambda} - \eta_{\mu\sigma}\eta_{\nu\lambda}) -ig^2 f^{ade}f^{bce}(\eta_{\mu\nu}\eta_{\sigma\lambda} - \eta_{\mu\lambda}\eta_{\nu\sigma})$$



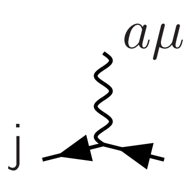
ghost propagator: $$C^{ab}(p) \quad = \quad \frac{i\delta^{ab}}{p^2 + i0}$$



$\bar{c}cg$ – vertex: $$\Gamma^{abc}(p) \quad = \quad g f^{abc} p_\mu$$



quark propagator: $$G^i_j(p) \quad = \quad \frac{i}{\not{p} - m + i0}\delta^i_j$$



quark - gluon vertex: $$(\Gamma^{a\mu})^j_i \quad = \quad i g \gamma^\mu (t^a)^j_i \tag{30}$$

Here 3–gluon and 4–gluon vertices came from the expansion of the first term in the lagrangian (26). We see that the 3–gluon vertex involves momenta coming from the derivatives $\partial_\mu A^a_\nu$ etc. The convention when all the momenta p, q, and r are outgoing so that $p+q+r=0$ is chosen. Also the ghost-ghost-gluon vertex involves the ghost momentum. The convention is that p_μ is the momentum of an outgoing ghost. All the vertices depend on the color indices of the crossing lines [adjoint indices a, b, c, d for gluons and ghosts and fundamental (antifundamental) indices i, j for quarks and antiquarks] in a non–trivial way. The external gluon lines present the transverse polarization vectors $e^{(T)}_\mu(k) \ = \ e^{(\alpha)}_\mu(k)$, $\alpha = 1, 2$, satisfying the properties $(e^{(\alpha)}_\mu)^2 = -1, \ e^{(\alpha)}_\mu(k) k^\mu = 0$. This is quite parallel to QED with the only difference that here $e_\mu(k)$ carries also the color index a which we will not display explicitly. The external quarks and antiquarks are represented by bispinors u_i and $\bar{u}^j$ carrying the fundamental or antifundamental color indices.

Ghosts and unitarity. Consider the process of annihilation of the quark–

antiquark pair into two gluons. Compared with the analogous process $e^+e^- \to 2\gamma$ in QED, it involves an extra graph with 3–gluon vertex (see Fig. 11). The amplitude of the process can be calculated according to the Feynman rules (29) and has the form $M^{\mu\nu}e_\mu^{(\alpha)}(k)e_\nu^{(\beta)}(q)$.

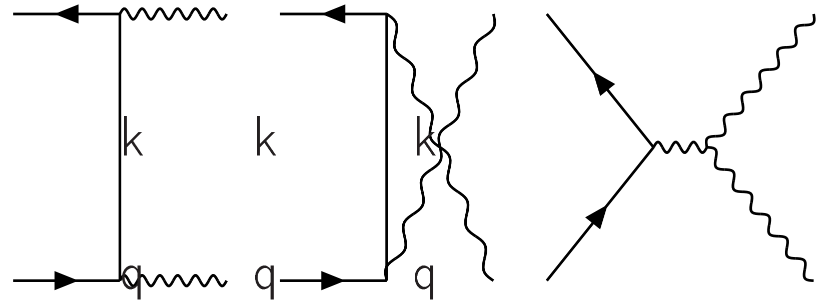


Figure 11: The process $\bar{q}q \to 2g$.

Let us first recall the situation in QED. For the process $e^+e^- \to 2\gamma$, one could write

$$iM^{\mu\nu}(k,q) \;=\; (ie_0)^2 \int d^4x d^4y e^{ikx+iqy} < 0|T\{j^\mu(x)j^\nu(y)\}|e^+e^- > \quad (31)$$

where $j^\mu(x) = \bar{e}(x)\gamma^\mu e(x)$ is the electromagnetic current, e_0 is the electron charge. In QED, the electromagnetic current is conserved: $\partial_\mu j^\mu = 0$ which implies $k_\mu M^{\mu\nu}(k,q) = q_\nu M^{\mu\nu}(k,q) = 0$. In non–abelian case, the situation is different. Indeed, the current $j^{a\mu} = \bar{q}\gamma^\mu t^a q$ satisfies only the property

$$\mathcal{D}_\mu j^{a\mu} = (\partial_\mu \delta^{ac} + f^{abc} A^b_\mu) j^{c\mu} = 0 \tag{32}$$

with covariant derivative. This relation does not describe any local conservation law [5]. Besides, $M^{\mu\nu}$ cannot be presented in the form (31) due to the third graph in Fig. 11.

Therefore, there is no reason for $k_\mu M^{\mu\nu}(k,q)$ to be zero, and it is not. If we multiply, however, $k_\mu M^{\mu\nu}$ by the transverse polarization vector of the second

[5] An analogy with general relativity can be drawn: as is well-known, in curved space a locally conserved energy–momentum tensor does not exist. The canonical energy–momentum tensor $T^{\mu\nu}$ satisfies only the property $T^{\mu\nu}_{\ ,\mu} = 0$ with covariant derivatives involving the Cristoffel symbols.

gluon, we obtain zero again:

$$k_\mu M^{\mu\nu}(k,q) e_\nu^{(T)}(q) \;=\; 0 \tag{33}$$

Indeed, an explicit calculation (do it !) reveals that

$$k_\mu M^{ab\;\mu\nu}(k,q) \;\sim\; (q^\nu q^\rho - \eta^{\nu\rho} q^2)\bar{u}\gamma_\rho t^c u f^{abc} \tag{34}$$

The term $\propto q^2$ is zero when the second gluon is on mass shell and the first term gives zero after multiplication by the transverse polarization vector $e_\nu^{(T)}(q)$. Note that the contribution of only two first graphs in Fig. 11 in the amplitude would not be transverse even in the limited sense (33); this property holds only when also the third graph is taken into account.

Generally, it is true that for any amplitude involving an ingoing or outgoing gluon with the momentum k and presentable thereby in the form $e_\mu^{(T)}(k)M^\mu(k)$ where M^μ involves also transverse polarization vectors of all other eventual external gluons, the relation $M^\mu(k)k_\mu \;=\; 0$ holds. It is one of the QCD Ward identities following from the gauge invariance and (after the gauge is fixed) the BRST symmetry (28). We will not give a general derivation here.

The physical meaning of the property (33) is the following. Besides two transverse polarizations, one could also consider two unphysical polarizations, a scalar polarization and a spatial longitudinal polarization:

$$e^{(s)} \; = \; (1, \vec{0}), \qquad e^{(l)} \; = \; \frac{1}{|\vec{k}|}(0, \vec{k}) \tag{35}$$

Or, alternatively, two light cone polarizations

$$e^{(+)}_\mu(k) \; = \; \frac{1}{\sqrt{2}|\vec{k}|}k_\mu, \qquad e^{(-)}_\mu(k) \; = \; \frac{1}{\sqrt{2}|\vec{k}|}(k_0, -\vec{k}) \tag{36}$$

The quantity $M^{\mu\nu}k_\mu e^{(T)}_\nu(q) \sim M^{\mu\nu}e^{(+)}_\mu(k)e^{(T)}_\nu(q)$ can be interpreted as the amplitude of the process of the production of transverse gluon with momentum q and a gluon with momentum k carrying an unphysical polarization $e^{(+)}_\mu(k)$.

That is something which we do not want: only transverse gluons are physical particles and nothing else can be produced in collision of quark and antiquark physical states [6]. The Ward identity (33) ensures that the amplitude of the process $\bar{q}q \to g^{(+)}g^{(T)}$ is zero. One can be convinced that the amplitude of the process $\bar{q}q \to g^{(-)}g^{(T)}$ is zero too.

This is not yet the end of the story, however. Consider the process $\bar{q}q \to g^{(+)}g^{(-)}$. Multiplying the tensor (34) by the vector $e_\nu^{(-)}(q) \sim (q_0, -\vec{q})$ we find with a dismay that the amplitude of the process with creation of *two* unphysical polarization is nonzero. One could try to find a way out of the paradox postulating that one is just not allowed to consider amplitudes of such unphysical processes. That is OK as far as tree amplitudes are concerned, but the trouble strikes back when loops

[6] Though we have not discussed it yet, a reader may know that, strictly speaking, gluons (whether transverse or not) and quarks are *not* physical asymptotic states of QCD. Due to the wonderful confinement phenomenon, the spectrum of the hamiltonian in QCD involves only colorless hadron states. Confinement is an experimental fact which is not proven yet. But what is quite definite is that it is a non–perturbative phenomenon and does not show up in any finite order of perturbation theory. In perturbative QCD, quarks and transverse gluons *are* physical asymptotic states in the same sense as electrons and positrons are.

are taken into account.

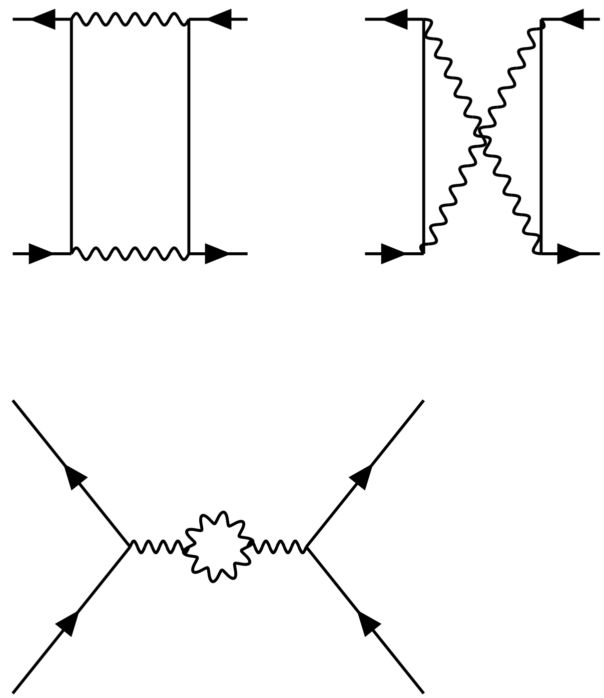


Figure 12: The process $\bar{q}q \to \bar{q}q$.

Consider the one loop contribution to the elastic zero angle scattering amplitude $\bar{q}q \to \bar{q}q$ with two–gluon intermediate state. The corresponding graphs are drawn in Fig. 12. Let us calculate its imaginary part using the Cutkosky rules so that the internal gluon lines are put on mass shell. If the graphs are calculated in the Feynman gauge, the residue of the propagators at the pole is $\eta_{\mu\nu}$ which can be decomposed as

$$\eta_{\mu\nu} = e_\mu^{(+)} e_\nu^{(-)} + e_\mu^{(-)} e_\nu^{(+)} - \sum_{\alpha=1,2} e_\mu^{(\alpha)} e_\nu^{(\alpha)} \tag{37}$$

(we choose the basis where $e_\mu^{(\alpha)}$ are real). Then the imaginary part of the elastic forward scattering amplitude is given by the sum of the cross section of the physical process $\bar{q}q \to 2g^{(T)}$ *and* the cross section of the unphysical process $\bar{q}q \to g^{(+)}g^{(-)}$. We have either to abandon the restriction that only physical polarizations are created (and thereby gauge invariance) or the optical theorem and thereby unitarity.

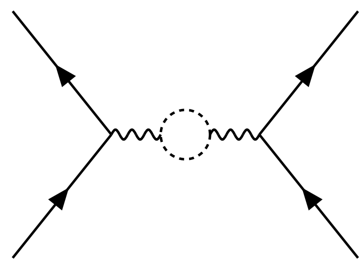


Figure 13: Ghost loop contribution in $M_{\bar{q}q\to\bar{q}q}$.

This paradox is cured by the ghosts. Consider another contribution in the amplitude drawn in Fig. 13 and involving the ghost loop. Its imaginary part due to the two–ghost "intermediate state" is non-zero. It is related to the "cross section" of the process $\bar{q}q \to \bar{c}c$. An explicit calculation shows that this "cross section" is negative and cancels exactly the positive "cross section" of another unphysical process $\bar{q}q \to g^{(+)}g^{(-)}$. Thereby, both contributions which cancel each other can be disregarded, only the physical cross section with transverse gluons in the final state contribute in the imaginary part, and gauge invariance and unitary can be reconciled again. A cancellation of this kind occurs actually also in QED if choosing as a basis for unphysical polarization not the light cone vectors (36) (in which case unphysical contributions to the cross section are zero right from the beginning), but the scalar and spatial longitudinal vectors (35). In that case, the "cross sections" $\sigma_{e^+e^- \to \gamma^{(s)}\gamma^{(s)}}$ and $\sigma_{e^+e^- \to \gamma^{(l)}\gamma^{(l)}}$ are not zero, and only the net contribution of all this garbage to the imaginary part $\mathrm{Im} M_{e^+e^- \to e^+e^-}$ is zero. In the non–abelian case, the cancellation occurs only when ghosts are taken into account.

PERTURBATIVE QCD. Regularization, Renormalization and β - function.

Regularization

Feynman integrals for the loop graphs diverge at large momenta. To handle this divergence, we should introduce an ultraviolet cutoff, to *regularize* the theory in ultraviolet. We should do it in a gauge–invariant way. If we will not bother to do so, huge gauge-non-invariant terms in the amplitudes would appear. For example, a simplistic momentum cutoff (which breaks gauge invariance) brings about a quadratically divergent gluon mass (and a quadratically divergent photon mass in QED).

Many gauge–invariant regularization procedures exit. The most "politically correct" one is probably the lattice regularization. As we have seen, the path integral symbol can be attributed a meaning (also beyond perturbation theory!)

if discretizing the space-time. The finite lattice spacing $a = \Lambda_{\mathrm{UV}}^{-1}$ serves as an ultraviolet cutoff. One *can* calculate loops in the Yang–Mills theory with lattice ultraviolet cutoff, but it is not so convenient. In particular, fixing the gauge for the lattice action brings about some (purely technical) problems.

For QED, the simplest and the most convenient regularization procedure is the Pauli-Villars procedure. It consists in subtracting from the each QED diagram involving electron-positron loops a similar diagram with loops of some extra fermions with a very large mass M. Heavy fields are irrelevant while the loop momenta are of order of physical external momenta $p_{\mathrm{char}} \ll M$. Both graphs (with electron loops and with heavy Pauli-Villars fermion loops) diverge, however, at large momenta. Their difference is finite but depends on the heavy mass M which plays the role of the ultraviolet regulator.

In the nonabelian case, the divergences come not only from quark, but also from gluon and ghost loops, and the Pauli-Villars method does not work. A possible gauge–invariant regularization procedure consists in modifying the lagrangian by adding a higher derivative term

$$\mathcal{L} \;\to\; -\frac{1}{2}\left[\mathrm{Tr}\{F_{\mu\nu}F^{\mu\nu}\} + \frac{1}{\Lambda^4}\mathrm{Tr}\{\mathcal{D}^2 F_{\mu\nu}\mathcal{D}^2 F^{\mu\nu}\}\right] \tag{38}$$

With this new term, the gluon propagator is modified according to

$$\frac{1}{k^2} \to \frac{1}{k^2 + k^6/\Lambda^4}$$

and the integrals are mostly convergent in the ultraviolet (those which are not can be handled in a Pauli–Villars–like way). The drawback of this regularization is the appearance of new vertices which makes explicit calculations complicated.

The most artificial and, physically, the less transparent but, technically, the most convenient way is the *dimensional regularization* used in the most practical calculations. It consists in changing the dimension of the space–time. Consider a typical divergent integral appearing in one–loop calculations:

$$\int \frac{d^4k}{(k^2 + M^2)^2} ,$$

where M^2 is an expression involving external momenta, Feynman parameters, and/or fermion masses. The integral diverges logarithmically at large momenta. Let us consider the same integral done over d^dk where d is an arbitrary (to start with, an integer) parameter. If $d < 4$, the integral is convergent and can be done explicitly. Using the property $d^dk \; = \; k^{d-1} dk V_{d-1}$ where

$$V_{d-1} \; = \; \frac{2\pi^{d/2}}{\Gamma(d/2)}$$

is the volume of the $(d-1)$–dimensional unit sphere, one can generally derive

$$\int \frac{d^d k}{(k^2+M^2)^n} = \frac{\pi^{d/2}\Gamma(n-d/2)}{(M^2)^{n-d/2}\Gamma(n)} \tag{39}$$

We next continue analytically this formula derived for an integer dimension d onto arbitrary real values of d and will be interested in the values of d just a little bit less than 4: $d = 4-2\epsilon$, $\epsilon \ll 1$. For $n=2$, the expression (39) develops a singularity in the limit $\epsilon \to 0$ due to the property

$$\Gamma(\epsilon) \sim \frac{1}{\epsilon} - \gamma + O(\epsilon)$$

(γ is the Euler constant) so that the original logarithmic singularity in our integral displays itself as a pole $\sim 1/\epsilon$. Note that, with dimensional regularization, also a quadratic divergence for the integral (39) with $n=1$ shows up as the same pole $[\Gamma(-1+\epsilon) = \Gamma(\epsilon)/(-1+\epsilon) \sim -1/\epsilon]$ with an extra dimensional factor $\sim M^2$; the logarithmic and the power divergence are somehow mixed up.

We have learned how to calculate scalar integrals like (39), but in practical calculations, the integrand may involve tensor structures depending on loop and external momenta. For fermion loops, also γ – matrices and their traces are present. A tensor like $\eta_{\mu\nu}$ has not a direct meaning in the space of fractional dimension. However, bearing in mind that eventually all such tensors are going to be contracted with other tensors an/or with polarization vectors of external gluons, it suffices to define formally $\eta^{\mu\nu} p_\mu q_\nu = (pq)$ and

$$\eta_{\mu\nu}\eta^{\mu\nu} = d = 4 - 2\epsilon \tag{40}$$

It is often necessary to keep here the term of order ϵ because the structure like $g_{\mu\nu} g_{\mu\nu}$ may be multiplied by a divergent integral $\sim 1/\epsilon$ and constant terms should be kept to provide for a correct gauge invariant answer.

Speaking of the γ – "matrices" in a fractional dimension space, they are defined formally as certain objects obeying the Clifford algebra

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\eta_{\mu\nu} \tag{41}$$

Bearing (41) and (40) in mind, the habitual relations of the Dirac matrices algebra are somewhat modified, e.g.

$$\begin{aligned} \gamma_\mu \gamma^\nu \gamma^\mu &= -2(1-\epsilon)\gamma^\nu \\ \gamma_\mu \gamma^\nu \gamma^\rho \gamma^\mu &= 4\eta^{\nu\rho} - 2\epsilon\gamma^\nu\gamma^\rho \end{aligned} \tag{42}$$

etc.

Problem 1. Using the Pauli–Villars regularization, prove that the photon mass is zero, indeed.

Problem 2. The same with dimensional regularization.

Renormalization.

QCD like QED is a renormalizable theory. That means that the *only* net effect of all the troublesome divergent contributions to different physical amplitudes consists in redefining the fundamental constants of the theory: electric charge and electron mass in QED and the coupling constant g and the quark masses in QCD. In other words, being expressed via a renormalized parameters $g^{\rm ren}$, $m_q^{\rm ren}$, all the physical quantities do not involve any ultraviolet divergences anymore. All such divergences are absorbed in the renormalized constants which depend on the ultraviolet cutoff Λ_0 , the bare coupling constants g_0 and m_0 and an arbitrary chosen scale μ. It is convenient to choose μ of order of characteristic energy scale of the process of interest.

The assertion of the renormalizability of QCD can be proven as an exact rigorous theorem. We will not do here, but rather provide a heuristic physical explanation. Ultraviolet divergences come from large loop momenta. When loop momenta are much larger than characteristic external momenta, the latter are not really important. Ii is instructive to consider a one–loop correction to, say, the photon propagator in the coordinate representation. Large momenta correspond to small distances in the loop. In that case, the loop "looses its structure" and can be treated as a point (see Fig.14). This point brings about a new quadratic contribution in the lagrangian which, by gauge and Lorentz invariance, is bound to have the structure $\sim (\partial_\mu A_\nu - \partial_\nu A_\mu)^2$. This *counterterm* is added to the similar structure in the tree lagrangian so that the effective value of e^2 is changed.

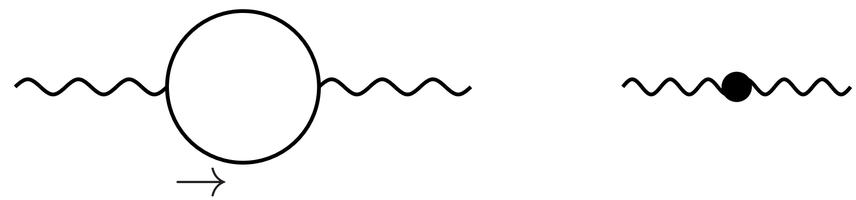


Figure 14: Appearance of counterterms.

Renormalizability is closely related to the general notion of "effective lagrangian" which is not even specific for field theories, but is heavily used also in the usual quantum and classical mechanics. Suppose we have a system whose hamiltonian involves two essentially different dimensionful parameters m and M associated with some "light" and "heavy" degrees of freedom ϕ_{light} and ϕ_{heavy}. Suppose that we are interested with the spectrum and other characteristics of the system at the energies of order $E_{\rm char} \sim m \ll M$. It is true then that these low–energy properties can be studied not with the full hamiltonian $H(\phi_{\rm light}, \phi_{\rm heavy})$, but with an *effective* hamiltonian $H^{\rm eff}(\phi_{\rm light})$ depending on the light degrees of freedom only. One can further build up a systematic expansion over the dimensionless small parameter m/M for $H^{\rm eff}(\phi_{\rm light})$ and analyze in the first place the effective hamiltonian in the leading order in this expansion.

A classical example of such an effective hamiltonian is the Born–Oppenheimer hamiltonian for a 2-atom molecule. In that case, heavy degrees of freedom were the positions and momenta of atomic electrons while the light degrees of freedom were the positions and momenta of atomic nuclei. The characteristic energies related to electronic excitations have atomic scale while a characteristic energy due to oscillations of the nuclei around their equilibrium position is much lower. The effective Born–Oppenheimer hamiltonian presents just an oscillator whose rigidity depends on how the energy of the lowest electronic term depends on the distance between the nuclei.

Another known example is the effective lagrangian (or hamiltonian) of quasi-static and quasi-homogeneous electromagnetic fields in QED. For simplicity, let the field be just static and homogeneous and purely magnetic. In that case, $\phi_{\rm light}$ is the magnetic field density $\vec{B}$ and $\phi_{\rm heavy}$ are the electron and positron field degrees of freedom. The effective hamiltonian presents a series over the dimensionless parameter $e^4\vec{B}^2/m^4$:

$$\mathcal{H}^{\rm eff} = \frac{1}{2}\vec{B}^2 - \frac{e^4}{360\pi^2 m^4}(\vec{B}^2)^2 + \cdots \tag{43}$$

The renormalization procedure is nothing else as the construction of the effective lagrangian in the Born–Oppenheimer spirit where the light degrees of freedom are field modes with the characteristic momenta of physical interest and the heavy degrees of freedom are the modes with momenta of order Λ_0. [7] There is a way

[7] To be precise, one should establish some separation scale μ and treat all the modes with momenta $p < \mu$ as light variables and the modes with momenta $p > \mu$ as heavy variables. This interpretation is due to Wilson.

to calculate the Wilsonean effective lagrangian directly using the background field method. We will use, however, a more standard approach when renormalization factors of the different structures in the lagrangian are calculated separately, and the full effective lagrangian is obtained on the second step. Being expanded in the powers of the fields, the lagrangian of Yang–Mills theory (26) (with the fermion term added) used for perturbative calculations involves quadratic, cubic and quartic terms giving rise to the gluon, ghost, and quark propagators and various vertices. When accounting for loop corrections, all these structures can in principle be renormalized:

$$
\begin{aligned}
[D^{ab}_{\mu\nu}(k)]^{\mathrm{trans.}} &= Z_g(g_0^2,\Lambda_0^2/\mu^2,\xi_0)[D^{ab\ (\mathrm{bare})}_{\mu\nu}(k)]^{\mathrm{trans}} \\
\xi(k) &= Z_\xi(g_0^2,\Lambda_0^2/\mu^2,\xi_0)\xi_0 \\
C^{ab}(k) &= Z_c(g_0^2,\Lambda_0^2/\mu^2,\xi_0)C^{ab\ (\mathrm{bare})}(k) \\
G^q(k) &= Z_q(g_0^2,\Lambda_0^2/\mu^2,\xi_0)\frac{i}{\not{p}-m^q(k)} \\
m^q(k) &= Z_m(g_0^2,\Lambda_0^2/\mu^2,\xi_0)m_0^q \\
\Gamma^{abc}_{\mu\nu\lambda}(p,q,r) &= Z^{-1}_{3g}(g_0^2,\Lambda_0^2/\mu^2,\xi_0)\Gamma^{abc\ (\mathrm{bare})}_{\mu\nu\lambda}(p,q,r) \\
\Gamma^{abcd}_{\mu\nu\lambda\sigma} &= Z^{-1}_{4g}(g_0^2,\Lambda_0^2/\mu^2,\xi_0)\Gamma^{abcd\ (\mathrm{bare})}_{\mu\nu\lambda\sigma} \\
\Gamma^{abc}_{\bar{c}cg} &= Z^{-1}_{\bar{c}cg}(g_0^2,\Lambda_0^2/\mu^2,\xi_0)\Gamma^{abc\ (\mathrm{bare})}_{\bar{c}cg} \\
\Gamma^{a}_{\mu} &= Z^{-1}_{\bar{q}qg}(g_0^2,\Lambda_0^2/\mu^2,\xi_0)\Gamma^{a\ (\mathrm{bare})}_{\mu}
\end{aligned}
\tag{44}
$$

where for the propagators we assume $k^2 = -\mu^2$ and, for the vertices, we choose a "symmetric normalization point" $k^2 = q^2 = r^2 = -\mu^2$ and similarly for the 4–gluon vertex. [8] The bare propagators and vertices are written in (29). The renormalization factors Z are just some numbers. As is written, they depend on the bare coupling g_0, bare gauge parameter ξ_0 and the dimensionless ratio Λ_0/μ. Generally, they depend also on the ratio m_0^q/μ. In this section, we will assume, however, that the quark masses are much less than the characteristic scale of interest μ and disregard this dependence. Note that the transverse part of the gluon Green's function and the longitudinal one involving the gauge parameter ξ are renormalized with their own factors. Actually, one can be convinced that the longitudinal part is not renormalized at all so that $Z_\xi = 1$. Two different spinor structures in the quark Green's function $G^{-1}(k)$ are also renormalized with their own factors, their ratio gives the mass renormalization. All the terms in the expression (29) for the 3-gluon vertex are multiplied by one and the same

[8] Green's functions at Euclidean momenta are more convenient to analyze as they do not involve imaginary parts.

renormalization factor however. This is guaranteed by the symmetry of the normalization point.

We hasten to comment that choosing the symmetric Euclidean normalization point is a pure convention. It makes calculations with several loops easier, but other conventions leading eventually to the same physical results are possible.

The Green's functions (44) "build up" an effective lagrangian

$$\mathcal{L}^{\text{eff}} = -\frac{1}{4Z_g}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)^2 - \frac{g_0}{Z_{3g}} f^{abc}(\partial_\mu A^a_\nu) A^{b\mu} A^{c\nu}$$
$$-\frac{g_0^2}{4Z_{4g}} f^{abe} f^{cde} A^a_\mu A^b_\nu A^{\mu c} A^{\nu d} + \frac{1}{Z_c}\partial_\mu \bar{c}^a \partial^\mu c^a - \frac{g_0}{Z_{\bar{c}cg}} f^{abc} \bar{c}^a A^b_\mu \partial^\mu c^c$$
$$-\frac{1}{2\xi_0}(\partial^\mu A^a_\mu)^2 + \frac{i}{Z_q}\bar{\psi} \partial\!\!\!/ \psi - \frac{m_0 Z_m}{Z_q}\bar{\psi}\psi + \frac{g_0}{Z_{\bar{q}qg}}\bar{\psi} A\!\!\!/ \psi \tag{45}$$

(with this expression in hand, the renormalized propagators and vertices (44) follow from the tree–level Feynman rules.).

To bring the kinetic terms to the standard form,it is convenient to redefine [9]

$$A \to Z_g^{1/2} A, \quad c \to Z_c^{1/2} c, \quad \psi \to Z_q^{1/2} \psi$$

Gauge invariance requires now that after that the effective lagrangian would coincide by form with the original one up to eventual renormalization of the constants. This requirement is rather rigid. It tells that the renormalization factors cannot be arbitrary, but satisfy the following *Slavnov – Taylor* identities

$$\frac{Z_g^3}{Z_{3g}^2} = \frac{Z_g^2}{Z_{4g}} = \frac{Z_g Z_c^2}{Z_{\bar{c}cg}^2} = \frac{Z_g Z_q^2}{Z_{\bar{q}qg}^2} \tag{46}$$

Their meaning is that the strength of coupling extracted from 3-gluon, 4-gluon, ghost-ghost-gluon, and quark-quark-gluon vertices coincides even after

[9] Note that such a redefinition results in that the effective gauge parameter ξ gets renormalized $\xi_0 \to \xi_0 / Z_g$.

the renormalization is performed. Any such vertex can be used to extract the renormalization factor for the effective charge given by the ratio (46)

One–loop calculations.

We will choose the $\bar{q}qg$ – vertex for this purpose. Technically, the calculations for the ghost-ghost-gluon vertex are a little bit more easy (ghosts are scalars while quarks are fermions involving extra spinor indices). The calculation with quarks is a little bit more instructive, however, because it is parallel to a similar calculation in QED, and both the similarities and differences between abelian and non–abelian theories are seen more clearly in this way. So we define the renormalized charge as

$$g^2(\mu) = Z_{\rm inv} g_0^2 = \frac{Z_g Z_q^2}{Z_{\bar{q}qg}^2} g_0^2 \qquad (47)$$

We start with calculating the renormalization factor for the gluon propagator.

The relevant graphs are drawn in Fig. 15. Z_g as such are not physical quantities and depend on the gauge. We will work in the Feynman gauge $\xi = 1$.

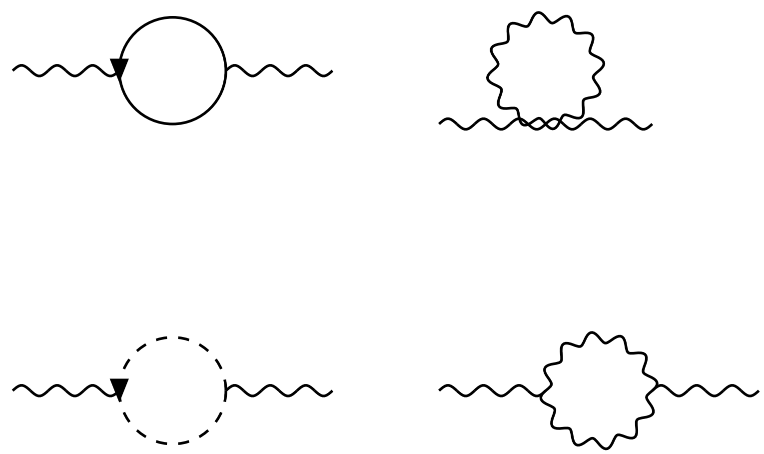


Figure 15: Gluon polarization operator

Note first that, as we have already mentioned, the longitudinal part of the gluon propagator is not renormalized:

$$k^\mu D_{\mu\nu}(k) \;=\; k^\mu D^{(\mathrm{bare})}_{\mu\nu}(k) \tag{48}$$

Actually, it is one of the Ward identities of QCD. From this and from the Dyson equation

$$D_{\mu\nu}(k) \;=\; D^{(\mathrm{bare})}_{\mu\nu}(k) - D^{(\mathrm{bare})}_{\mu\alpha}(k)\Pi^{\alpha\beta}D_{\beta\nu}(k)$$

it follows that the gluon polarization operator is transverse:

$$k_\mu \Pi^{\mu\nu}(k) = 0 \tag{49}$$

In QED the property (49) follows trivially from the current conservation. As was discussed earlier, the colored current is not conserved in QCD, but the property (49) holds nevertheless. Lorentz invariance and transversality dictate

$$\Pi^{ab}_{\mu\nu}(k) = \delta^{ab}\left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right)\Pi(k^2)$$

So, like in QED, it suffices to calculate $\Pi^{aa}_{\mu\mu}(k) = 3(N_c^2 - 1)\Pi(k^2)$ which simplifies the calculations. For $\Pi^{ab}_{\mu\nu}(k)$ to be non-singular, $\Pi(0)$ (the gluon mass) should be zero and it is. With the dimensional regularization, it is seen immediately. The graph with fermion loop gives the same expression as for QED and gives zero for $\Pi(0)$ by the same reason. Also the contributions to $\Pi(0)$ from all other graphs with massless particles in the loop are proportional, if any, to

$$\int \frac{d^4p}{p^2}$$

which, according to the rule (39), gives zero identically as soon as the dimension of the space–time is changed $4 \to 4 - 2\epsilon$.

We will calculate $\Pi(k^2)$ for massless quarks (we are interested now only in the renormalization factor which, as was mentioned does not depend on mass when the latter is small compared to characteristic momenta) via its imaginary part using the dispersion relation along the same lines as it is done in the book [14] for QED. This method is much more physical than the dimensional regularization (though its direct implication is associated with certain technical difficulties for complicated graphs in higher orders). The fact that the gluon mass is zero allows us to write for $\Pi(k^2)$ the dispersive relation with one subtraction:

$$\Pi(s) = \frac{s}{\pi} \int_0^\infty \frac{\mathrm{Im}\Pi(s')\, ds'}{s'(s' - s - i0)} \tag{50}$$

We will calculate the imaginary part as a half of the discontinuity of the polarization operator at the cut, the later being determined with the help of the Cutkosky rules when the fermions, gluons, and ghosts in the loop are put on the mass shell and their propagators are substituted by δ – functions:

$$\frac{1}{p^2 + i0} \to -2\pi i \delta(p^2)$$

Consider first the diagram with quark loop. The integrand in the corresponding

Feynman integral involves the factor

$$(i^2)_{\rm propagators}[(ig)^2]_{\rm vertices}(-1)_{\rm ferm.loop}{\rm Tr}\{t^a t^b\}{\rm Tr}\{\gamma_\mu \not{p}\gamma^\mu(\not{p}-\not{k})\} \quad (51)$$

It is the same as in QED up to the factor ${\rm Tr}\{t^a t^b\} = \frac{1}{2}\delta^{ab}$. Calculating the trace and noting that, as this expression is multiplied by $\delta(p^2)\delta[(p-k)^2]$, we can safely put $p^2 \equiv 0$ and $(pk) \equiv k^2/2$, we obtain

$$\mathrm{Im}\,\Pi \propto 4g^2[p^2 - pk]\delta^{ab} = -2g^2k^2\delta^{ab} \quad (52)$$

It is the same as in QED up to the factor ${\rm Tr}\{t^a t^b\} = \frac{1}{2}\delta^{ab}$. Restoring all the relevant numerical factors or just multiplying by $N_f/2$ the known QED result (N_f is the number of quark flavours), we obtain

$$\mathrm{Im}\,\Pi^q(s) \;=\; \frac{g^2 s N_f}{24\pi} \quad (53)$$

Let us now consider 3 other graphs in Fig. 15. The easiest is the graph with 4–gluon vertex: its imaginary part is just zero. The graphs with triple gluon vertices and with the ghost loop contribute, however. Consider first the ghost loop contribution. Calculating the imaginary part of the corresponding graph with Cutkosky rules, we obtain on the place of Eq.(52)

$$(i^2)_{\text{propagators}}(-1)_{\text{ferm.loop}} g^2 f^{dac} f^{cbd} p_\mu (p-k)^\mu = -N g^2 [p^2 - pk] \delta^{ab} \equiv \frac{g^2 N_c k^2}{2} \delta^{ab}$$

comparing with Eq.(52) and the latter with Eq.(53), we derive for the fourth term of the proportion

$$\text{Im}\, \Pi^c(s) \;=\; -\frac{g^2 N_c s}{96\pi} \tag{54}$$

The imaginary part of the ghost loop contribution has the negative sign which in the light of the discussion at the end of the previous sections is very natural: in fact, the Cutkosky trick we are using amounts to calculating the unphysical amplitude $<$ virtual gluon$(k^2)|$virtual gluon$(k^2) >$ by unitarity saturating it, in the first case, by physical quark–antiquark states and, in the second case, by unphysical ghost states. The graph in Fig. 15c) is related to the graph of Fig. 13 for the physical process $\bar{q}q \to \bar{q}q$. As was mentioned, the cross section for the production of the unphysical ghost degrees of freedom is *negative*.

Finally, let us calculate the graph with the gluon loop in Fig. 15d). Proceeding in the same way as before, we obtain for the integrand

$$\left(\frac{1}{2}\right)_{\text{symmetry}} \times (i^2)_{\text{propagators}} \times g^2 f^{acd} f^{bdc} \left[(p+k)_\beta \eta_{\mu\alpha} + (k-2p)_\mu \eta_{\alpha\beta} \right.$$

$$\left. +(p-2k)_\alpha \eta_{\mu\beta}\right]^2 = 9g^2 N_c (k^2 - pk + p^2)\delta^{ab} \equiv \frac{9g^2 N_C k^2}{2}\delta^{ab} \quad (55)$$

which gives

$$\text{Im}\,\Pi^c(s) = -\frac{3g^2 N_c s}{32\pi} \quad (56)$$

We see that the gluon loop contribution is 9 times larger than that from the ghost loop and has the same negative sign. It comes from the unphysical gluon polarizations which, according to Eq.(37), appear in the residue $g_{\mu\nu}$ of the gluon propagator in Feynman gauge. Notice that the contribution of unphysical gluon polarization and the ghosts does not cancel as it was the case for the inclusive cross section $\mathrm{Im} M_{\bar{q}q\to\bar{q}q} \propto \sigma_{\bar{q}q\to 2g} + \sigma_{\bar{q}q\to\bar{c}c}$ considered in the previous section. There is nothing wrong here: the decay of the virtual gluon is not a physical process, and there is no reason for such a cancellation to occur. Actually, the calculation we have just done is nothing else as the calculation of two of the graphs contributing to $\mathrm{Im} M_{\bar{q}q\to\bar{q}q}$: the graph in Fig. 13 and the last graph in Fig. 12; the cancellation occurs only if taking two other graphs in Fig. 12 into account.

Adding all pieces together, we get

$$\mathrm{Im}\,\Pi(s) \;=\; g^2 s \left[-\frac{5N_C c}{48\pi} + \frac{N_f}{24\pi}\right] \tag{57}$$

Substituting it in the dispersive relation (50), we finally obtain

$$Z_g \;=\; 1 + \frac{g_0^2}{48\pi^2}[5N_c - 2N_f] \ln \frac{\Lambda_0^2}{\mu^2} \tag{58}$$

If calculating the same graphs in an arbitrary ξ gauge, the result is

$$Z_g \;=\; 1 + \frac{g_0^2}{48\pi^2}\left[\frac{13-3\xi}{2}N_c - 2N_f\right] \ln \frac{\Lambda_0^2}{\mu^2} \tag{59}$$

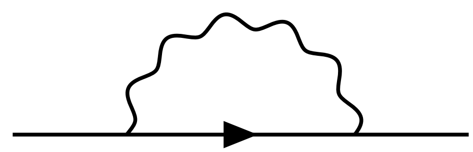


Figure 16: Quark polarization operator.

Our next task is the renormalization factors for the quark propagator and for the $\bar{q}qg$ vertex. The former can actually be found without calculation: the corresponding graph depicted in Fig. 16 has exactly the same structure as in QED up to the color factor

$$t^a t^a \;=\; \frac{N_c^2 - 1}{2N_c} \;\equiv\; c_F \;, \tag{60}$$

the Casimir eigenvalue in the fundamental representation. Taking the known result from QED, we can immediately write in the arbitrary ξ gauge

$$Z_q \;=\; 1 - \frac{g_0^2 c_F \xi}{16\pi^2} \ln \frac{\Lambda_0^2}{\mu^2}. \tag{61}$$

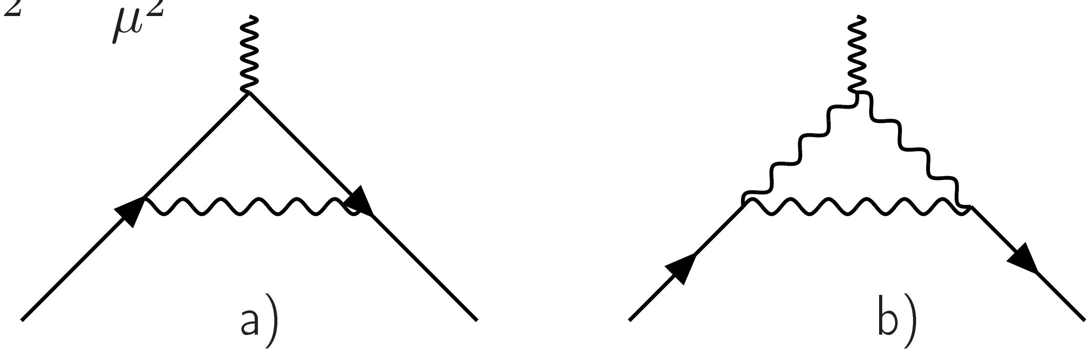


Figure 17: $\bar{q}qg$ vertex.

Consider now the renormalization of the vertex. On the 1–loop level, two graphs depicted in Fig. 17 contribute. The first graph has the same structure as in QED, again, and involves the color factor $c_F - N_c/2$ which is obtained from the equality

$$
\begin{aligned}
t^b t^a t^b &= [t^b, t^a] t^b + c_F t^a = c_F t^a - i f^{bac} t^c t^b = c_F t^a - \frac{1}{2} i f^{bac} [t^c, t^b] \\
&= (c_F - N_c/2) t^a \quad (62)
\end{aligned}
$$

The second diagram involving the 3–gluon vertex should be calculated anew. We will not do it here and just quote the result for the sum of two graphs:

$$
Z_{\bar{q}qg} = 1 - \frac{g_0^2}{16\pi^2} \left[c_F \xi + \frac{3+\xi}{4} N_c \right] \ln \frac{\Lambda_0^2}{\mu^2} \quad (63)
$$

Note that in QED $Z_e = Z_{\bar{e}e\gamma}$; this is one of QED Ward identities following from the current conservation. In QCD, the corresponding Z – factors do not coincide and their ratio contribute in the effective charge renormalization.

Substituting the Z – factors (58, 61, 63) in Eq.(47), we obtain the final result

$$g^2(\mu) = Z_{\rm inv} g_0^2 \;=\; \left[1 + \frac{g_0^2}{16\pi^2}\left(\frac{11N_c}{3} - \frac{2N_f}{3}\right)\ln\frac{\Lambda_0^2}{\mu^2} + O(g_0^4)\right] g_0^2 \quad (64)$$

The result (64) is valid only in the leading non-trivial order in g_0^2 and does not alone allow one to draw far–reaching conclusion. Pretty soon we'll derive an improved expression taking account the leading contributions in all orders in the coupling constant. A remarkable fact can, however, be observed right now: if the number of flavours is not too large [10] , the coefficient of g_0^2 in the square bracket has the opposite sign compared to QED [the QED result can be obtained from Eq.(64) if setting $N_c \to 0, \;\; N_f/2 \to 1$]. As a result, the effective charge $grows$ when μ goes down. In QED on the contrary, it decreases.

Let us recall the reason why the sign of the term $\sim e_0^2 \ln\frac{\Lambda_0^2}{\mu^2}$ in the renormalization

[10] In the real world, $N_f = 6$ or even effectively less if a characteristic scale of interest μ is less than the heavy quark masses.

factor $Z_{\rm inv}$ in QED was negative. First, in QED $Z_e = Z_{\bar{e}e\gamma}$ and the charge renormalization comes exclusively from the renormalization of the photon propagator. The renormalization factor of the propagator follows from the dispersive relation (50): the *Källen – Lehmann representation*. The imaginary part of the photon polarization operator can be expressed as

$$\mathrm{Im}\,\Pi(s) \;\propto\; -\sum_n < 0|\hat{j}_\mu|n >< n|\hat{j}^\mu|0 > \delta(P_n^2 - s) \tag{65}$$

where $\hat{j}_\mu$ is the Heisenberg operator for the electromagnetic current and the sum runs over all physical intermediate states. In QED, the current is conserved which dictates $P_n^\mu < 0|\hat{j}_\mu|n >= 0$. Going in the frame where the virtual photon is at rest, we see that the vectors $< 0|\hat{j}_\mu|n >$ are purely spacelike from which we derive that $\mathrm{Im}\,\Pi^{\rm QED}(s)$ is strictly positive; it is an exact theorem of QED [11]

[11] Another way of reasoning is noticing that $\mathrm{Im}\,\Pi^{\rm QED}(s)$, the decay probability for the virtual photon, can be related with a physical cross section. Consider some extra fermions E carrying a very small charge $e' \ll e$. In the leading order in e'^2, the inclusive cross section $\sigma_{\bar{E}E\to{\rm all}}(s)$ (with *all* orders in e^2 taken into account) is

Thereby, $Z_{\rm inv} = Z_\gamma < 1$ in QED and the physical charge is necessarily smaller that the bare one.

Due to the lack of current conservation on one hand and, on the other hand, impossibility to introduce some extra quarks with small color charge (in non–abelian case, the latter is in a sense quantized, the ordinary quarks having the minimal one corresponding to the fundamental representation), this kind of reasoning does not apply to QCD and $\mathrm{Im}\ \Pi(s)$ does not need to be positive. In fact, we have seen that it is not due to contribution of unphysical polarizations when calculating it with Cutkosky rules. Besides, $Z_q \neq Z_{\bar{q}qg}$ which affects the charge renormalization. Nothing prevents $Z_{\rm inv}$ to be greater than unity and we have seen, indeed, that it is *greater* than 1 on the one–loop level.

Renormalization group. Asymptotic freedom and infrared slavery.

The result (64) can be trusted when the correction $\sim g_0^2$ is small. Note, however,

---

proportional to $\mathrm{Im}\ \Pi^{\rm QED}(s)$. Needless to say, a physical cross section is always positive.

that even if g_0^2 is small, the correction can be of order one or larger due to the presence of the large logarithm

$$L = \ln \frac{\Lambda_0^2}{\mu^2}$$

as a factor. In that case, higher loop corrections are of the same order as the first one and we are in a position to take them into account.

Fortunately, this can be done. There are two ways to do it. The first way is to single out accurately the *leading logarithmic contributions* $\sim (g_0^2)^n L^n$ in the relevant graphs in all orders in perturbation theory and sum up all such terms. In QED, it is relatively easy. One can notice that the higher loop corrections to the photon polarization operator in QED involve only one power of L in all orders [basically, this follows from the fact (which needs to be first proven, of course) that $\mathrm{Im}\,\Pi^{\mathrm{QED}}(s)$ is finite $\propto s$ with a coefficient presenting a well–defined series over the coupling constant and from the dispersive relation (50)]. Thereby in the leading order, it suffices to consider only the one–loop graph in $\Pi(s)$ and to sum all these one–loop bubbles with the Dyson equation. As a result, we have a geometric progression

$$e^2(\mu) \;=\; \frac{e_0^2}{1 - \frac{e_0^2}{12\pi^2}\ln\frac{\mu^2}{\Lambda_0^2}} \tag{66}$$

The same program can in principle be carried out for QCD, but it is much more difficult and I even do not know whether somebody has done it explicitly. People usually use here another very powerful method known as the *renormalization group*.

Let me first quote the result. In the leading order, it is very simple and presents, again, a sum of the geometric progression

$$g^2(\mu) \;=\; \frac{g_0^2}{1+\frac{g_0^2 b_1}{16\pi^2}\ln\frac{\mu^2}{\Lambda_0^2}} \tag{67}$$

where

$$b_1 \;=\; \frac{11N_c}{3}-\frac{2}{3}N_f \tag{68}$$

Let us now derive it. The renormalizability of the theory means that all the physical results should not depend on the bare charge g_0^2 and the ultraviolet cutoff Λ_0, but only on the effective charge $g^2(\mu)$. Let us consider the effective charges $g^2(\mu)$ and $g^2(\mu')$ at two different scales and let us first assume that $\mu' \gg \mu$.

The effective charges are related to g_0^2 and Λ_0^2 according to

$$\begin{aligned}
g^2(\mu) &= Z_{\rm inv}\left(\ln\frac{\Lambda_0^2}{\mu^2},\ g_0^2\right) g_0^2 \\
g^2(\mu') &= Z_{\rm inv}\left(\ln\frac{\Lambda_0^2}{\mu'^2},\ g_0^2\right) g_0^2 \qquad (69)
\end{aligned}$$

with one and the same invariant function $Z_{\rm inv}$.

Suppose we are interested in a process with a characteristic energy scale μ and hence, eventually, in $g^2(\mu)$. A crucial observation is that we can treat the parameters μ' and $g^2(\mu')$ exactly on the same footing as Λ_0 and g_0^2. In the Wilsonean effective action spirit, we proceed here in several steps. On the $0^{\underline{th}}$ step, we define our theory with the coupling constant g_0^2 at the scale Λ_0. On the first step, we integrate over the modes with momenta greater than μ' and derive thereby a renormalized effective action on the scale μ' which, as far as the modes with momenta less than μ' are concerned, plays exactly the same role as the bare one. We can perform now the second step and integrate over the modes with momenta less than μ', but greater than μ to obtain the effective action at the scale μ and the effective charge $g^2(\mu)$. Nothing is changed except the upper and low scale from and where we are going. That means that the charge $g^2(\mu)$ will be related to $g^2(\mu')$ as

$$g^2(\mu) \;=\; Z_{\rm inv}\left(\ln\frac{\mu'^2}{\mu^2},\; g^2(\mu')\right) g^2(\mu') \tag{70}$$

with *the same* function $Z_{\rm inv}(L, g^2)$ as in Eq.(69). Substituting here the effective charges (69), we see that the function $Z_{\rm inv}$ satisfy an equation

$$Z_{\rm inv}\left(\ln\frac{\Lambda_0^2}{\mu^2},\; g_0^2\right) \;=\; Z_{\rm inv}\left(\ln\frac{\mu'^2}{\mu^2},\; g^2(\mu')\right) Z_{\rm inv}\left(\ln\frac{\Lambda_0^2}{\mu'^2},\; g_0^2\right) \tag{71}$$

This equation defines the renormalization group (or the group of multiplicative renormalizations) which just consists in tuning the effective charge when scale is changed with the factor $Z_{\rm inv}$ which depends on the *ratio* of two scales and on the charge defined at upper scale, but no explicit dependence on the scale as such appears. The elements of the group are the *functions* $Z_{\rm inv}(L, g^2)$ and the product of two such functions gives a third one.

The functional equation (71) is the main magic trick. It provides stringent restrictions on the form of the function $Z_{\rm inv}(L, g^2)$. The best way to handle Eq. (71) is to differentiate it over μ and to set $\mu = \mu'$ afterwards. This amounts to presenting one of the group elements in Eq. (71) in the infinitesimal Lee form. It is convenient to differentiate over $\ln \mu$ and not Eq. (71), but rather directly Eq. (70). We obtain

$$\mu \frac{dg^2(\mu^2)}{d\mu} \;=\; \beta[g^2(\mu^2)] \tag{72}$$

where

$$\beta[g^2(\mu^2)] = -2g^2 \,\frac{\partial}{\partial L} Z(L, g^2)\bigg|_{L=0} \tag{73}$$

is called the *Gell-Mann-Low function.* $\beta(g^2)$ presents a series over g^2. The first term of this series can be inferred from the one–loop result (64):

$$\beta(g^2) = -\frac{b_1 g^4}{8\pi^2} + O(g^6) \tag{74}$$

When only the term (74) is taken into account, the equation (72) can be easily integrated. For

$$\alpha_s = \frac{g^2}{4\pi} \ ,$$

we obtain

$$\alpha_s(\mu^2) \ = \ \frac{2\pi}{b_1 \ln \frac{\mu}{\Lambda_{\rm QCD}}} \tag{75}$$

where $\Lambda_{\rm QCD}$ is the integration constant. One can be easily convinced that

Eq.(75) coincides actually with Eq.(67) if identifying

$$\Lambda^2_{\rm QCD} \; = \; \Lambda_0^2 \exp\left\{ -\frac{16\pi^2}{b_1 g_0^2} \right\} \tag{76}$$

That means that we summed up a geometric progression, indeed. The form (75) of the result is much more illuminating, however. First of all, it does not involve the dependence of unphysical parameters Λ_0 and g_0, but rather on their combination $\Lambda_{\rm QCD}$ which is the *only* physical coupling parameter of QCD. A miracle has happened. The original Yang–Mills lagrangian involves a *dimensionless* coupling g^2. We see, however, that a *real* physical parameter of the theory is $\Lambda_{\rm QCD}$ which carries the dimension. This phenomenon is called *dimensional transmutation*. Together with the physical quark masses, $\Lambda_{\rm QCD}$ sets a scale for all relevant dimensionful quantities in QCD: in particular, to all hadron masses.

Both from Eqs.(72, 74) (telling us that the derivative of the effective charge over the scale is negative) and from their solution (75), it follows that the effective charge falls down when characteristic energy grows. It means that the larger is the energy (the smaller are the distances), the smaller is the coupling constant and hence the more trustable is the perturbation theory. This behavior is called the *asymptotic freedom.* It is just opposite to what we had in QED: the growing of charge at small distances so that eventually the perturbation theory breaks down, however small the initial large–distance charge was.

The latter property discovered first in QED by Landau, Abrikosov, and Khalatnikov and known as a *zero charge* problem is rather troublesome and means in fact, that QED seems not to be a self-consistent theory. Indeed, to define a quantum field theory, we should attribute a meaning to the path integral symbol. That can be done by putting the theory on the lattice and introducing thereby an ultraviolet cutoff. Naturally, we want that the results would not really depend on the ultraviolet cutoff in the limit when it is sent to infinity. However, in QED and in many other field theories where coupling grows at small distances ($\lambda\phi^4$ theory, Yukawa theory, etc) , we cannot do it. If we tend $\Lambda_0 \to \infty$ while keeping g_0 fixed, the effective charge (66) at any physical energy scale would go to zero. Of course, the result (66) was obtained only in perturbation theory, and does not allow to make definite conclusions concerning non-perturbative regime. But no serious reason why this trouble should be rectified in the full theory is seen, and probably it is not.

On the contrary, QCD is very nice in this respect. A continuum limit when the cutoff is sent to infinity and the bare coupling to zero exists and presents no difficulties. Everything depends on the combination (76).

There is, however, another side of the coin. According to (75), when the physical scale μ goes down, the effective charge $rises$. Eventually, at $\mu \sim \Lambda_{\rm QCD}$, it becomes of order 1, and perturbative calculations in terms of quarks and gluons lose any sense. This growth of charge is sometimes called "infrared slavery". Indeed, we know from experiment that, at large distances, there is no trace of perturbative quarks or gluons whatsoever. The confinement occurs and, instead of quarks, we have hadrons with a characteristic energy scale $\sim \Lambda_{\rm QCD}$.

Unfortunately, we cannot *prove* now that confinement (i.e. the absence of free quarks and gluons in the physical spectrum) occurs, indeed: this is a problem (a hard one) for the next century. But the connection of confinement with the infrared slavery phenomenon is obvious.

Asymptotic freedom and the infrared slavery are not, however, related so rigidly. To understand it, remind that the result (75) corresponds to the summation of the leading logarithmic terms $\propto \alpha_s^n L^n$ in the Green's function. The equation (72) allows one to sum up also next–to–leading logarithms $\propto \alpha_s^n L^{n-1}$, $\propto \alpha_s^n L^{n-2}$ etc. at almost no cost.

We need only to know the higher–order terms in the β–function expansion. According to Eq. (73), the term $\sim g^6$ in $\beta(g^2)$ and hence all the subleading terms $\propto \alpha_s^n L^{n-1}$ in Z_{inv} can be determined if the coefficients of the two–loop overlapping ultraviolet logarithm $\alpha_s^2 L$ in the Green's functions are determined. The term $\sim g^8$ in $\beta(g^2)$ gives all the subleading terms $\alpha_s^n L^{n-2}$ and is obtained from the 3–loop calculation of the terms $\sim \alpha_s^3 L$, etc.

We quote here the result of the two–loop calculations:

$$\beta(g^2) \;=\; -\frac{b_1 g^4}{8\pi^2} - \frac{b_2 g^6}{128\pi^4} \tag{77}$$

where

$$b_2 \;=\; \frac{34}{3}N_c^2 - \left(\frac{13}{3}N_c - \frac{1}{N_c}\right) N_f \tag{78}$$

Now look. When $N_c = 3$, $N_f = 6$, the first term is well negative and so is the second term. The coefficient b_1 given by (68) falls down, however, if N_f is increased: quarks provide for a conventional screening of charge like in QED and if their effect overshoots the antiscreening effect due to gluons and ghosts which happens at $N_f > 16$, the asymptotic freedom is lost. Consider an imaginary World with $N_c = 3,\ N_f = 16$. Then the first term in (77) is still negative, but very small. The second term is now positive and is not particularly small. When we start to evolve the equation (72) from *very* small distances (with *very* small coupling constant) into the infrared, the second term in (77) is at first unessential, only the first term works, and the effective charge grows. Sooner or later, the coupling will grow up to a point when the second term would balance the first one. Due to the chosen boundary value $N_f = 16$ when the coefficient b_1 is artificially small, this balancing occurs at the point when the coupling constant is rather small

$$\alpha_s^* = \frac{2\pi}{151}, \tag{79}$$

and third and higher terms in the β – function can be ignored.

The coupling constant is freezed at the *fixed point* (79) and does not grow anymore however large the distance is. In this theory, we *do* have asymptotic freedom, but do *not* have infrared slavery and confinement.

Problem. *a)* Show that the explicit solution of the differential equation (72) can be written in the form

$$\mu' = \mu \exp\left\{\int_{g^2(\mu)}^{g^2(\mu')} \frac{dy}{\beta(y)}\right\} \tag{80}$$

b) From that deduce that, in any order in coupling constant, Λ_{QCD} defined as the scale where the effective charge $g^2(\Lambda_{\mathrm{QCD}})$ becomes singular is related to Λ_0 and g_0^2 as

$$\Lambda_{\mathrm{QCD}} = \Lambda_0 \exp\left\{\int_{g_0^2}^{\infty} \frac{dy}{\beta(y)}\right\} \tag{81}$$

QCD VACUUM

Strong interactions as described by Quantum Chromodynamics (QCD) is a remarkable branch of physics where the observable entities – hadrons and nuclei – are very far from quarks and gluons in terms of which the theory is formulated. To make matters worse, the scale of strong interactions 1 fm is nowhere to be found in the QCD Lagrangian. If we restrict ourselves to hadrons 'made of' u, d, s quarks and glue, the masses of those quarks can be to a good accuracy set to zero. In this so-called chiral limit the nucleon is just 5% lighter than in reality. In the chiral limit there is not a single dimensional parameter in the QCD Lagrangian. The 1 fm scale surfaces via a mechanism named the '*transmutation of dimensions*'. QCD is a quantum field theory and being such it is not defined without introducing of some kind of ultra-violet cutoff μ. There is also a dimensionless gauge coupling constant given at that cutoff $\alpha_s(\mu)$. The dimensionful quantity Λ determining the scale of the strong interactions is the combination of μ and $\alpha_s(\mu)$:

$$\Lambda \quad = \quad \mu \exp\left(-\frac{2\pi}{b_1\alpha_s(\mu)}\right)\left(\frac{4\pi}{b_1\alpha_s(\mu)}\right)^{\frac{b_2}{2b_1^2}}\left(1+O\left(\alpha_s\right)\right), \tag{1}$$

$$b_1 \quad = \quad \frac{11}{3}N_c - \frac{2}{3}N_f, \qquad b_2 = \frac{34}{3}N_c^2 - \frac{13}{3}N_cN_f + \frac{N_f}{N_c}, \tag{2}$$

where $N_c = 3$ is the number of quark colours and N_f is the number of acting quark flavours. The ultra-violet cutoff μ sets in the dimension of mass but the exponentially small factor makes Λ much less than μ. To ensure that Λ is actually independent of the cutoff, one has to add that $\alpha_s(\mu)$ has to decrease with μ according to

$$\frac{2\pi}{\alpha_s(\mu)} = b_1 \ln\frac{\mu}{\Lambda} + \frac{b_2}{2b_1}\ln\ln\frac{\mu^2}{\Lambda^2} + O\left(\frac{1}{\ln\frac{\mu}{\Lambda}}\right). \tag{3}$$

This formula is called 'asymptotic freedom': at large scales α_s decreases.

All physical observables in strong interactions, like the nucleon mass, the pion decay constant F_π, total cross sections, etc. are proportional to Λ in the appropriate power. That is how the strong interactions scale, 1 fm, appears in the theory. One of the theory's goals is to get, say, the nucleon mass in the form of eq. (1) and to find the numerical proportionality coefficient. Doing lattice simulations the first thing one needs to check is whether an observable scales with α_s as prescribed by eq. (1). If it does not, the continuum limit is not achieved. In analytical approaches, getting an observable in the form of eq. (1) is extremely difficult. It implies doing non-perturbative physics. The only analytical approach to QCD I know of where one indeed gets observables through the transmutation of dimensions is the approach based on instantons, and it will be the subject of this paper.

Instantons are certain large non-perturbative fluctuations of the gluon field discovered by Belavin, Polyakov, Schwartz and Tyupkin in 1975 [15, 16], and the name has been suggested in 1976 by 't Hooft [17], who made a major contribution to the investigation of the instantons properties. The QCD instanton vacuum has been studied starting from the pioneering works in the end of the seventies [18, 19]; a quantitative treatment of the instanton ensemble has been developed in refs. [20, 21]. The basic ideas of the instanton vacuum are presented in section 2.

Instantons are not the only possible large non-perturbative fluctuations of the gluon field: one can think also of merons, monopoles, vortices, etc. However, instantons are the best studied non-perturbative effects. It may happen that they are not the whole truth but they are definitely present in the QCD vacuum, and they are working quite effectively in reproducing many remarkable features of the strong interactions. For example, instantons lead to the formation of the gluon condensate [22] and of the so-called topological susceptibility needed to cure the $U(1)$ paradox [17, 23]. The most striking success of instantons is their capacity to provide a beautiful microscopic mechanism of the spontaneous chiral symmetry breaking [24, 25, 26, 27]. Moreover, instantons enable one to understand it from different angles and using different mathematical formalisms. These topics are central in the review and are presented in sections 4,5 and 6.

We know that, were the chiral symmetry of QCD unbroken, the lightest hadrons would appear in parity doublets. The large actual splitting between, say, $N(\frac{1}{2}^-, 1535)$ and $N(\frac{1}{2}^+, 940)$ implies that chiral symmetry is spontaneously broken as characterized by the nonzero quark condensate $<\bar{\psi}\psi> \simeq -(250\,\mathrm{MeV})^3$. Equivalently, it means that nearly massless ('current') quarks obtain a sizable non-slash term in the propagator, called the dynamical or constituent mass $M(p)$, with $M(0) \simeq 350\,\mathrm{MeV}$. The ρ-meson has roughly twice and nucleon thrice this mass, *i.e.* are relatively loosely bound. The pion is a (pseudo) Goldstone boson and is very light. The hadron size is typically $\sim 1/M(0)$ whereas the size of constituent quarks is given by the slope of $M(p)$. In the instanton approach the former is much larger than the latter. It explains, at least on the qualitative level, why constituent quark models are so phenomenologically successful.

It should be stressed that literally speaking instantons do not lead to confinement, although they induce a growing potential for heavy quarks at intermediate separations [28]; asymptotically it flattens out [18]. However, it has been realized long ago [19, 29], that it is chiral symmetry breaking and not confinement that determines the basic characteristics of nucleons and pions as well as their first excitations. After all, 99% of the mass around us is due to the spontaneous generation of the quark constituent mass. Probably one would need an explicit confinement to get the properties of short-living highly excited hadrons. According to a popular wisdom, moving a quark away from a diquark system in a baryon generates a string, also called a flux tube, whose energy rises linearly with the separation. This is expected in the "pure-glue" world with no dynamical quarks. However, in the real world with light quarks and the spontaneous chiral symmetry breaking the string energy exceeds the pion mass $m_\pi = 140\,\mathrm{MeV}$ at a modest separation of about $0.26\,\mathrm{fm}$, see Fig. 1. At larger separations the would-be linear potential is screened since it is energetically favorable to tear the string and produce a pion. Virtually, the linear potential can stretch to as much as $0.4\,\mathrm{fm}$ where its energy exceeds $2m_\pi$ but that can happen only for a short time of $1/m_\pi$. Meanwhile, the ground-state baryons are stable, and their sizes are about $1\,\mathrm{fm}$. The pion-nucleon coupling is huge, and there seems to be no suppression of the string breaking by pions. The paradox is that the linear potential of the pure glue world, important as it might be to explain why quarks are not observed as a matter of principle, can hardly have a direct impact on the properties of lightest hadrons.

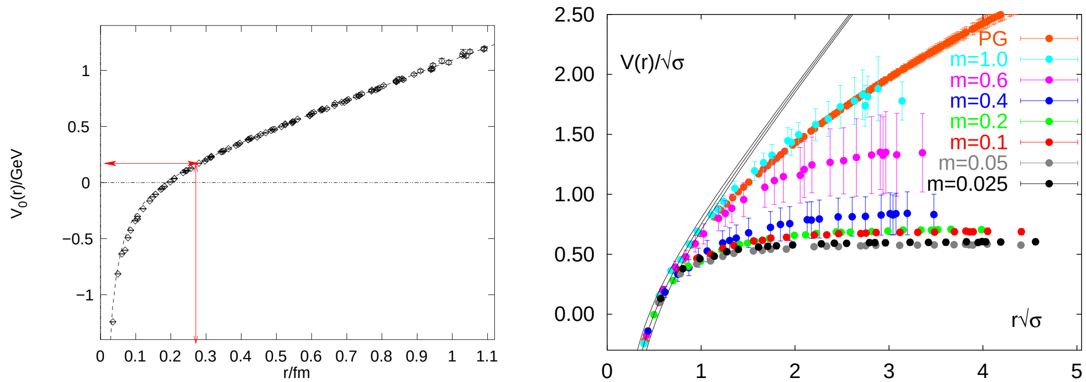


Figure 18: The lattice-simulated potential between static quarks in pure glue theory [30] exceeds m_π at the separation of $0.26\,\mathrm{fm}$ (left). The screening of the linear potential by dynamical quarks is clearly seen in simulations at high temperatures but below the phase transition [31] (right). As one lowers the pion mass the string breaking happens at smaller distances; the scale is $\sqrt{\sigma} \simeq 425\,\mathrm{MeV} \simeq (0.47\,\mathrm{fm})^{-1}$.

Even for highly excited hadrons lying on linear Regge trajectories the situation is not altogether clear. The usual explanation of resonances lying on linear trajectories is that they are rotating confining flux tubes attached to quarks at end points moving with the speed of light. Why should $350\,\mathrm{MeV}$ quarks bound by a $\sqrt{\sigma} \simeq 425\,\mathrm{MeV}$ string move with the speed of light is not clear, but if they do not, the trajectories are not linear. In this picture, the finite (and experimentally large) width of resonances is due to the same string breaking by meson production. However, if there is no string in the ground-state nucleon, why should it be excited in a collision? The lightest degrees of freedom in the real world are pions and one might expect that they are the first to be excited. An alternative explanation of resonances lying on linear trajectories is that they are rotating elongated lumps of pion field [32], and their decay is due to the normal pion radiation. It follows then that the dominant decay of a large-J baryon resonance is a cascade $\mathrm{Bar}_J \to \mathrm{Bar}_{J-1} + \pi \to \mathrm{Bar}_{J-2} + \pi\pi \to \ldots$ whereas if it is due to the string breaking it rather has a different pattern $\mathrm{Bar}_J \to \mathrm{Bar}_{\sim J/2} + \mathrm{Mes}_{\sim J/2}$. Studying the decay patterns of high-J resonances could be illuminating for understanding the relation between confining and chiral forces.

Leaving aside the unsettled question of highly excited resonances, the situation with the lightest and most important hadrons π, ρ, N, Δ... is, to my mind, clear: it is the spontaneous chiral symmetry breaking (SCSB) rather than the expected linear confining potential of the pure glue world which is the key to understanding of their properties. Therefore, since the instanton vacuum describes successfully the physics of the chiral symmetry breaking, one would expect that instantons do explain the properties of light hadrons, both mesons and baryons. Indeed, a detailed numerical study of dozens of correlation functions with different quantum numbers in the instanton medium undertaken by Shuryak, Verbaarschot and Schäfer [35] (earlier certain correlation functions were computed analytically in refs. [25, 26]) demonstrated an impressing agreement with the phenomenology [33] and with direct lattice measurements [36], see ref. [37] for a review. In fact, instantons induce strong interactions between quarks, leading to bound-state baryons with calculable and reasonable properties. There are specialized reviews on this subject, therefore I touch it only briefly here (sections 8 and 9).

More recently, there has been much activity in applying instantons to explain various phenomena in high energy processes including heavy ion collisions. For that reason, I have included section 7 which suggests a new point of view on the pomeron which might be also related to instantons.

QCD INSTANTONS

Being a quantum field theory, QCD deals with the fluctuating gluon and quark fields. A fundamental fact [38, 39] is that the potential energy of the gluon field is a periodic function in one particular direction in the infinite-dimensional functional space; in all other directions the potential energy is oscillator-like. This is illustrated in Fig. 2.

To observe this periodicity, let us temporarily work in the $A_0^a = 0$ gauge, called Weyl or Hamiltonian gauge, and forget about fermions for a while. The remaining pure Yang–Mills or "pure glue" theory is nonetheless non-trivial, since gluons are self-interacting. For simplicity I start from the $SU(2)$ gauge group.

The spatial YM potentials $A_i^a(\mathbf{x}, t)$ can be considered as an infinite set of the coordinates of the system, where $i = 1, 2, 3, \quad a = 1, 2, 3$ and $\mathbf{x}$ are "labels" denoting various coordinates. The YM action is

$$S = \frac{1}{4g^2}\int d^4x\ F^a_{\mu\nu}F^a_{\mu\nu} = \int dt\left(\frac{1}{2g^2}\int d^3\mathbf{x}\ \mathbf{E}^2 - \frac{1}{2g^2}\int d^3\mathbf{x}\ \mathbf{B}^2\right) \qquad (4)$$

where $\mathbf{E}$ is the electric field strength,

$$E^a_i(\mathbf{x},t) = \dot{A}^a_i(\mathbf{x},t) \qquad (5)$$

(the dot stands for the time derivative), and $\mathbf{B}$ is the magnetic field strength,

$$B^a_i(\mathbf{x},t) = \frac{1}{2}\epsilon_{ijk}\left(\partial_j A^a_k - \partial_k A^a_j + \epsilon^{abc}A^b_jA^c_k\right). \qquad (6)$$

Apparently, the first term in eq. (4) is the kinetic energy of the system of coordinates $\{A_i^a(\mathbf{x},t)\}$ while the second term is minus the potential energy being just the magnetic energy of the field. The simple and transparent form of eq. (5) is the advantage of the Weyl gauge. Upon quantization the electric field is replaced by the variational derivative, $E_i^a(x) \to -ig^2\delta/\delta A_i^a(x)$, if one uses the 'coordinate representation' for the wave functional. The functional Schrödinger equation for the wave functional $\Psi[A_i^a(x)]$ takes the form

$$\mathcal{H}\Psi[A_i] = \int d^3x \left\{ -\frac{g^2}{2}\frac{\delta^2}{(\delta A_i^a(x))^2} + \frac{1}{2g^2}(B_i^a(x))^2 \right\} \Psi[A_i] = \mathcal{E}\Psi[A_i] \quad (7)$$

where $\mathcal{E}$ is the eigen-energy of the state in question. The YM vacuum is the ground state of the Hamiltonian (7), corresponding to the lowest energy $\mathcal{E}$.

Let us introduce an important quantity called the Pontryagin index or the four-dimensional topological charge of the YM fields:

$$Q_T = \frac{1}{32\pi^2}\int d^4x \; F^a_{\mu\nu}\tilde{F}^a_{\mu\nu}, \qquad \tilde{F}^a_{\mu\nu} \equiv \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^a_{\alpha\beta}. \tag{8}$$

The integrand in eq. (8) happens to be a full derivative of the four-vector K_μ:

$$\frac{1}{32\pi^2}F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} = \partial_\mu K_\mu, \qquad K_\mu = \frac{1}{16\pi^2}\epsilon_{\mu\alpha\beta\gamma}\left(A^a_\alpha\partial_\beta A^a_\gamma + \frac{1}{3}\epsilon^{abc}A^a_\alpha A^b_\beta A^c_\gamma\right). \tag{9}$$

Therefore, assuming the fields A_μ are decreasing rapidly enough at spatial infinity, one can rewrite the 4-dimensional topological charge (8) as

$$Q_T = \int d^4x(\partial_0 K_0 - \partial_i K_i) = \int dt \frac{d}{dt} \int d^3\mathbf{x} K_0. \tag{10}$$

Introducing the *Chern–Simons number*

$$N_{CS} = \int d^3\mathbf{x}\, K_0 = \frac{1}{16\pi^2} \int d^3\mathbf{x}\, \epsilon^{ijk} \left(A_i^a \partial_j A_k^a + \frac{1}{3} \epsilon^{abc} A_i^a A_j^b A_k^c \right) \tag{11}$$

we see from eq. (10) that Q_T can be rewritten as the difference of the Chern–Simons numbers characterizing the fields at $t = \pm\infty$:

$$Q_T = N_{CS}(+\infty) - N_{CS}(-\infty). \tag{12}$$

The Chern–Simons number of the field has an important property that it can change by integers under large gauge transformations. Indeed, under a general time-independent gauge transformation,

$$A_i \to U^\dagger A_i U + iU^\dagger \partial_i U, \quad A_i \equiv A_i^a \frac{\tau^a}{2}, \tag{13}$$

the Chern–Simons number transforms as follows:

$$N_{CS} \to N_{CS} + N_W + \frac{i}{8\pi^2} \int d^3x \, \epsilon^{ijk} \partial_j \, \mathrm{Tr} \, (\partial_i U U^\dagger A_k). \qquad (14)$$

The last term is a full derivative and can be omitted if, *e.g.*, A_i decreases sufficiently fast at spatial infinity. N_W is the winding number of the gauge transformation (13):

$$N_W = \frac{1}{24\pi^2} \int d^3\mathbf{x} \, \epsilon^{ijk} \left[(U^\dagger \partial_i U)(U^\dagger \partial_j U)(U^\dagger \partial_k U) \right]. \qquad (15)$$

The $SU(2)$ unitary matrix U of the gauge transformation (13) can be viewed as a mapping from the 3-dimensional space onto the 3-dimensional sphere of parameters S^3. If at spatial infinity we wish to have the same matrix U independently of the way we approach the infinity (and this is what is usually assumed), then the spatial infinity is in fact one point, so the mapping is topologically equivalent to that from S^3 to S^3. This mapping is known to be non-trivial, meaning that mappings with different winding numbers are irreducible by smooth transformations to one another. The winding number of the gauge transformation is, analytically, given by eq. (15). As it is common for topological characteristics, the integrand in (15) is in fact a full derivative. For example, if we take the matrix $U(\mathbf{x})$ in a "hedgehog" form, $U = \exp[i(r \cdot \tau)/r\, P(r)]$, eq. (15) can be rewritten as

$$N_W = \frac{2}{\pi} \int dr \frac{dP}{dr} \sin^2 P = \frac{1}{\pi} \left[P - \frac{\sin 2P}{2} \right]_0^\infty = \text{integer} \qquad (16)$$

since $P(r)$ both at zero and at infinity needs to be multiples of π if we wish $U(\vec{r})$

to be unambiguously defined at the origin and at the infinity.

Let us return now to the potential energy of the YM fields,

$$\mathcal{V} = \frac{1}{2g^2} \int d^3\mathbf{x} \, (B_i^a)^2 \, . \tag{17}$$

One can imagine plotting the potential energy surfaces over the Hilbert space of the coordinates $A_i^a(\mathbf{x})$. It will be some complicated mountain country. If the field happens to be a pure gauge, $A_i = iU^\dagger \partial_i U$, the potential energy at such points of the Hilbert space is naturally zero. Imagine that we move along the "generalized coordinate" being the Chern–Simons number (11), fixing all other coordinates whatever they are. Let us take some point $A_i^a(\mathbf{x})$ with the potential energy $\mathcal{V}$. If we move to another point which is a gauge transformation of $A_i^a(\mathbf{x})$

with a winding number N_W, its potential energy will be exactly the same as it is strictly gauge invariant. However the Chern–Simons "coordinate" of the new point will be shifted by an integer N_W from the original one. We arrive to the conclusion first pointed out by Faddeev [38] and Jackiw and Rebbi [39] in 1976, that the potential energy of the YM fields is *periodic* in the particular coordinate called the Chern–Simons number.

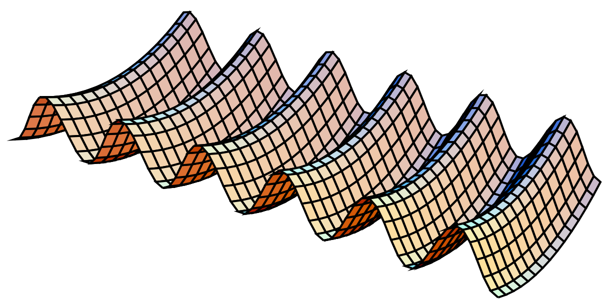


Figure 19: Potential energy of the gluon field is periodic in one direction and oscillator-like in all other directions in functional space.

Instantons in simple words

In perturbation theory one deals with zero-point quantum-mechanical fluctuations of the YM fields near one of the minima, say, at $N_{CS} = 0$. The non-linearity of the YM theory is taken into account as a perturbation, and results in series in g^2 where g is the gauge coupling. In that approach one is apparently missing a possibility for the system to tunnel to another minimum, say, at $N_{CS} = 1$. The tunneling is a typical non-perturbative effect in the coupling constant.

Instanton is a large fluctuation of the gluon field in imaginary (or Euclidean) time corresponding to quantum tunneling from one minimum of the potential energy to the neighbor one. Mathematically, it was discovered by Belavin, Polyakov, Schwarz and Tyupkin; [15] the tunneling interpretation was given by V. Gribov, see [16]. The name 'instanton' has been introduced by 't Hooft [17] who studied many of the key properties of those fluctuations. Anti-instantons are similar fluctuations but tunneling in the opposite direction in Fig. 2. Physically, one can think of instantons in two ways: on the one hand it is a tunneling *process* occurring in time, on the other hand it is a localized *pseudoparticle* in the Euclidean space.

Following the WKB approximation, the tunneling amplitude can be estimated as $\exp(-S)$, where S is the action along the classical trajectory in imaginary time, leading from the minimum at $N_{CS} = 0$ at $t = -\infty$ to that at $N_{CS} = 1$ at $t = +\infty$. According to eq. (12) the 4-dimensional topological charge of such trajectory is $Q_T = 1$. To find the best tunneling trajectory having the largest amplitude one has thus to minimize the YM action (4) provided the topological charge (8) is fixed to be unity. This can be done using the following trick [15]. Consider the inequality

$$0 \leq \int d^4x \left(F^a_{\mu\nu} - \tilde{F}^a_{\mu\nu}\right)^2$$

$$= \int d^4x \left(2F^2 - 2F\tilde{F}\right) = 8g^2 S - 64\pi^2 Q_T, \tag{18}$$

hence the action is restricted from below:

$$S \geq \frac{8\pi^2}{g^2} Q_T = \frac{8\pi^2}{g^2}. \tag{19}$$

Therefore, the minimal action for a trajectory with a unity topological charge is equal to $8\pi^2/g^2$, which is achieved if the trajectory satisfies the *self-duality* equation:

$$F^a_{\mu\nu} = \tilde{F}^a_{\mu\nu}. \tag{20}$$

Notice that any solution of eq. (20) is simultaneously a solution of the general YM equation of motion $D_\mu^{ab} F_{\mu\nu}^b = 0$: that is because the "second pair" of the Maxwell equations, $D_\mu^{ab} \tilde{F}_{\mu\nu}^b = 0$, is satisfied identically.

Thus, the tunneling amplitude can be estimated as

$$\mathcal{A} \sim \exp(-\mathrm{Action}) = \exp\left(-\frac{1}{4g^2}\int d^4x\, F_{\mu\nu}^2\right) = \exp\left(-\frac{8\pi^2}{g^2}\right) = \exp\left(-\frac{2\pi}{\alpha_s}\right). \tag{21}$$

It is non-analytic in the gauge coupling constant and hence instantons are missed in all orders of the perturbation theory. However, it is not a reason to ignore tunneling. For example, tunneling of electrons from one atom to another in a metal is also a non-perturbative effect but we would get nowhere in understanding metals had we ignored it.

Instanton configurations

To solve eq. (20) let us recall a few facts about the Lorentz group $SO(3,1)$. Since we are talking about the tunneling fields which can only develop in imaginary time, it means that we have to consider the fields in Euclidean space-time, so that the Lorentz group is just $SO(4)$ isomorphic to $SU(2)\times SU(2)$. The gauge potentials A_μ belong to the $(\frac{1}{2},\frac{1}{2})$ representation of the $SU(2)\times SU(2)$ group, while the field strength $F_{\mu\nu}$ belongs to the reducible $(1,0)+(0,1)$ representation. In other words it means that one linear combination of $F_{\mu\nu}$ transforms as a vector of the left $SU(2)$, and another combination transforms as a vector of the right $SU(2)$. These combinations are

$$F_L^A=\eta^A_{\mu\nu}(F_{\mu\nu}+\tilde{F}_{\mu\nu}),\qquad F_R^A=\bar{\eta}^A_{\mu\nu}(F_{\mu\nu}-\tilde{F}_{\mu\nu}), \tag{22}$$

where $\eta,\bar{\eta}$ are the so-called 't Hooft symbols described in ref. [17], see also below.

We see therefore that a self-dual field strength is a vector of the left $SU(2)$ while its right part is zero. Keeping that experience in mind we look for the solution of the self-dual equation in the form

$$A^a_\mu = \bar{\eta}^a_{\mu\nu}\, x_\nu\, \frac{1+\Phi(x^2)}{x^2}. \tag{23}$$

Using the formulae for the η symbols from ref. [17] one can easily check that the YM action can be rewritten as

$$S = \frac{8\pi^2}{g^2}\frac{3}{2}\int d\tau \left[\frac{1}{2}\left(\frac{d\Phi}{d\tau}\right)^2 + \frac{1}{8}(\Phi^2-1)^2\right], \qquad \tau = \ln\left(\frac{x^2}{\rho^2}\right). \tag{24}$$

This can be recognized as the action of the double-well potential whose minima lie at $\Phi = \pm 1$, and τ plays the role of "time"; ρ is an arbitrary scale. The trajectory which tunnels from 1 at $\tau = -\infty$ to -1 at $\tau = +\infty$ is

$$\Phi = -\tanh\left(\frac{\tau}{2}\right), \tag{25}$$

and its action (24) is $S = 8\pi^2/g^2$, as needed. Substituting the solution (25) into (23) we get

$$A^a_\mu(x) = \frac{2\bar{\eta}^{\mu}_{\nu a}\rho^2}{x^2(x^2+\rho^2)}. \tag{26}$$

The correspondent field strength is

$$F^a_{\mu\nu} = -\frac{4\rho^2}{(x^2+\rho^2)^2}\left(\bar\eta^\mu_{\nu a} - 2\bar\eta^\mu_{\alpha a}\frac{x_\alpha x_\nu}{x^2} - 2\bar\eta^\beta_{\nu a}\frac{x_\mu x_\beta}{x^2}\right), \qquad F^a_{\mu\nu}F^a_{\mu\nu} = \frac{192\rho^4}{(x^2+\rho^2)^4}, \tag{27}$$

and satisfies the self-duality condition (20).

The *anti-instanton* corresponding to tunneling in the opposite direction, from $N_{CS} = 1$ to $N_{CS} = 0$, satisfies the *anti*-self-dual equation, with $\tilde F \to -\tilde F$; its concrete form is given by eqs.(26, 27) with the replacement $\bar\eta \to \eta$.

Eqs.(26, 27) describe the field of the instanton in the singular Lorenz gauge; the singularity of A_μ at $x^2 = 0$ is a gauge artifact: the gauge-invariant field strength squared is smooth at the origin. Formulae for instantons are more compact in the Lorenz gauge, and I shall use it further on [12].

[12] Jackson and Okun [40] recommend to call the $\partial_\mu A_\mu = 0$ gauge by the name of the Dane Ludvig Lorenz

Instanton collective coordinates

The instanton field, eq. (26), depends on an arbitrary scale parameter ρ which we shall call the instanton size, while the action, being scale invariant, is independent of ρ. One can obviously shift the position of the instanton to an arbitrary 4-point z_μ – the action will not change either. Finally, one can rotate the instanton field in colour space by constant unitary matrices U. For the $SU(2)$ gauge group this rotation is characterized by 3 parameters, $e.g.$ by Euler angles. For a general $SU(N_c)$ group the number of parameters is $N_c^2 - 1$ (the total number of the $SU(N_c)$ generators) *minus* $(N_c - 2)^2$ (the number of generators which do not affect the left upper 2×2 corner where the standard $SU(2)$ instanton (26) is residing), that is $4N_c - 5$. These degrees of freedom are called instanton orientation in colour space. All in all there are

---

and not the Dutchman Hendrik Lorentz who certainly used this gauge too but several decades later.

$$4\,(centre) \quad + \quad 1\,(size) \quad + \quad (4N_c - 5)\,(orientations) \quad = \quad 4N_c \qquad (28)$$

so-called collective coordinates describing the field of the instanton, of which the action is independent.

It is convenient to introduce 2×2 matrices

$$\sigma^{\pm}_{\mu} = (\pm i\overrightarrow{\sigma}, 1), \qquad x^{\pm} = x_{\mu}\sigma^{\pm}_{\mu}, \qquad (29)$$

such that

$$2i\tau^{a}\eta^{\mu}_{\nu a} = \sigma^{+}_{\mu}\sigma^{-}_{\nu} - \sigma^{+}_{\nu}\sigma^{-}_{\mu}, \qquad 2i\tau^{a}\bar{\eta}^{\mu}_{\nu a} = \sigma^{-}_{\mu}\sigma^{+}_{\nu} - \sigma^{-}_{\nu}\sigma^{+}_{\mu}, \qquad (30)$$

then the instanton field with arbitrary center z_μ, size ρ and colour orientation U in the $SU(N_c)$ gauge group can be written as

$$A_\mu = A^a_\mu t^a = \frac{-i\rho^2 U[\sigma^-_\mu (x-z)^+ - (x-z)_\mu] U^\dagger}{(x-z)^2[\rho^2 + (x-z)^2]}, \qquad \mathrm{Tr}\,(t^a t^b) = \frac{1}{2}\delta^{ab}, \tag{31}$$

or as

$$A^a_\mu = \frac{2\rho^2 O^{ab} \bar\eta^\mu_{\nu b} (x-z)_\mu}{(x-z)^2[\rho^2 + (x-z)^2]}, \qquad O^{ab} = \mathrm{Tr}\,(U^\dagger t^a U \sigma^b), \qquad O^{ab} O^{ac} = \delta^{bc}.$$

$$(32)$$

This is the explicit expression for the $4N_c$-parameter instanton field in the $SU(N_c)$ gauge theory, written down in the singular Lorenz gauge.

QCD INSTANTON VACUUM

Gluon condensate

The QCD perturbation theory implies that the fields $A_i^a(\mathbf{x})$ are performing quantum zero-point oscillations; in the lowest order these are just plane waves with arbitrary frequencies. The aggregate energy of these zero-point oscillations, $(\mathbf{B}^2+\mathbf{E}^2)/2$, is divergent as the fourth power of the cutoff frequency, however for any state one has $\langle F_{\mu\nu}^2\rangle = 2\langle \mathbf{B}^2-\mathbf{E}^2\rangle = 0$, which is just a manifestation of the virial theorem for harmonic oscillators: the average potential energy is equal the kinetic one (I am temporarily in the Minkowski space). One can prove that this is also true in any order of the perturbation theory in the coupling constant, provided one does not violate the Lorentz symmetry and the renormalization properties of the theory. Meanwhile, we know from the QCD sum rules phenomenology that the QCD vacuum possesses what is called *gluon condensate* [22]:

$$\frac{1}{32\pi^2}\langle F^a_{\mu\nu}F^a_{\mu\nu}\rangle = \frac{1}{16\pi^2}\langle \mathbf{B}^2 - \mathbf{E}^2\rangle \simeq (200\ MeV)^4 \ \ > \ 0. \tag{33}$$

Instantons suggest an immediate explanation of this basic property of QCD. Indeed, instanton is a tunneling process, it occurs in imaginary time; therefore in Minkowski space one has $E_i^a = \pm i B_i^a$ (this is actually the duality eq. (20)). Therefore, during the tunneling $\mathbf{B}^2 - \mathbf{E}^2$ is positive, and one gets a chance to explain the gluon condensate. In Euclidean space the electric field is real as well as the magnetic one, and the gluon condensate is just the average action density. Let us make a quick estimate of its value. Let the total number of instantons and anti-instantons (henceforth I's and $\bar{I}$'s for short) in the 4-dimensional volume V be N. Assuming that the average separations of instantons are larger than their average sizes (to be justified below), we can estimate the total action of the ensemble as the sum of individual actions (see eq. (19)):

$$\langle F_{\mu\nu}^2 \rangle V = \int d^4 x F_{\mu\nu}^2 \simeq N \cdot 32\pi^2, \tag{34}$$

hence the gluon condensate is directly related to the instanton density in the 4-dimensional Euclidean space-time:

$$\frac{1}{32\pi^2} \langle F_{\mu\nu}^a F_{\mu\nu}^a \rangle \simeq \frac{N}{V} \equiv \frac{1}{\bar{R}^4}. \tag{35}$$

In order to get the phenomenological value of the condensate one needs thus to have the average separation between pseudoparticles [22, 19]

$$\bar{R} \simeq \frac{1}{200\,\mathrm{MeV}} = 1\,\mathrm{fm}. \tag{36}$$

There is another point of view on the gluon condensate which I describe briefly. In principle, all information about field theory is contained in the partition function being the functional integral over the fields. In the Euclidean formulation it is

$$\mathcal{Z} = \int DA_\mu \exp\left(-\frac{1}{4g^2}\int d^4x F_{\mu\nu}^2\right) \stackrel{T\to\infty}{\longrightarrow} e^{-\mathcal{E}T}, \tag{37}$$

where I have used that at large (Euclidean) time T the partition function picks up the ground state or vacuum energy $\mathcal{E}$. For the sake of brevity I do not write the gauge fixing and Faddeev–Popov ghost terms. If the state is homogeneous, the energy can be written as $\mathcal{E} = \theta_{44}V^{(3)}$ where $\theta_{\mu\nu}$ is the stress-energy tensor and $V^{(3)}$ is the 3-volume of the system. Hence, at large 4-volumes $V = V^{(3)}T$ the partition function is $\mathcal{Z} = \exp(-\theta_{44}V)$. This θ_{44} includes zero-point oscillations and diverges badly. A more reasonable quantity is the partition function, normalized to the partition function understood as a perturbative expansion about the zero-field vacuum[13],

$$\frac{\mathcal{Z}}{\mathcal{Z}_{\mathrm{P.T.}}} = \exp\left[-(\theta_{44} - \theta_{44}^{\mathrm{P.T.}})V\right]. \tag{38}$$

We expect that the non-perturbative vacuum energy density $\theta_{44} - \theta_{44}^{\mathrm{P.T.}}$ is a negative quantity since we have allowed for tunneling: as usual in quantum mechanics, it lowers the ground state energy. If the vacuum is isotropic, one has $\theta_{44} = \theta_{\mu\mu}/4$. Using the trace anomaly,

$$\theta_{\mu\mu} = \frac{\beta(\alpha_s)}{16\pi\alpha_s^2}\left(F_{\mu\nu}^a\right)^2 \simeq -b\frac{F_{\mu\nu}^2}{32\pi^2}, \tag{39}$$

[13] The latter can be distinguished from the former by imposing a condition that it does not contain integration over singular Yang–Mills potentials; recall that the instanton potentials are singular at the origins.

where $\beta(\alpha_s)$ is the Gell-Mann–Low function,

$$\beta(\alpha_s) \equiv \frac{d\alpha_s(\mu)}{d\ln\mu} = -b_1\frac{\alpha_s^2(\mu)}{2\pi} - \frac{b_2}{2}\frac{\alpha_s^3(\mu)}{(2\pi)^2} - ..., \tag{40}$$

with $b_{1,2}$ given by eq. (2), one gets [21]:

$$\frac{\mathcal{Z}}{\mathcal{Z}_{\rm P.T.}} = \exp\left(\frac{b}{4}V\langle F_{\mu\nu}^2/32\pi^2\rangle_{\rm NP}\right) \tag{41}$$

where $\langle F_{\mu\nu}^2\rangle_{\rm NP}$ is the gluon field vacuum expectation value which is due to non-perturbative fluctuations, i.e. the gluon condensate. The aim of any QCD-vacuum builder is to minimize the vacuum energy or, equivalently, to maximize the gluon condensate. It is important that it is a renormalization-invariant quantity, meaning that its dependence on the ultraviolet cutoff μ and the bare charge $\alpha_s(\mu)$ given at this cutoff is such that it is actually cutoff-independent. Such a combination is called Λ, see eq. (1). The gluon condensate has to be proportional to Λ^4 by dimensions.

The fact that the vacuum energy or, equivalently, the gluon condensate is a renormalization-invariant quantity leads to an infinite number of low-energy theorems [34]. Translated into the instanton-vacuum language, the renormalizability of the QCD implies that the probability that there are N I's and $\bar{I}$'s in the vacuum is [21, 41]

$$P(N) \sim \exp\left[-\frac{b}{4}\left(\ln\frac{N}{\langle N\rangle} - 1\right)\right], \tag{42}$$

where $\langle N\rangle \simeq V\langle F^a_{\mu\nu}F^a_{\mu\nu}\rangle/(32\pi^2)$ is the *average* number of I's and $\bar{I}$'s .

One-instanton weight

The notion "instanton vacuum" implies that one assumes that the QCD partition function (37) is mainly saturated by an ensemble of interacting I's and $\bar{I}$'s , together with quantum fluctuations about them. Instantons are necessarily present in the QCD vacuum if only because they lower the vacuum energy with respect to the purely perturbative (divergent) one. The question is whether they give the dominant contribution to the gluon condensate, and to other basic quantities. To answer this question one has to compute the partition function (37) assuming that it is mainly saturated by instantons, and to compare the obtained gluon condensate with the phenomenological one.

The starting point of this calculation [21, 41] is the contribution of one isolated instanton to the partition function (37), or the one-instanton weight. We have already estimated the tunneling amplitude in eq. (21) but it is not sufficient: the prefactor is very important. To the 1-loop accuracy, it has been first computed by 't Hooft [17] for the $SU(2)$ colour group, and generalized to arbitrary $SU(N)$ by Bernard [42].

The general field can be decomposed as a sum of a classical field of an instanton $A^I_\mu(x,\xi)$ where ξ is a set of $4N_c$ collective coordinates characterizing a given instanton (see eq. (31)), and of a presumably small quantum field $a_\mu(x)$:

$$A_\mu(x) = A^I_\mu(x,\xi) + a_\mu(x). \tag{43}$$

There is a subtlety in this decomposition due to the gauge freedom: an interested reader is addressed to ref. [21] where this subtlety is treated in detail. The action is

$$\text{Action} = \frac{1}{4g^2}\int d^4x F_{\mu\nu}^2 = \frac{8\pi^2}{g^2} + \frac{1}{g^2}\int d^4x D_\mu F_{\mu\nu} a_\nu + \frac{1}{2g^2}\int d^4x a_\mu W_{\mu\nu} a_\nu + O(a^3). \tag{44}$$

Here the term linear in a_μ drops out because the instanton field satisfies the equation of motion. The quadratic form $W_{\mu\nu}$ has $4N_c$ zero modes related to the fact that the action does not depend on $4N_c$ collective coordinates. This brings in a divergence in the functional integral over the quantum field a_μ which, however, can and should be qualified as integrals over the collective coordinates: centre, size and orientations. Formally the functional integral over a_μ gives

$$\frac{1}{\sqrt{\det\ W_{\mu\nu}(A^I)}}, \tag{45}$$

which must be i) normalized (to the determinant of the free quadratic form, i.e. with no background field), ii) regularized (for example by using the Pauli–Villars method), and iii) accounted for the zero modes. Actually one has to compute a "quadrupole" combination,

$$\left[\frac{\det' W \;\det(W_0+\mu^2)}{\det W_0 \;\det(W+\mu^2)}\right]^{-\frac{1}{2}}, \tag{46}$$

where W_0 is the quadratic form with no background field and μ^2 is the Pauli–Villars mass playing the role of the ultraviolet cutoff; the prime reminds that the zero modes should be removed and treated separately. The resulting one-instanton contribution to the partition function (normalized to the free one) is [17, 42]:

$$\frac{\mathcal{Z}_{1-\mathrm{inst}}}{\mathcal{Z}_{\mathrm{P.T.}}} = \int d^4 z_\mu \int d\rho \int d^{4N_c-5} U \, d_0(\rho), \tag{47}$$

$$d_0(\rho) = \frac{C(N_c)}{\rho^5} \left[\frac{2\pi}{\alpha_s(\mu)}\right]^{2N_c} (\mu\rho)^{\frac{11}{3}N_c} \exp\left(-\frac{2\pi}{\alpha_s(\mu)}\right). \tag{48}$$

The fact that there are all in all $4N_c$ integrations over the collective coordinates z_μ, ρ, U reflects $4N_c$ zero modes in the instanton background. The numerical coefficient $C(N_c)$ depends implicitly on the regularization scheme used. In the Pauli–Villars scheme exploited above [42]

$$C(N_c) = \frac{4.60 \exp(-1.68 N_c)}{\pi^2 (N_c-1)!(N_c-2)!}. \tag{49}$$

If the scheme is changed, one has to change the coefficient $C(N_c) \to C'(N_c) = C(N_c) \cdot (\Lambda/\Lambda')^b$. One has [43]: $\Lambda_{\mathrm{P.V.}} = e^{\frac{1}{22}} \Lambda_{\overline{\mathrm{MS}}} = 40.66\, e^{-\frac{3\pi^2}{11 N_c^2}} \Lambda_{\mathrm{lat}} = ...$

Eq. (48) cannot yet be expressed through the 2-loop renormalization-invariant combination Λ (1) as it is written to the 1-loop accuracy only. In the 2-loop approximation the instanton weight is given by [44, 21]

$$
\begin{aligned}
d_0(\rho) &= \frac{C(N_c)}{\rho^5} \beta(\rho)^{2N_c} \exp\left[-\beta^{\mathrm{II}}(\rho) + \left(2N_c - \frac{b_2}{2b_1} \right) \frac{b_2}{2b_1} \frac{\ln \beta(\rho)}{\beta(\rho)} + O\left(\frac{1}{\beta(\rho)} \right) \right] \\
&\sim \frac{1}{\rho^5} (\Lambda\rho)^{\frac{11}{3} N_c},
\end{aligned}
\tag{50}
$$

where $\beta(\rho) \equiv 2\pi/\alpha_s(\rho)$ and $\beta^{\mathrm{II}}(\rho)$ are the inverse charges to the 1-loop and 2-loop accuracy, respectively (not to be confused with the Gell-Mann–Low function!):

$$\beta^{\mathrm{II}}(\rho) = \beta(\rho) + \frac{b_2}{2b_1} \ln \frac{2\beta(\rho)}{b_1}, \tag{51}$$

$$\beta(\rho) = b_1 \ln \frac{1}{\Lambda\rho}, \qquad b_1 = \frac{11}{3} N_c, \qquad b_2 = \frac{34}{3} N_c^2. \tag{52}$$

These equations express the one-instanton weight $d_0(\rho)$ through the cutoff-independent combination Λ (1), and the instanton size ρ. This is how the 'transmutation of dimensions' occurs in the instanton calculus and how Λ enters into the game. Henceforth all dimensional quantities will be expressed through Λ, which is very much welcome.

Notice that the integral over the instanton sizes in eq. (47) diverges as a high power of ρ at large ρ: this is of course the consequence of asymptotic freedom. It means that individual instantons tend to swell. This circumstance plagued the instanton calculus for many years. If one attempts to cut the ρ integrals "by hand", one violates the renormalization properties of the YM theory, as mentioned in the previous section. Actually the size integrals appear to be cut from above due to instanton interactions.

Instanton ensemble

To get a volume effect from instantons one needs to consider an $I\bar{I}$ ensemble, with their total number N proportional to the 4-dimensional volume V. Immediately a mathematical difficulty arises: any superposition of I's and $\bar{I}$'s is not, strictly speaking, a solution of the equation of motion, therefore, one cannot directly use the semiclassical approach of the previous section. One way to overcome this difficulty is to use a variational principle [21]. Its idea is to use a modified YM action for which a chosen $I\bar{I}$ ansatz *is* a saddle point. Exploiting the convexity of the exponent one can prove that the true vacuum energy is *less* than that obtained from the modified action. One can therefore use variational parameters (or even functions) to get a best upper bound for the vacuum energy. It is not the Rayleigh-Ritz but rather the Feynman variational principle since the method has been suggested by Feynman in his famous study of the polaron problem. The gauge theory is more difficult, though: one has not to loose either gauge invariance or the renormalization properties of the YM theory. These difficulties were overcome in ref. [21], see also [41]. It should be kept in mind that we are dealing with "strong interactions", meaning that all dimensionless quantities are generally speaking of the order of unity – there are no small parameters in the theory. Therefore, one has to use certain approximate methods, and the variational principle is among the best. Todays direct lattice investigation of the $I\bar{I}$ ensemble seem to indicate that we have obtained rather accurate numbers in this difficult problem.

In the variational approach, the normalized (to perturbative) and regularized YM partition function takes the form of a partition function for a grand canonical ensemble of interacting pseudoparticles of two kind, I's and $\bar{I}$'s :

$$\frac{\mathcal{Z}}{\mathcal{Z}_{\rm P.T.}} \geq \sum_{N_+,N_-} \frac{1}{N_+!}\frac{1}{N_-!} \prod_n^{N_++N_-} \int d^4 z_n d\rho_n dU_I \; d_0(\rho_n) \; \exp(-U_{\rm int}), \quad (53)$$

where $d_0(\rho)$ is the 1-instanton weight (50). The integrals are over the collective coordinates of (anti)instantons: their coordinates z, sizes ρ and orientations given by $SU(N_c)$ unitary matrices U; dU means the Haar measure normalized to unity. The instanton interaction potential $U_{\rm int}$ (to be discussed below) depends on the separation between pseudoparticles, $z_m - z_n$, their sizes $\rho_{m,n}$ and their relative orientations $U_m U_n^\dagger$. In the variational approach the interaction between instantons arise from *i*) the defect of the classical action, *ii*) the non-factorization of quantum determinants and *iii*) the non-factorization of Jacobians when one passes to integration over the collective coordinates. All three factors are ansatz-dependent, but there is a tendency towards a cancellation of the ansatz-dependent pieces. Qualitatively, in any ansatz the interactions between I's and $\bar{I}$'s resemble those of molecules: at large separations there is an attraction, at smaller separations there is a repulsion. It is very important that the interactions depend on the relative orientations of instantons: if one averages over orientations (which is the natural thing to do if the $I\bar{I}$ medium is in a disordered phase; if not, one would expect a spontaneous breaking of both Lorentz and colour symmetries [21]), the interactions seem to be repulsive at any separations.

In general, the mere notion of the instanton interactions is notorious for being ill-defined since instanton + antiinstanton is not a solution of the equation of motion. Such a configuration belongs to a sector with topological charge zero, thus it seems to be impossible to distinguish it from what is encountered in perturbation theory. The variational approach uses brute force in dealing with the problem, and the results appear to be somewhat dependent on the ansatz used. Thanks to the inequality for the vacuum energy mentioned above, we still get quite a useful information. However, recently a mathematically unequivocal definition of the instanton interaction has been suggested, based on the one hand on analyticity and unitarity [45] and on the other hand on certain singular solutions of the YM equations of motion [46]. Both definitions cut off automatically contributions of the perturbation theory. The first three leading terms for the interaction potential at large separations has been computed by the two very different methods [45, 46] with coinciding results. At smaller separations one observes a strong repulsion [46].

At this point I should mention certain experience one gains from a simpler 2-dimensional so-called CP^N model, also possessing instantons as classical Euclidean solutions. Contrary to the $4d$ YM theory, the instanton measure in that model is known exactly [47, 48]. In the dilute limit the instanton measure reduces to the product of integrals over instanton sizes, positions and orientations, as in eq. (53). The exact measure, however, is written in terms of the so-called 'instanton quarks' which does not suppose that instantons are dilute. The statistical mechanics of I's and $\bar{I}$'s in this model has been studied in ref. [49] both by analytical methods and by numerical simulations. Although the 'instanton quark' parameterization allows for complete 'melting' of instantons and is quite opposite in spirit to the dilute-gas ansatz, it has been observed that, owing to a combination of purely geometric and dynamic reasons, the vast majority of 'instanton quarks' form neutral clusters which can be identified with well-separated instantons. Of course, there is always a fraction of overlapping instantons in the vacuum, however, it is small even in the $2d$ case; in the $4d$ YM case both reasons mentioned above are expected to be even stronger.

Summing up the discussion, I would say that today there exists no evidence that a variational calculation with the simplest sum ansatz used in ref. [21] is qualitatively or even quantitatively incorrect, therefore I will cite the numerics from those calculations in what follows. The main finding [21, 41] is that the $I\bar{I}$ ensemble (53) stabilizes at a certain density related to the Λ parameter (there is no other dimensional quantity in the theory!)

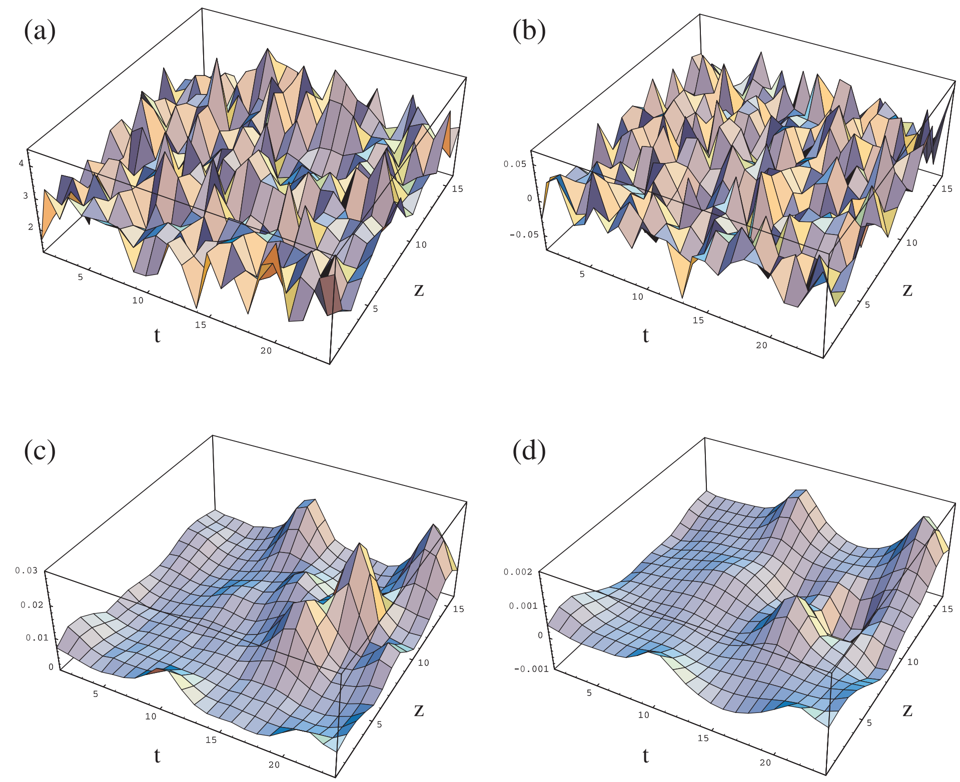


Figure 20: "Cooling" the normal zero-point oscillations reveals large fluctuations of the gluon field, which were identified with instantons and anti-instantons with random positions and sizes [36]. The left column shows the action density and the right column shows the topological charge density for the same snapshot.

$$\frac{N}{V} \simeq \langle F_{\mu\nu}^2/32\pi^2\rangle \simeq \frac{1}{V}\langle Q_T^2\rangle \geq (0.75\Lambda_{\overline{\rm MS}})^4. \tag{54}$$

The average instanton size and the average separation between instantons are, respectively,

$$\bar{\rho} \simeq 0.48/\Lambda_{\overline{\rm MS}} \simeq 0.35\,{\rm fm}, \tag{55}$$

$$\bar{R} = \left(\frac{N}{V}\right)^{-\frac{1}{4}} \simeq 1.35/\Lambda_{\overline{\rm MS}} \simeq 0.95\,{\rm fm}, \tag{56}$$

if one uses $\Lambda_{\overline{\rm MS}} = 280\,{\rm MeV}$ as it follows from the DIS data. Earlier, very similar characteristics, $\bar{\rho} = \frac{1}{3}\,{\rm fm}$, $\bar{R} = 1\,{\rm fm}$, have been suggested by Shuryak [19] from studying the phenomenological applications of instantons.

Instanton interactions lead to the modification of the (divergent) size distribution function $d_0(\rho)$ (50) by a distribution decreasing at large ρ. The use of the variational principle yields a Gaussian cutoff for large sizes [21, 41]:

$$d_0(\rho) \to d(\rho) = d_0(\rho)\,\exp\left(-\,\mathrm{const.}\sqrt{\frac{N}{V}}\,\rho^2\right). \tag{57}$$

In fact, it is a rather narrow distribution peaked around $\bar{\rho}$ (55); therefore for practical estimates in what follows I shall just replace all instantons by the average-size one.

It should be said that, strictly speaking, nothing can prevent some instantons to be anomalously large and overlapping with other. For overlapping instantons the notion of size distribution becomes senseless. The question is quantitative: how often and how strong do instantons overlap. Given the estimate (55,56), it seems that the majority of instantons in the vacuum ensemble are well-isolated.

In the recent years instantons have been intensively studied by direct numerical simulations of gluon fields on the lattice, using various configuration-smoothing methods [36, 50, 51]. A typical snapshot of gluon fluctuations in the vacuum is shown in Fig. 3 borrowed from ref. [36]. Naturally, it is heavily dominated by normal perturbative UV-divergent zero-point oscillations of the field. However, after "cooling" down these oscillations one reveals a smooth background field which was shown in ref. [36] to be nothing but an ensemble of instantons and anti-instantons with random positions and sizes [14]. The lower part of Fig.20 is what is left of the upper part after "cooling" that particular configuration. The average sizes and separations of instantons found vary somewhat depending on the concrete smearing method used. Ref. [36] gives the following values

[14] Quite recently a more involved fluctuation-smearing procedure carried on the lattice has indicated that instantons might have an additional structure, see the next section.

$$\bar{\rho} \simeq 0.36\,\mathrm{fm}, \qquad \bar{R} = (N/V)^{-\frac{1}{4}} \simeq 0.89\,\mathrm{fm}, \tag{58}$$

which are not far from the estimate from the variational principle. The ratio,

$$\frac{\bar{\rho}}{\bar{R}} \simeq \frac{1}{3}, \tag{59}$$

seems to be more stable: it follows from phenomenological [19], variational [21, 41] and lattice [36, 50, 51] studies. It means that the *packing fraction*, i.e. the fraction of the 4-dimensional volume occupied by instantons appears to be rather small, $\pi^2\bar{\rho}^4/\bar{R}^4 \simeq 1/8$. This small packing fraction of the instantons gives an *a posteriori* justification for the use of the semi-classical methods. As I shall show in the next sections, it also enables one to identify adequate degrees of freedom to describe the low-energy QCD.

NONPERTURBATIVE QCD

The QCD Lagrangian with N_f massless flavours is known to posses a large global symmetry, namely a symmetry under $U(N_f) \times U(N_f)$ independent rotations of left- and right-handed quark fields. This symmetry is called *chiral* [15]. Instead of rotating separately the 2-component Weyl spinors corresponding to left- and right-handed components of quark fields, one can make independent vector and axial $U(N_f)$ rotations of the full 4-component Dirac spinors – the QCD lagrangian is invariant under these transformations too.

Meanwhile, axial transformations mix states with different P-parities. Therefore, were that symmetry exact, one would observe parity degeneracy of all states with otherwise the same quantum numbers. In reality the splittings between states with the same quantum numbers but opposite parities are huge. For example, the splitting between the vector ρ and the axial a_1 meson is $(1260 - 770) \simeq$

[15] The word was coined by Lord Kelvin in 1894 to describe molecules not superimposable on its mirror image.

$500\,\mathrm{MeV}$; the splitting between the nucleon and its parity partner is even larger: $(1535-940) \simeq 600\,\mathrm{MeV}$.

The splittings are too large to be explained by the small bare or current quark masses which break the chiral symmetry from the beginning. Indeed, the current masses of light quarks are: $m_u \simeq 4\,\mathrm{MeV}, \quad m_d \simeq 7\,\mathrm{MeV}, \quad m_s \simeq 150\,\mathrm{MeV}$. The conclusion one can draw from these numbers is that the chiral symmetry of the QCD Lagrangian is broken down *spontaneously*, and very strongly. Consequently, one should have light (pseudo) Goldstone pseudoscalar hadrons – their role is played by pions which indeed are by far the lightest hadrons.

The order parameter associated with chiral symmetry breaking is the so-called *chiral* or *quark condensate*:

$$\langle \bar{\psi}\psi \rangle \simeq -(250\ MeV)^3. \tag{60}$$

It should be noted that this quantity is well defined only for massless quarks, otherwise it is somewhat ambiguous. By definition, this is the quark Green function taken at one point; in momentum space it is a closed quark loop:

$$\langle\bar{\psi}\psi\rangle = -N_c \int \frac{d^4k}{(2\pi)^4 i}\,\mathrm{Tr}\,\frac{Z(k)}{M(k) - \not{k}}. \tag{61}$$

If the quark propagator is massless and has only the 'slash' term, the trace over the spinor indices in the loop gives an identical zero. Therefore, chiral symmetry breaking implies that a massless (or nearly massless) quark develops a non-zero dynamical mass $M(k)$, *i.e.* a 'non-slash' term in the propagator. There are no reasons for this quantity to be a constant independent of the momentum; moreover, we understand that it should anyhow vanish at large momentum. Sometimes it is called the constituent quark mass, however a momentum-dependent *dynamical quark mass* $M(k)$ is a more adequate term which I shall use below.

The spontaneous generation of the dynamical quark mass (equivalent to the spontaneous chiral symmetry breaking, SCSB) is the most important feature of QCD being key to the whole hadron phenomenology. The theory's task is to get $M(k)$ in the form

$$M(k) = \Lambda f(k/\Lambda) \tag{62}$$

where Λ is the renormalization-invariant combination (1) and f is some function. Instantons enable one to get $M(k)$ in the needed form and to find the function. But first let us derive some general relations.

We start by writing down the QCD partition function. Functional integrals are well defined in Euclidean space which is obtained by the following formal substitutions of Minkowski space quantities:

$$\begin{aligned} ix_{M0} &= x_{E4}, \quad x_{Mi} = x_{Ei}, \quad A_{M0} = iA_{E4}, \quad A_{Mi} = A_{Ei}, \\ i\bar{\psi}_M &= \psi_E^\dagger, \quad \gamma_{M0} = \gamma_{E4}, \quad \gamma_{Mi} = i\gamma_{Ei}, \quad \gamma_{M5} = \gamma_{E5}. \end{aligned} \tag{63}$$

Neglecting for brevity the gauge fixing and Faddeev–Popov ghost terms, the QCD partition function with quarks can be written as

$$\begin{aligned} \mathcal{Z} &= \int DA_\mu D\psi D\psi^\dagger \exp\left[-\frac{1}{4g^2}\int F_{\mu\nu}^2 + \sum_f^{N_f} \int \psi_f^\dagger (i\nabla\!\!\!/ + im_f)\psi_f\right] \\ &= \int DA_\mu \exp\left[-\frac{1}{4g^2}\int F_{\mu\nu}^2\right] \prod_f^{N_f} \det(i\nabla\!\!\!/ + im_f). \end{aligned} \tag{64}$$

The chiral condensate of a given flavour f is, by definition,

$$\langle\bar{\psi}_f\psi_f\rangle_M = -i\langle\psi_f^\dagger\psi_f\rangle_E = -\frac{1}{V}\frac{\partial}{\partial m_f}\left(\ln\mathcal{Z}\right)_{m_f\to 0}. \tag{65}$$

The Dirac operator has the form

$$i\nabla\!\!\!/ = \gamma_\mu(i\partial_\mu + A_\mu^{\mathrm{I\bar{I}}} + a_\mu) \tag{66}$$

where $A_\mu^{\mathrm{I\bar{I}}}$ denotes the classical field of the $I\bar{I}$ ensemble and a_μ is a presumably small field of quantum fluctuations about that ensemble, which I shall neglect as it has little impact on chiral symmetry breaking. Integrating over DA_μ in eq. (64) means averaging over the $I\bar{I}$ ensemble with the partition function (53), therefore one can write

$$\mathcal{Z} = \overline{\det(i\nabla\!\!\!/ + im)} \tag{67}$$

where I temporarily restrict the discussion to the case of only one flavour for simplicity. Because of the im term the Dirac operator in (67) is formally not Hermitian; however the determinant is real due to the following observation. Suppose we have found the eigenvalues and eigenfunctions of the Dirac operator,

$$i\nabla\!\!\!/\Phi_n = \lambda_n \Phi_n, \tag{68}$$

then for any $\lambda_n \neq 0$ there is an eigen-function $\Phi_{n'} = \gamma_5 \Phi_n$ whose eigenvalue is $\lambda_{n'} = -\lambda_n$. This is because γ_5 anticommutes with $i\nabla\!\!\!/$. Owing to this the fermion determinant can be written as

$$\det(i\nabla\!\!\!/ + im) = \prod_n(\lambda_n + im) = \sqrt{\prod(\lambda_n^2 + m^2)} = \exp\left[\frac{1}{2}\sum_n \ln(\lambda_n^2 + m^2)\right]$$

$$= \exp\left[\frac{1}{2}\int_{-\infty}^{\infty} d\lambda\,\nu(\lambda)\ln(\lambda^2 + m^2)\right], \quad \nu(\lambda) \equiv \sum_n \delta(\lambda - \lambda_n), \tag{69}$$

Instanton vacuum field is assumed as a superposition of N_+ instantons and N_- antiinstantons:

$$A_\mu(x) = \sum_I^{N_+} A_\mu^I(\xi_I, x) + \sum_A^{N_-} A_\mu^A(\xi_A, x). \tag{70}$$

Here $\xi = (\rho, z, U)$ are (anti)instanton collective coordinates– size, position and color orientation.

The first step of the derivation [25, 41] was the splitting of the total quark determinant to the low and high frequencies parts as $\mathrm{Det} = \mathrm{Det}_{\mathrm{high}} \cdot \mathrm{Det}_{\mathrm{low}}$, where $\mathrm{Det}_{\mathrm{high}}$ gets a contribution from fermion modes with Dirac eigenvalues from the interval M_1 to the Pauli–Villars mass M, and $\mathrm{Det}_{\mathrm{low}}$ is accounted eigenvalues less than M_1. The product of these determinants is independent on the scale M_1. But, we may calculate both of them only approximately. In [25] was demonstrated week dependence of the product on M_1in the wide range of M_1, which serves as a check of the approximations.

The high-momentum part $\mathrm{Det}_{\mathrm{high}}$ can be written as a product of the determinants in the field of individual instantons, while the low-momentum one $\mathrm{Det}_{\mathrm{low}}$ has to be treated approximately, would-be zero modes being taken into account only.

LIGHT QUARKS IN THE INSTANTON VACUUM

Low-frequency part of the quark determinant beyond the chiral limit.

The starting point of the consideration is the zero-mode approximation formulated in [52, 25, 41]

$$S_i = \frac{1}{\hat{p} + \hat{A}_i + im} = \frac{1}{\hat{p}} + \frac{|\Phi_{i,0}\rangle\langle\Phi_{i,0}|}{im}. \tag{71}$$

Here the zero-modes $\Phi_{i,0}$ are also functions of the instanton collective coordinates ξ_i. This approximation is good for small values of m (chiral limit) and indicates that the main contribution to the quark propagator is due to the zero-modes.

We would like to go beyond chiral limit. First of all we have to extend Ed. (71)

for a non-small m case. Then, our main assumption is the interpolation formula:

$$S_i = S_0 + S_0 \hat{p} \frac{|\Phi_{0i} >< \Phi_{0i}|}{c_i} \hat{p} S_0 \tag{72}$$

where

$$c_i = - < \Phi_{0i} | \hat{p} S_0 \hat{p} | \Phi_{0i} >= im < \Phi_{0i} | S_0 \hat{p} | \Phi_{0i} >= im < \Phi_{0i} | \hat{p} S_0 | \Phi_{0i} > \tag{73}$$

The advantage of this interpolation is shown by the projection of S_i to the zero-modes:

$$S_i | \Phi_{0i} >= \frac{1}{im} | \Phi_{0i} >, \;\; < \Phi_{0i} | S_i =< \Phi_{0i} | \frac{1}{im} \tag{74}$$

as it must be, while the similar projection of S_i given by Eq. (71) has a wrong component, negligible only in the $m \to 0$ limit.

Now we switch on external flavour fields (v, a, s, p), where v_μ and a_μ are vector and axial-vector fields, s and p are scalar and pseudoscalar fields. They also may have flavour content. We assume that v_μ and a_μ has a trivial topological properties and their topological charges are equal to zero. Total quark propagator $\tilde{S}$ and the single instanton quark propagator $\tilde{S}_i$ become:

$$\tilde{S} = \frac{1}{\hat{p} + \hat{A} + \hat{V} + im}, \quad \tilde{S}_i = \frac{1}{\hat{p} + \hat{A}_i + \hat{V} + im}, \tag{75}$$

where $\hat{V} = \hat{v} + \hat{a}\gamma_5 + s + p\gamma_5$ and $\hat{v} = \gamma_\mu v_\mu$, $\hat{a} = \gamma_\mu a_\mu$. Defining also the quark propagator with only external flavour fields $\tilde{S}_0$ and the free one S_0 as follows:

$$\tilde{S}_0 = \frac{1}{\hat{p} + \hat{V} + im}, \quad S_0 = \frac{1}{\hat{p} + im}, \tag{76}$$

we can expand the quark propagator $\tilde{S}$ with respect to a single instanton:

$$\tilde{S} = \tilde{S}_0 + \sum_i (\tilde{S}_i - \tilde{S}_0) + \sum_{i \neq j} (\tilde{S}_i - \tilde{S}_0) \tilde{S}_0^{-1} (\tilde{S}_j - \tilde{S}_0) + \cdots . \tag{77}$$

In order to specify the gauge dependence, we rewrite $\tilde{S}_i$ and $\tilde{S}_0$ in the following form:

$$\begin{aligned}
\tilde{S}_i &= L_i S_i' L_i^{-1}, \;\; S_i' = \frac{1}{\hat{p} + \hat{A}_i + \hat{V}_i' + +im}, \\
\tilde{S}_0 &= L_i S_{0i}' L_i^{-1}, \;\; S_{0i}' = \frac{1}{\hat{p} + \hat{V}_i' + im},
\end{aligned} \tag{78}$$

where $\hat{V}_i' = L_i^{-1}(\hat{p} + \hat{V})L_i$ and the gauge connection L_i can be written as the path-ordered exponent:

$$L_i(x, z_i) = \mathrm{P}\exp\left(i\int_{z_i}^{x} dy_\mu (v_\mu(y) + a_\mu(y)\gamma_5)\right), \;\; L_i^{-1}(x, z_i) = \gamma_0 L_i^\dagger(x, z_i)\gamma_0 \tag{79}$$

where z_i denotes an instanton position.

We do not include fields s and p into L-factor since they transform homogenously under local gauge transformations.

Expanding S_i' over $\hat{V}_i'$ and re-summing it we get

$$S_i' = S_i(1 + \sum_n (-1)^n (\hat{V}_i' S_i)^n) = S_{0i}' + S_{0i}'\hat{p}\frac{|\Phi_{0i} >< \Phi_{0i}|}{c_i - b_i}\hat{p}S_{0i}' \tag{80}$$

where

$$\begin{aligned} b_i &= < \Phi_{0i}|\hat{p}(S_{0i}' - S_0)\hat{p}|\Phi_{0i} >, \qquad (81) \\ c_i - b_i &= - < \Phi_{0i}|\hat{p}S_{0i}'\hat{p}|\Phi_{0i} > \\ &= < \Phi_{0i}|(im + \hat{V}_i')|\Phi_{0i} > - < \Phi_{0i}|(im + \hat{V}_i')S_{0i}'(im + \hat{V}_i')|\Phi_{0i} > \end{aligned}$$

Rearrangement of the Eq. (77) for the total propagator leads to:

$$\begin{aligned}\tilde{S} &= \tilde{S}_0 + \tilde{S}_0 \sum_{i,j} L_i \hat{p} |\Phi_{i0} > (\frac{1}{-D} + \frac{1}{-D} C \frac{1}{-D} + ...)_{ij} < \Phi_{0j} | \hat{p} L_j^{-1} \tilde{S}_0 \\ &= \tilde{S}_0 + \tilde{S}_0 \sum_{i,j} L_i \hat{p} |\Phi_{i0} > (\frac{1}{-V-T}) ij < \Phi_{0j} | \hat{p} L_j^{-1} \tilde{S}_0 \qquad (82)\end{aligned}$$

where

$$V_{ij} =< \Phi_{0i} | \hat{p} (L_i^{-1} \tilde{S}_0 \bar{L}_j^{-1}) \hat{p} | \Phi_{0j} > - < \Phi_{0i} | \hat{p} S_0 (L_i^{-1} L_j \hat{p} | \Phi_{0j} >, \qquad (83)$$

$$T_{ij} = (1 - \delta_{ij}) < \Phi_{0i} | \hat{p} S_0 (L_i^{-1} L_j \hat{p} | \Phi_{0j} >,$$

$$D_{ij} = \delta_{ij} V_{ij} \equiv (b_i - c_i) \delta_{ij}, \ \ C_{ij} = (1 - \delta_{ij}) V_{ij}.$$

It is natural to introduce

$$|\phi_0 >= \frac{1}{\hat{p}} L\hat{p}|\Phi_0 > \tag{84}$$

which has the same chiral properties as zero-mode function $|\Phi_0 >$. Then

$$\tilde{S} - \tilde{S}_0 = -\tilde{S}_0 \sum_{i,j} \hat{p}|\phi_{0i} >< \phi_{0i}|(\frac{1}{V+T})|\phi_{0j} >< \phi_{0j}|\hat{p}\tilde{S}_0 \tag{85}$$

with

$$V + T = \hat{p}\tilde{S}_0\hat{p}. \tag{86}$$

The final explicit form for (85) is

$$\mathrm{Tr}\,(\tilde{S}-\tilde{S}_0) \tag{87}$$

$$= -\sum_{i,j} < \phi_{0,j,f,g_1}|\hat{p}\,(\tilde{S}_0^2)_{g_1 g_2}\,\hat{p}|\phi_{0,i,g_2,g_3} >< \phi_{0,i,g_3,g_4}|(\frac{1}{\hat{p}\tilde{S}_0\hat{p}})_{g_4 g_5}|\phi_{0,j,g_5,f} >$$

Introducing now the operator

$$\tilde{B}(m)_{ij}^{fg} =< \phi_{0,i,f,f_1}|(\hat{p}\tilde{S}_0\hat{p})_{f_1 g_1}|\phi_{0,j,g_1,g} > \tag{88}$$

it is easy to show that

$$\mathrm{Tr}\,\int_{M_1}^{m} idm'(\tilde{S}(m')-\tilde{S}_0(m')) \tag{89}$$

$$= \sum_{i,j}\int_{\tilde{B}(M_1)}^{\tilde{B}(m)} d\tilde{B}(m')_{ij}^{fg}(\frac{1}{\tilde{B}(m')})_{ji}^{gf} = \tilde{\mathrm{Tr}}\,\ln\frac{\tilde{B}(m)}{\tilde{B}(M_1)}$$

Here $\tilde{\mathrm{Tr}}$ means the trace on the flavour and only on zero-mode ($|\Phi_{0j}>$) space. The explicit form of the matrix $\tilde{B}(m)$ in this space is:

$$\tilde{B}(m)_{ij}^{fg} =< \Phi_{0i}|\hat{p}L^{-1}_{i,f,f_1}\,\tilde{S}_{0,f_1,g_1}L_{j,g_1,g}\hat{p}|\Phi_{0j}> \tag{90}$$

Then

$$\mathrm{Det}_{\mathrm{low}}[v,a,s,p,m] = \det\tilde{B}(m) \tag{91}$$

We see that $\tilde{B}$ is the extension of Lee-Bardeen's matrix B beyond chiral limit in the presence of the external flavour fields (v,a,s,p). and with account of quark current mass m without making expansion over current mass m and also extended to a few flavours case.

If turn off the external fields and expand over m, keeping only $O(m)$ term, we obtain the quark determinant $\mathrm{Det}_{\mathrm{low}}$ for small m case [52, 25].

Fermionized representation of the determinant

First, by introducing the Grassmanian variables Ω_i $\bar{\Omega}_j$ we represent

$$\det \tilde{B} = \int d\Omega d\bar{\Omega} \exp(\bar{\Omega}\tilde{B}\Omega), \tag{92}$$

where

$$\bar{\Omega}\tilde{B}\Omega = \bar{\Omega}_i < \Phi_{0i}|\hat{p}L^{-1}_{i,f,f_1}\tilde{S}_{0,f_1,g_1}L_{j,g_1,g}\hat{p}|\Phi_{0j} > \Omega_j \tag{93}$$

The next step is to introduce the sources η_i and $\bar{\eta}_j$ defined as:

$$\bar{\eta}_i = -\bar{\Omega}_i < \Phi_{0i}|\hat{p}L^{-1}_{i,f,f_1}, \eta_j = \hat{p}|\Phi_{0j} > \Omega_j \tag{94}$$

Then $(\bar{\Omega}\tilde{B}\Omega)$ can be rewritten as

$$(\bar{\Omega}\tilde{B}\Omega) \;=\; -(\bar{\eta}\tilde{S}_0\eta) \tag{95}$$

Compare this representation with the generation functional of fermionic fields correlators to get

$$\begin{aligned}
\det \tilde{B} \;&=\; \int d\Omega d\bar{\Omega}\exp(\bar{\Omega}\tilde{B}\Omega) \\
&=\; \left(\det(\tilde{S}_0^{-1})\right)^{-1}\int d\Omega d\bar{\Omega} D\psi D\psi^\dagger \exp\int dx(\psi^\dagger(x)\tilde{S}_0^{-1}\psi(x) \\
&+\; \sum_i(\bar{\eta}_i(x)\psi(x) \;+\; \psi^\dagger(x)\eta_i(x)))
\end{aligned} \tag{96}$$

The integration over Grassmanian variables Ω and $\bar{\Omega}$ (with the account of the N_f flavours $\det_N = \prod_f \det B_f$) provides finally the fermionized representation of the determinant (93) in the form:

$$\mathrm{Det}_{\mathrm{low}}[v,a,s,p,m] = \det \tilde{B} \tag{97}$$

$$= \int \prod_f D\psi_f D\psi_f^\dagger \exp\left(\int d^4x \sum_{f,g} \psi_f^\dagger (\hat{p} + \hat{V} + im)_{fg}\psi_g\right)$$

$$\times \prod_f \left\{ \prod_{+}^{N_+} \tilde{V}_{+,f}[\psi^\dagger,\psi] \prod_{-}^{N_-} \tilde{V}_{-,f}[\psi^\dagger,\psi] \right\} \ ,$$

where

$$\begin{aligned} \tilde{V}_{\pm,f}[\psi^\dagger,\psi] &= \sum_{f_1,f_2} \int d^4x \left(\psi_{f_1}^\dagger(x)\, L_{\pm,f_1 f}(x,z)\, \hat{p}\Phi_{\pm,0}(x;\xi_\pm)\right) \\ &\times \int d^4y \left(\Phi_{\pm,0}^\dagger(y;\xi_\pm)(\hat{p}\, L_{\pm,f f_2}^{-1}(y,z)\psi_{f_2}(y)\right), \end{aligned} \tag{98}$$

and

$$L_{\pm}(x,z) = \mathrm{P}\exp\left(i\int_z^x dy_\mu(v_\mu(y) \pm a_\mu(y))\right), \tag{99}$$

$$L_{\pm}^{-1}(x,z) = \mathrm{P}\exp\left(-i\int_z^x dy_\mu(v_\mu(y) \mp a_\mu(y))\right).$$

This fermionization was suggested in [55] and we interpret the fermions $\psi^\dagger, \psi$ as a constituent quarks. Note that external v_μ and a_μ fields gauges not only the kinetic term of the effective action but also its interaction term $\tilde{V}_{\pm,f}[\psi^\dagger, \psi]$ in Eq. (99). The reason is obvious: It is the nonlocal interaction induced by instantons. The external v_μ and a_μ fields are presented here due to the factor L attached to each fermionic line. This factor provides us a gauge invariance of the interaction term $\tilde{V}_{\pm,f}[\psi^\dagger, \psi]$ under the gauge transformation. Instead of we have path dependence of $\tilde{V}_{\pm,f}[\psi^\dagger, \psi]$ via L factors. The reason is obvious – we did not take into account non-zero quark modes in $\mathrm{Det}_{\mathrm{low}}$, which spoil completeness of the states. In each specific case we have to estimate the accuracy of the calculation and to find the way of the minimization of an uncertainty introduced by this path dependence.

The remaining problem is to average the quark determinant over collective coordinates $\xi_\pm$. It is a rather simple procedure, since the low density of the instanton medium $(\pi^2 \left(\frac{\rho}{R}\right)^4 \sim 0.1)$ allows us to average over positions and orientations of the instantons independently.

Now the averaging over collective coordinates $\xi_\pm$ become trivial problem. By using Fourier-transformed zero-modes $\exp(-ikz)\Phi_{\pm,0}(k;\xi_\pm) = \exp(-ikz)\int d^4(x-z)\exp(-ik(x-z))\Phi_{\pm,0}(x-z;\xi_\pm)$ we get:

$$
\begin{aligned}
&\Phi_{\pm,0,i\alpha}(k_1;\xi_\pm)\Phi^\dagger_{\pm,0,j\beta}(k_2;\xi_\pm) \\
&= \frac{(2\pi\rho)^2 F(k_1)F(k_2)}{8k_1^2k_2^2}(\hat{k}_1\gamma_\mu\gamma_\nu\hat{k}_2\frac{1\mp\gamma_5}{2})_{ij}(U_\pm\tau^\mp_\mu\tau^\pm_\nu U^\dagger_\pm)_{\alpha\beta}
\end{aligned}
\tag{100}
$$

where z is a position and $U_\pm$ is a matrix of the color orientations of the (anti)instanton. Form-factor $F(k)$ is a normalized zero-mode. It has explicit form:

$$F(k) = -\frac{d}{dt}[I_0(t)K_0(t) - I_1(t)K_1(t)]_{t=\frac{|k|\rho}{2}} \tag{101}$$

Also, it was used simplified version of the form-factor $F(p)$ (with corrected high momentum dependence for actual numerical calculations):

$$\begin{aligned} F(p) \qquad &= \frac{L^2}{L^2+p^2}, \;\; p < 2GeV \\ &= \frac{1.414}{p^3}, \;\; p > 2GeV \end{aligned} \tag{102}$$

where $L \approx \frac{\sqrt{2}}{\bar{\rho}} = 848MeV$.

The integration over color orientations, which are given by the formulae:

$$\int dU = 1, \int dU U^i_k U^{\dagger j}_l = \frac{1}{N_c}\delta^i_l \delta^j_k \tag{103}$$

$$\int dU U^{i_1}_{k_1} U^{i_2}_{k_2} U^{\dagger j_1}_{l_1} U^{\dagger j_2}_{l_2} = \frac{1}{N_c^2 - 1}[\delta^{i_1}_{l_1}\delta^{i_2}_{l_2}(\delta^{j_1}_{k_1}\delta^{j_2}_{k_2} - \frac{1}{N_c}\delta^{j_2}_{k_1}\delta^{j_1}_{k_2}) + (1 \leftrightarrow 2)]$$

provide specific a la t'Hooft structures of quark-quark interaction term.

Also, it is evident that the integration over z leads to the energy-momentum conservation delta-function.

Exponentiation.

With very good accuracy we may apply saddle-point approximation to the following integral to get

$$\frac{i}{(2\pi)^{0.5}} \int_{\alpha+i\infty}^{\alpha-i\infty} d\lambda \exp\left(N \ln \frac{N}{\lambda} + y\lambda - N\right) = y^N \tag{104}$$

at $N >> 1$.

Bosonization of partition function.

After applying the formula (104) partition function is:

$$\begin{aligned} Z[m] \quad = \quad & \int D\psi D\psi^\dagger \exp[\int d^4x \sum_f \psi_f^\dagger (i\hat{\partial} + im_f)\psi_f \qquad (105) \\ & +\lambda_+ Y_{N_f}^+ + \lambda_- Y_{N_f}^- + N_+(\ln \frac{N_+}{\lambda_+ V M_1^{N_f}} - 1) + N_-(\ln \frac{N_-}{\lambda_- V M_1^{N_f}} - 1)] \end{aligned}$$

Take $N_\pm = N$, $\lambda_\pm = \lambda$, $N_f = 2$, $m_f = m$. Then

$$Y_2^\pm = \frac{1}{N_c^2 - 1} \int d^4x\nu(\rho)d\rho[(1-\frac{1}{2N_c}) \det(iJ^\pm(\rho,x)+\frac{1}{8N_c} \det(iJ^\pm_{\mu\nu}(\rho,x)], \tag{106}$$

where

$$J^\pm_{fg}(\rho,x) = \int \frac{d^4k_f d^4l_g}{(2\pi)^8} \exp i(k_f - l_g)x q_f^+(k_f)\frac{1\pm\gamma_5}{2} q_g(l_g) \tag{107}$$

$$J^\pm_{\mu\nu,fg}(\rho,x) = \int \frac{d^4k_f d^4l_g}{(2\pi)^8} \exp i(k_f - l_g)x q_f^+(k_f)\frac{1\pm\gamma_5}{2} \sigma_{\mu\nu} q_g(l_g), \tag{108}$$

where $q(x) = \int \frac{d^4k}{(2\pi)^4} \exp(ikx)\, q(k)$, $g^2 = \frac{(N_c^2-1)2N_c}{2N_c-1}$,
$q(k) = 2\pi\rho F(k\rho)\psi(k)$.

Introduce meson fields taking into account the chiral transformations:
$\delta q = i\gamma_5 \vec{\tau}\vec{\alpha} q$, $\delta q^+ = q^+ i\gamma_5 \vec{\tau}\vec{\alpha}$,
$\delta\sigma = 2\vec{\alpha}\vec{\phi}$, $\delta\vec{\phi} = -2\vec{\alpha}\sigma$, $\delta\eta = -2\vec{\alpha}\vec{\sigma}$, $\delta\vec{\sigma} = 2\eta\vec{\alpha}$.
Then, as it must be, $\delta q^+(\sigma + i\gamma_5\vec{\tau}\vec{\phi})q = 0$, $\delta q^+(\vec{\tau}\vec{\sigma} + i\gamma_5\eta)q = 0$.

The bosonization now means

$$\exp\int d^4x[\lambda(\det\frac{iJ^+}{g} + \det\frac{iJ^-}{g})] \equiv \exp\int d^4x \frac{\lambda}{8g^2}[-(q^+q)^2 - (q^+ i\gamma_5\vec{\tau}q)^2$$
$$+(q^+ i\gamma_5 q)^2 + (q^+\vec{\tau}q)^2] = \int D\sigma D\vec{\phi} D\eta D\vec{\sigma} \tag{109}$$
$$\times \exp\int d^4x \left[\frac{\lambda^{0.5}}{2g} q^+ i(\sigma + i\gamma_5\vec{\tau}\vec{\phi} + i\vec{\tau}\vec{\sigma} + \gamma_5\eta)q - \frac{1}{2}(\sigma^2 + \vec{\phi}^2 + \vec{\sigma}^2 + \eta^2)\right]$$

The partition function become

$$Z[m] = \int d\lambda D\sigma D\vec{\phi} D\eta D\vec{\sigma} \exp[N \ln \frac{K}{\lambda} - N - \frac{1}{2}\int d^4x(\sigma^2 + \vec{\phi}^2 + \vec{\sigma}^2 + \eta^2)$$

$$+\mathrm{Tr}\ \ln \frac{\hat{p} + im + i\frac{\lambda^{0.5}}{2g}(2\pi\rho)F(\sigma + i\gamma_5\vec{\tau}\vec{\phi} + i\vec{\tau}\vec{\sigma} + \gamma_5\eta)(2\pi\rho)F}{\hat{p} + im}] \tag{110}$$

Here K some unessential constant to make under-log expression dimensional-less. Tr means $\int d^4x \mathrm{tr}\,_c \mathrm{tr}\,_D \mathrm{tr}\,_f$, $p_\mu = i\partial_\mu$, $[p_\mu, x_\nu] = i\delta_{\mu\nu}$.

Dynamical quark mass

One of the main advantages of the instanton vacuum model is the natural description of the SχSB, which is signaled by non-zero vacuum quark condensate $\langle \bar{q}q \rangle$. The quark-quark interaction term (106) leads to the strong attraction in the channels with vacuum (and pion) quantum numbers. As a consequence, there appear the nonzero vacuum expectation σ of scalar-isoscalar component of meson fields Φ and related with it $\langle \bar{q}q \rangle$. For evaluation of the partition function $Z[m]$, it is very convenient to use the formalism of the effective action [62, 63] $\Gamma_{eff}[m, \lambda, \Phi]$, defined as:

$$Z_N[m] = \int d\lambda Z_N[m, \lambda] = \int d\lambda \exp(-\Gamma_{eff}[m, \lambda, \Phi]) \tag{111}$$

where for the sake of simplicity we dropped all the external currents which are not essential in this section, and the field Φ is the solution of the vacuum equation

$$\frac{\partial \Gamma_{eff}[m,\lambda,\Phi]}{\partial \Phi} = 0. \qquad (112)$$

Notice that the solution depends on λ, *i.e.* $\Phi = \Phi(\lambda)$. Here it will be assumed that the only nonzero vacuum field is a condensate $\Phi = \sigma$, which is independent of coordinates, so the effective action $\Gamma_{eff}[m,\lambda,\Phi]$ may be replaced with effective potential $V_{eff}[m,\lambda,\sigma]$.

In the leading order, the effective action just coincides with the action. Shifting $\Phi \to \sigma + \Phi'$ and integrating over the fluctuations, we get for the meson loop correction

$$\Gamma^{mes}_{eff}[m,\lambda,\sigma] = \frac{1}{2}\mathrm{Tr}\,\ln\left(4\delta_{ij} - \frac{1}{\sigma^2}\mathrm{Tr}\,\frac{M(p)}{\hat{p}+i\mu(p)}\Gamma_i\frac{M(p)}{\hat{p}+i\mu(p)}\Gamma_j\right), \qquad (113)$$

where $\mu(p) = m + M(p)$ and we introduced the dynamical quark mass $M(p) =$

$MF^2(p)$; $M = \frac{(2\pi\rho)^2\lambda^{0.5}}{2g}\sigma.$

It is convenient to introduce notations for the leading order meson propagators

$$\Pi_i^{-1}(q) = 4 + \frac{1}{\sigma^2}V_2^i(q), \tag{114}$$

where $V_2^i(q) = Tr\left(Q(p)\Gamma_i Q(p+q)\Gamma_i\right)$, and $Q(p) = S(p)iM(p) \equiv \frac{iM(p)}{\hat{p}+i\mu(p)}$. With these notations vacuum equation (112) turns into

$$4\sigma^2 - \frac{1}{V}Tr\left(Q(p)\right) - \frac{1}{\sigma^2}\int\frac{d^4q}{(2\pi)^4} \quad \sum_i \quad V_3^i(q)\Pi_i(q) = 0,$$

where $V_3^i(q) = Tr\left(Q^2(p)\Gamma_i Q(p+q)\Gamma_i\right)$.

There is an important difference between the instanton vacuum and traditional NJL-type models – the coupling λ is not an external parameter of the model, but is defined from the saddle-point equation. We have to integrate over the coupling

λ in Eq. (111) to obtain partition function Z_N. The saddle-point approximation for the result becomes exact in the large-N limit. The saddle-point equation for λ has a form

$$\frac{N}{V} - \frac{1}{2V} Tr\,(Q(p)) + + \frac{1}{2\sigma^2} \int \frac{d^4q}{(2\pi)^4} \sum_i \left(V_2^i(q) - V_3^i(q)\right) \Pi_i(q) = 0 \tag{115}$$

Notice that V_2-term in (115) requires special attention. Formally it is next to leading order correction, while numerically it is strongly enhanced (about a factor of 30 compared to the other $1/N_c$-corrections), which indicates the failure of the large-N_c expansion in (115).

If we solve the equation (115), expanding it in powers of $1/N_c$ with the set of our parameters, we'll get

$$M_0 = 0.567 - 2.362\, m \tag{116}$$

$$M_1 = \frac{1}{N_c}(-0.687 - 0.808\, m - 4.197\, m \ln m).$$

Here and in the following M and m are given in GeV.

We can see that the meson loop correction M_1 is of comparable size with the LO term M_0, so we can try to *solve the equations (115,115) numerically in chiral limit and then evaluate the chiral corrections to it.* Such procedure gives

$$M(m) = 0.36 - 2.36\, m - \frac{m}{N_c}(0.808 + 4.197 \ln m) \tag{117}$$

The accuracy of the solutions(116,117) is $\mathcal{O}(m^2, \frac{1}{N_c^2})$.

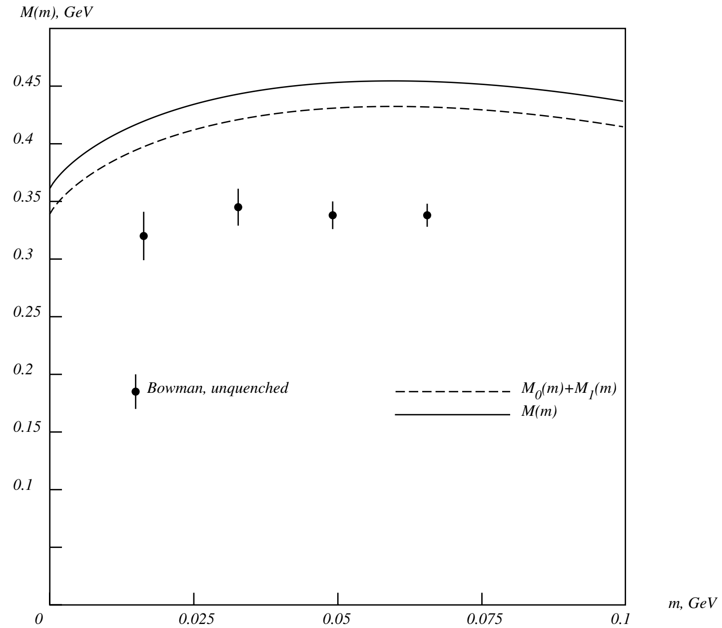


Figure 21: m-dependence of the dynamical quark mass M. The solid curve – the exact numerical solution (117) of vacuum Eqs. (115, 115). The dashed curve – the solution (116), obtained by the iterations ($1/N_c$-expansion) with the same accuracy. Data points are from [65]. Notice that the scale of the lattice data is $1.64\,GeV$, not $\rho^{-1} \approx 0.6\,GeV$.

Fig.21 represents the $M(m)$-dependence obtained from Eqs. (115,115). For the sake of comparison we also plotted the lattice data from [65]. From the first point of view our results are $\approx 30\%$ higher than the lattice data. However, they are given in different gauges and on different scales. Since $M(p)$ is essentially nonperturbative object, it is not very easy to rescale the data and make comparison. Rough estimates in perturbative QCD show that the discrepancy may be attributed to the scales difference. We may conclude that we have a qualitative correspondence between our model result for $M(m)$-dependence and unquenched lattice data [65], as it was expected.

Quark condensate

The presence of the quark condensate $\langle\bar{q}q\rangle$ is one of the most important properties of the QCD vacuum. Its value characterizes the SχSB. In the chosen framework we can extract it directly from the effective action taking derivative over the current quark mass [59]

$$\langle\bar{q}q\rangle = \frac{1}{2}\frac{\partial\Gamma_{eff}}{\partial m} = -\frac{1}{2}\mathrm{Tr}\left(\frac{i}{\hat{p}+i\mu(p)} - \frac{i}{\hat{p}+im}\right)$$
$$+\frac{1}{2}\int\frac{d^4q}{(2\pi)^4}\mathrm{Tr}\left(\frac{MF^2(p)}{(\hat{p}+i\mu(p))^2}\Gamma_i\frac{MF^2(p+q)}{\hat{p}+\hat{q}+i\mu(p+q)}\Gamma_i\right)\tilde{\Pi}_i(q). \quad (118)$$

Evaluation of (118) gives

$$-\langle\bar{q}q\rangle(m) = ((0.005 - 0.034\,m)\ N_c + (0.002 - 0.05\,m - 0.058\,m\ \ln m))\,[GeV^3]$$

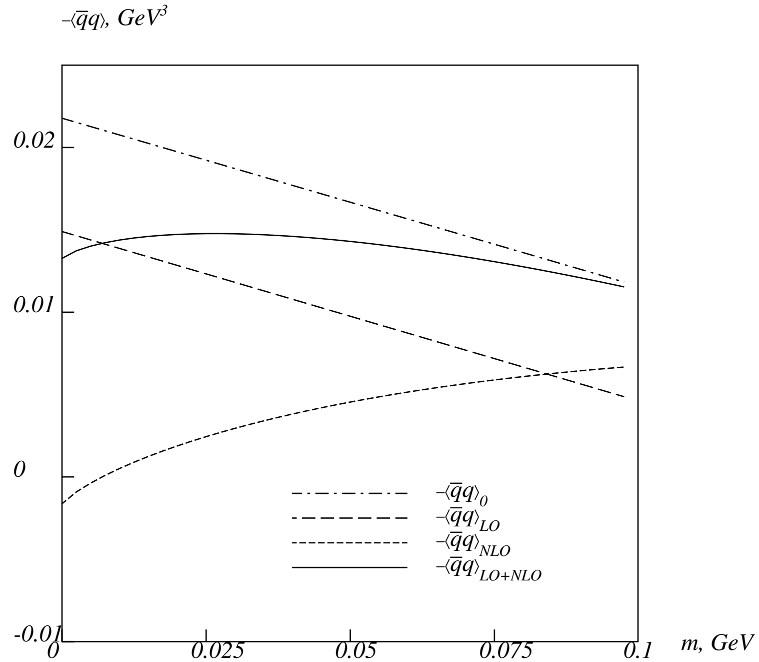


Figure 22: m-dependence of the quark condensate $-\langle\bar{q}q\rangle$. The long-dashed curve is the LO result $-\langle\bar{q}q\rangle_{LO}$, the short-dashed curve is the NLO contribution $-\langle\bar{q}q\rangle_{NLO}$, the solid curve is the total contribution $-\langle\bar{q}q\rangle_{LO+NLO}$. The dot-dashed line represents the leading-order in $1/N_c$-expansion result, evaluated with the mass M_0 from (116) (see text).

The $\langle\bar{q}q\rangle(m)$-dependence is depicted on the Fig. 22. We can see again that due to the chiral logarithm the m-dependence is not linear and meson loops change the $m-$dependence of the quark condensate drastically.

Here and below we define as leading order (LO) the result calculated with the $1/N_c$-corrections dropped everywhere except in the mass $M(m)$, which is taken from (117) without the last $\mathcal{O}(m/N_c, m/N_c \ln m)$-terms. As next-to-leading order (NLO) we define the contribution of the last two terms in (117) plus "direct" contribution of meson loops.

For the sake of comparison, we also plotted the value $\langle\bar{q}q\rangle_0$ which one would get using LO formulae with the mass $M_0(m)$ taken from (116). Recall that the value $\langle\bar{q}q(m=0)\rangle = (255\,MeV)^3$, as well as $F_\pi(m=0) = 88\,MeV$, was used as the input in order to fix the parameters (ρ, R) in (**??**). In LO we have $\langle\bar{q}q\rangle(m=0.1)/\langle\bar{q}q\rangle(m=0)|_{LO} = 0.31$, while with NLO corrections we have $\langle\bar{q}q\rangle(m=0.1)/\langle\bar{q}q\rangle(m=0) = 0.89$, a noticeably different result showing the size of the chiral logarithmic term in the NLO-corrections.

PION PHYSICS FROM THE INSTANTON VACUUM

Quarks in external axial-vector field and pion properties

The formalism of the effective action used in one of the previous sections may be successfully applied in the presence of the external axial-vector isovector field $a_\mu = a^i_\mu \tau_i/2$. The general partition function is reduced in this case to the form

$$Z_N[m, \vec{a}_\mu] = \int d\lambda \exp(-\Gamma_{eff}[m, \lambda, \vec{u}, \vec{a}_\mu]) \tag{119}$$

The external field $\vec{a}_\mu$ can generate nonzero vacuum average $\langle\vec{\phi}\rangle = \vec{u}$ and shift the value of the vacuum filed $\langle\sigma\rangle$ and saddle-point value λ.

In this paper we restrict ourselves to the case of the soft and weak external field $a_\mu(q)$, which can be treated in perturbative fashion, and $q \sim M_\pi \ll \rho^{-1}$. For the purpose of this paper it is sufficient to keep only $\mathcal{O}(\vec{a}_\mu^2, \vec{a}_\mu \partial_\mu \vec{u}, \partial_\mu \vec{u} \partial_\mu \vec{u})$-terms in (119).

On general grounds, one can guess that the vacuum expectation value $\vec{u} \sim \vec{a}_\mu$, whereas shifts of the vacuum field $\langle \sigma \rangle$ and saddle-point value λ are proportional to the second power of $\vec{a}_\mu$. Using the saddle-point equation

$$\frac{\partial \Gamma_{eff}[m, \lambda, \vec{u}, \vec{a}_\mu]}{\partial \lambda} = 0 \qquad (120)$$

and the vacuum equations

$$\frac{\partial \Gamma_{eff}[m, \lambda, \vec{u}, \vec{a}_\mu]}{\partial \sigma} = 0, \ \frac{\partial \Gamma_{eff}[m, \lambda, \vec{u}, \vec{a}_\mu]}{\partial \vec{u}} = 0, \qquad (121)$$

we may easily get that the shifts of $\langle \sigma \rangle, \lambda$ contribute only to $\mathcal{O}(a^4)$-terms and thus may be safely omitted in this paper.

In the leading order the effective action simply coincides with the action. Using vacuum equations (121), one may show that $\langle\sigma\rangle^2 + \langle\vec{\phi}\rangle^2 = const$. This inspires us to introduce a unitary matrix U with the properties

$$U = u_0 + i\vec{\tau}\vec{u},\ U^\dagger U = UU^\dagger = 1, \tag{122}$$

$$\langle\sigma\rangle = \sigma u_0,\ \langle\vec{\phi}\rangle = \sigma\vec{u}. \tag{123}$$

where σ is the value found in Section . In this representation the vacuum meson field is represented as $\Phi_{cl} = \sigma\, U$.

In the next to leading order one has to take into account the fluctuations of the field $\Phi \to \sigma U + \Phi'$ and integrate over Φ'.

Meson loop contribution to Γ_{eff} has a form

$$\Gamma^{mes}_{eff}[m,\lambda,\vec{u},\vec{a}_\mu] = \frac{1}{2}\mathrm{Tr}\;\ln\frac{\delta^2 S[m,\lambda,\sigma,\vec{u},\vec{a}_\mu,\Phi']}{\delta\Phi'_i\delta\Phi'_j}|_{\Phi'=0}$$
$$= \quad \Gamma^{mes}_{eff}[m,\lambda,\vec{u}=0,\vec{a}_\mu=0] + \Delta\Gamma^{mes}_{eff}[m,\lambda,\vec{u},\vec{a}_\mu] \tag{124}$$

After simple but very tedious evaluations it is possible to show that in agreement with chiral symmetry expectations, the structure of the effective action is

$$\Gamma_{eff} = \alpha_0(\vec{a}_\mu+\partial_\mu\vec{u})^2 + m\,\alpha_1\partial_\mu\vec{u}(\vec{a}_\mu+\partial_\mu\vec{u}) + \tag{125}$$
$$+m\,\alpha_2\vec{u}^2 = \frac{1}{2}\left[F^2_{aa}\vec{a}^2_\mu + F^2_{uu}(\partial_\mu\vec{u})^2 + 2F^2_{au}\vec{a}_\mu\partial_\mu\vec{u} + \right.$$
$$\left. +F^2_{uu}M^2_\pi\vec{u}^2\right] + \mathcal{O}(a^3,u^3,m^2),$$

where beyond chiral limit $F^2_{aa} - F^2_{uu} = 2\left(F^2_{au} - F^2_{uu}\right) = -\alpha_1\,m$. Now, one can get that the two-point axial-isovector currents correlator has a form:

$$\int d^4 x e^{-iq\cdot x} \langle j_\mu^{A,i}(x) j_\nu^{A,j}(0)\rangle = \delta_{ij} F_\pi^2 \left(\delta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2 + M_\pi^2}\right) + \mathcal{O}(q^2) \quad (126)$$

We can see that M_π has a meaning of pion mass and F_π – pion decay constant. Numerically it is much easier to calculate F_π^2 as a constant in front of $\delta_{\mu\nu}$-term in (126), taking $\vec{u}(x) = 0$, $a_\mu(x) = const$ (this corresponds to $a_\mu(q \approx 0)$). Also, it is possible to show that the result of such evaluation is independent of the path choice in the transporter L. In a similar way, we can put $a = 0, u = u(q)$ and evaluate in NLO the quantities F_{uu}^2 and pion mass M_π. Both quantities F_π and M_π naturally have the chiral log terms due to the pion loops contributions. The coefficients in these chiral log terms are controlled by the low-energy theorems [66] and are reproduced analytically.

We can see that the specific structure of the Γ_{eff} (125) provides a check of the numerical calculations. Moreover, the chiral log theorems provide another check of the numerical calculations.

Pion decay constant F_π from a^2-term

The basic diagrams which contribute to this quantity in the leading order and in the next-to-leading order are shown schematically in the Fig. 23 and Fig. 24 respectively. Notice that in integration over $\vec{\phi}$ the saddle-point is shifted to $\langle\vec{\phi}\rangle \sim \vec{a}_\mu$, where the "proportionality" sign implies some nonlocal linear operator. All the vertices on these plots should be understood as a sum of the local and nonlocal parts.

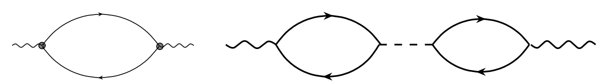


Figure 23: The basic diagrams which contribute to the a^2 term in Γ_{eff}. The wavy line corresponds to the external field $a_\mu(x)$, the dashed line corresponds to the intermediate meson, the bulbs correspond to all the possible (local and nonlocal) couplings of the field a to the constituent quarks.

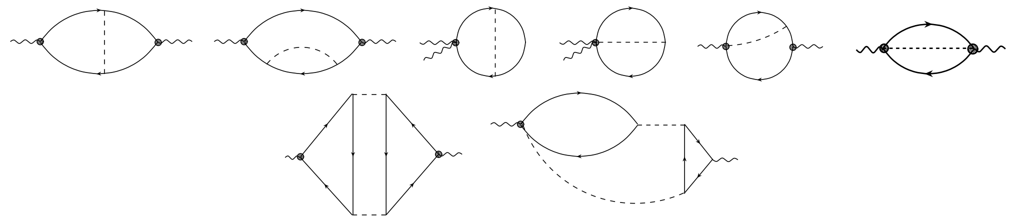


Figure 24: The meson loop corrections to the a^2 term. The notations are the same as in Fig. 23.

Finally we have

$$F_\pi^2 = N_c \left(\left(2.85 - \frac{0.869}{N_c} \right) - \left(3.51 + \frac{0.815}{N_c} \right) m - \right.$$
$$\left. - \frac{44.25}{N_c} m \ln m + \mathcal{O}(m^2) \right) \cdot 10^{-3} \; [GeV^2] = \tag{127}$$
$$(7.67 - 11.35\, m - 44.25\, m \ln m) \cdot 10^{-3} \; [GeV^2]$$

The $F_\pi(m)$-dependence is shown in the Fig.25. For the sake of comparison, we also plotted the value $F_{\pi,0}$ which one would get using LO formulae with the mass $M_0(m)$ taken from (116). Recall that the value $F_\pi(m=0) = 88\,MeV$, as well as $\langle \bar{q}q(m=0)\rangle = (255\,MeV)^3$, was used as the input in order to fix the parameters (ρ, R). The comparison between the solid curve and the long-dashed one shows that the effect of the NLO-corrections grows with m and is about 40% at $m = 0.1\,GeV$.

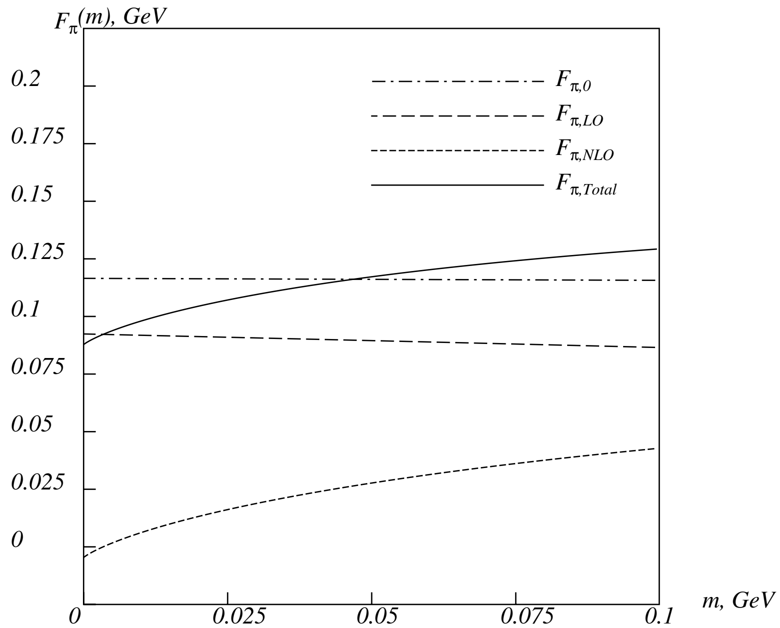


Figure 25: m-dependence of the pion decay constant F_π. The long-dashed curve is the LO contribution, the short-dashed curve is the NLO contribution, the solid curve is the total LO+NLO contribution. The dot-dashed line represents the leading-order in $1/N_c$-expansion result, evaluated with the mass M_0 from (116) (see text).

The pion mass M_π.

Effective action Γ_{eff}, Eq. (125) at $a = 0$ has a meaning of inverse π-meson propagator at small external momentum $q \sim M_\pi$ with account of meson loops. For our purpose it is sufficient to have only the $\mathcal{O}(q^0)$ and $\mathcal{O}(q^2)$ terms.

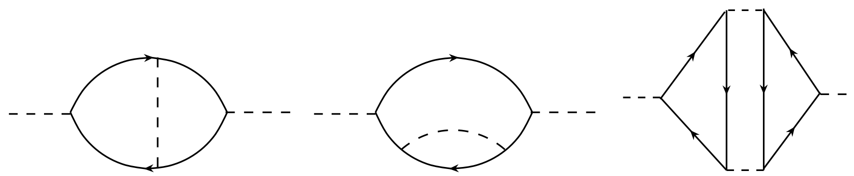


Figure 26: Diagrams corresponding to the meson loop terms.

The value of the pion mass M_π is defined as a pole position in the propagator. Evaluating F^2_{uu}-term in (125) with account of meson loops we obtain the pion mass

$$M_\pi^2 = m\left(\left(3.49 + \frac{1.63}{N_c}\right) + m\left(15.5 + \frac{18.25}{N_c} + \frac{13.5577}{N_c}\ln m\right)\right) \quad (128)$$

The $M_\pi(m)$-dependence of the pion mass is shown on the Fig.27. For the sake of comparison, we also plotted the value $M_{\pi,0}$ which one would get using LO formulae with the mass $M_0(m)$ taken from (116). Altogether, for this observables the NLO-corrections turn out to be small.

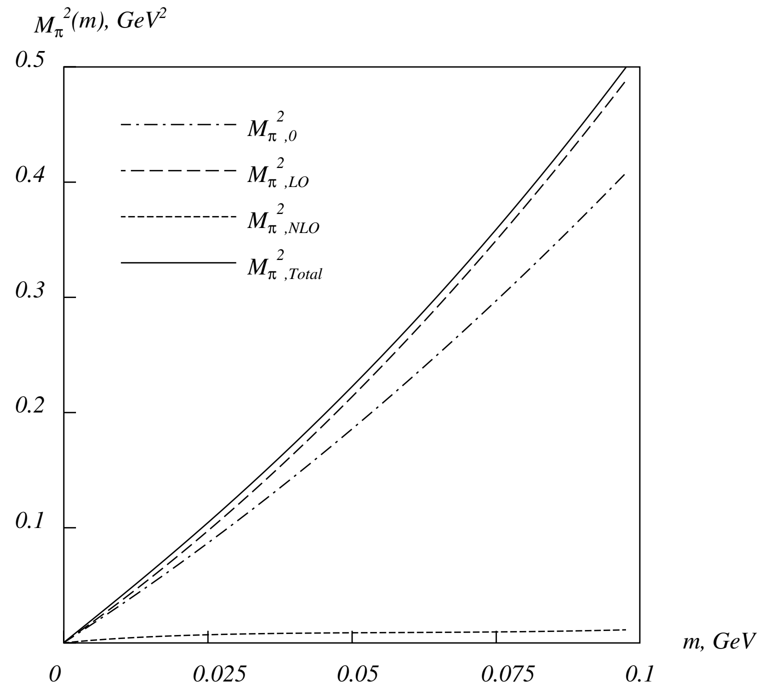


Figure 27: m-dependence of the pion mass M_π. The long-dashed curve is the LO contribution, the short-dashed curve is the NLO contribution, the solid curve is the total LO+NLO contribution. The dot-dashed line represents the leading-order in $1/N_c$-expansion result, evaluated with the mass M_0 from (116) (see text).

Finite width and "tensor" term corrections

In the previous sections it was assumed that the instanton size distribution has a zero width, i.e. $d(\rho) = \delta(\rho - \bar{\rho})$ and all instantons have the same size $\bar{\rho}$. This approximation is justified by the small parameter $1/N_c$, i.e. $\frac{\langle\rho^2\rangle - \langle\rho\rangle^2}{\langle\rho\rangle^2} \sim \mathcal{O}\left(\frac{1}{N_c}\right)$. For numerical evaluations we take the value $\delta\rho^2 = \langle\rho^2\rangle - \langle\rho\rangle^2 \approx \frac{0.5599\, GeV^{-2}}{N_c}$ which follows from the two-loop size distribution. Since we are interested in all $1/N_c$ corrections, we must take the finite width into account. To do this, we must return to the formula (99). Additional integration over ρ doesn't change the exponentiation procedure, and we get the standard $2N_f$-interaction term in the effective action S. However, for bosonization we should slightly modify the standard procedure.

The final results for the corrections are

$$\delta F_\pi^2 = \left(0.00045 - 0.0037m + 0.0036m^2\right) [GeV^2] \tag{129}$$
$$\delta\langle \bar{q}q\rangle = \left(-0.00045 + 0.011m - 0.062m^2\right) [GeV^3]$$

Thus we can see that these corrections are relatively small, $\approx 5\%$ for F_π^2 and $\approx 2.6\%$ for $\langle \bar{q}q\rangle$.

The modification of the dynamical mass $M(p)$ is shown in the Fig. 28. We can see that for $p = 0$ the increase of the dynamical quark mass is $\delta M/M = 10\%$.

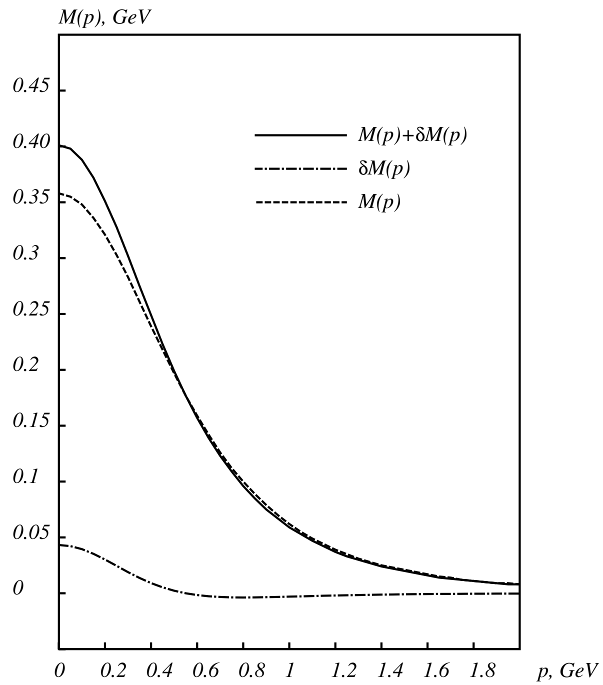


Figure 28: Change of the p-dependence of the constituent quark mass $M(p)$ due to the finite width corrections. The dashed curve is the contribution of the leading order result, the dot-dashed curve is the contribution of the finite width correction, the solid curve is a total result

"Tensor" terms contribution to the axial currents correlator.

This contribution may be represented as a Feynman diagram shown in the Fig.29.

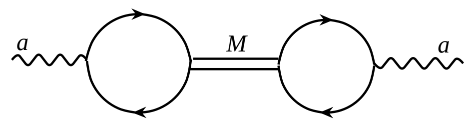


Figure 29: Contribution to the axial currents correlator. The intermediate state is the tensor meson $\Phi_{\mu\nu}$, the diagram is $1/N_c$-correction.

Straightforward evaluation of the tensor-axial coupling is

$$2i\epsilon_{\mu\nu\rho\lambda}\Phi_{\mu\nu}q_\rho a_\lambda(q) \times c_A, \tag{130}$$

$$c_A = -8N_c \int \frac{d^4p}{(2\pi)^4} \frac{2\mu^3(p) + pMf(p)f'(p)(p^2 - 3\mu^2(p))}{(p^2 + \mu^2(p))^2}$$

and the total contribution of the diagram in the Fig. 29 is

$$Fig.\ 29 \sim 8c_A^2 q^2 \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) \times \tag{131}$$

$$\left(1 - \left(\frac{\beta}{\alpha \langle \sigma \rangle} \right)^2 2N_c \int \frac{d^4p}{(2\pi)^4} \frac{M^2 F^4(p) \mu^2(p)}{(p^2 + \mu^2(p))^2} \right)^{-1}$$

Thus we can see that this diagram is just $O(q^2)$-correction to the axial correlator. Since we are interested only in the LO over q^2 (evaluation of F_π), we should not evaluate this diagram.

Gasser-Leutwyler couplings

According to [68], the low-energy constants $\bar{l}_i$ of the chiral lagrangian may be extracted from the $\mathcal{O}(m)$-corrections to physical quantities, e.g.

$$M_\pi^2 = m_\pi^2 \left(1 - \frac{m_\pi^2}{32\pi^2 F^2}\bar{l}_3 + ...\right), \quad F_\pi^2 = F^2 \left(1 + \frac{m_\pi^2}{8\pi^2 F^2}\bar{l}_4 + ...\right), \tag{132}$$

where M_π, F_π are the pion mass and decay constants, $m_\pi^2 = 2\,m\,B$ and B, F are the phenomenological parameters of the chiral lagrangian. Using our results

(127, 128), we can obtain (m is given in GeV)

$$F^2 = 0.00284777 N_c - 0.000868917 + \mathcal{O}\left(\frac{1}{N_c}\right), \tag{133}$$

$$B = 1.7467 + \frac{0.8183}{N_c} + \mathcal{O}\left(\frac{1}{N_c^2}\right) \tag{134}$$

$$\bar{l}_3 = 0.0738267 - 1.14251\, N_c - 0.999\, \ln m + \mathcal{O}\left(\frac{1}{N_c}\right) \tag{135}$$

$$\bar{l}_4 = -0.0793814\, N_c + 0.0187608 - 1.000\, \ln m + \mathcal{O}\left(\frac{1}{N_c}\right) \tag{136}$$

which gives

$$F = 88\, MeV,\ B = 2.019\, GeV,\ \bar{l}_3 = 1.84,\ \bar{l}_4 = 4.98 \tag{137}$$

at $m = 0.0055\ GeV$, corresponding $M_\pi = 0.142\ GeV$, $F_\pi = 0.0937\ GeV$.

The values of $F, -\langle\bar{q}q(m=0)\rangle = -F^2B$ in (137) were taken as input when we fixed the parameters $\rho = 0.350fm,\ R = 0.856fm$. Our values of $(\bar{l}_3, \bar{l}_4)$ should be compared with the phenomenological estimates [69, 70] as well as lattice predictions [71, 72] given in Table 1.

	χPT [73, 68, 70]	MILC [74]	Del Debbio *et. al.* [72]	ETM [75]	Our prediction
$\bar{l}_3$	2.9 ± 2.4	0.6 ± 1.2	3.0 ± 0.5	3.62 ± 0.12	1.84
$\bar{l}_4$	4.4 ± 0.2	3.9 ± 0.5	—	4.52 ± 0.06	4.98

Table 1: Estimates and predictions of the low-energy constants. The first column contains phenomenological estimates, the next three columns are lattice results from different collaborations, the last column contains our results. The first four columns of the table are taken from [70].

Nonperturbative QCD. Discussion.

1. The would-be linear confining potential of the pure glue world is necessarily screened by pion production at very moderate separations between quarks. Therefore, light hadrons need not be sensitive to confinement forces but rather to the dynamics of the spontaneous chiral symmetry breaking (SCSB).

2. Very likely, the SCSB is driven by instantons – large non-perturbative fluctuations of the gluon field having the meaning of tunneling. The SCSB is due to 'hopping' of quarks from one randomly situated instanton to another, each time flipping the helicity. The instanton theory of the SCSB is in agreement with the low-energy phenomenology ($cf.$ the chiral condensate $< \bar{\psi}\psi >$, the dynamical quark mass $M(p)$, F_π, $m_{\eta'}$...) and seems to be confirmed by direct lattice methods. Furthermore, lattice simulations indicate that instantons alone are responsible for the properties of lightest hadrons $\pi, \rho, N, ...$.

3. Instantons induce not only very strong non-perturbative quark interactions but also new and interesting vertices with an additional gluon emission. In particular, they induce a large anomalous chromomagnetic moment which can play an important role in soft high-energy hadron scattering, $e.g.$ in spin phenomena.

4. Summing up instanton-induced quark interactions in baryons leads to the Chiral Quark–Soliton Model where baryons appear to be bound states of constituent quarks pulled together by the chiral field. The model enables one to compute numerous parton distributions, as well as 'static' characteristics of baryons – with no fitting parameters whatsoever.

5. For highly excited baryons $(m = 1.5 - 3\,\mathrm{GeV})$ the relative importance of confining forces $vs.$ those of the SCSB may be reversed. One can view a large-spin J resonance as due to a short-time stretch of an unstable string or, alternatively, as a rotating elongated pion cloud [32]. What picture is more

adequate is a question to experiment. In the first case the dominant decay is on the average of the type $\mathrm{Bar}_J \to \mathrm{Bar}_{\sim J/2} + \mathrm{Mes}_{\sim J/2}$; in the second case it is mainly a cascade $\mathrm{Bar}_J \to \mathrm{Bar}_{J-1} + \pi \to \mathrm{Bar}_{J-2} + \pi\pi \to$ ... Studying resonances can elucidate the relation between chiral and confining forces.

6. One of the aim of our work was the study of the pion physics beyond the chiral limit in the framework of the instanton vacuum model. We found the generating functional of the hadronic correlators with account of $\mathcal{O}(1/N_c,\, m,\, m/N_c,\, m\ln m/N_c)$-corrections and exploited it for evaluation of the corrections to different physical observables. The corrections considered in this paper include meson loops, finite width of instanton size distribution and quark-quark tensor interactions term. In contrast to the expectations, we found that numerically the $1/N_c$-corrections to dynamical quark mass are large and mostly come from meson loops. As a consequence, we have large $1/N_c$-corrections to all the other quantities. To provide the values of

$F_\pi(m = 0), \langle\bar{q}q(m = 0)\rangle$ in agreement with χPT, we offer a new set of parameters $\rho = 0.350 fm,\ R = 0.856 fm$. Remarkably, this set of parameters is still in agreement with current phenomenological and lattice estimates.

It was evaluated the $F_\pi(m)$ and $M_\pi(m)$-dependence with account of $O(1/N_c, m, m/N_c, m/N_c \ln m)$-corrections. From comparison with χPT we extract the values of the low energy constants $\bar{l}_3, \bar{l}_4$. Our results for the values of $\bar{l}_3$ and $\bar{l}_4$ are in a satisfactory agreement with phenomenological as well as lattice estimates (See Table 1). This means that the instanton vacuum is applicable for understanding of the low-energy physics, at least on the qualitative level. Evaluation of the other LEC's is in progress.

References

[1] C. Amsler et al. (Particle Data Group), *The Review of Particle Physics*,

Physics Letters B667(2008)1, http://pdg.lbl.gov/

[2] M. Gell-Mann, Phys. Lett. 8 (1964) 214;
G. Zweig, CERN preprints CERN-TH 401 and 412 (1964), unpublished.

[3] M. Gell-Mann, Acta Phys. Austriaca Suppl. 9 (1972) 733;
M. Han and Y. Nambu, Phys. Rev. 139B (1965) 1006.

[4] S.L. Adler and W.A. Bardeen, Phys. Rev. 182 (1969) 1517.

[5] J.D. Bjorken, Phys. Rev. 179 (1969) 1547.

[6] C.J. Callan and D.J. Gross, Phys. Rev. Lett. 22 (1969) 156.

[7] W.B. Attwood, in Proc. 1979 SLAC Summer Institute on Particle Physics (SLAC-224, 1980), ed. A. Mosher, Vol. 3.

[8] I.J.R. Aitchison and A.J.G. Hey, *Gauge Theories in Particle Physics*, Graduate Student Series in Physics (IOP Publishing Ltd, Bristol, 1989).

[9] D.H. Perkins, *Introduction to High Energy Physics* (Addison–Wesley, Menlo Park, 1980) 3^{rd} edition.

[10] H. Fritzsch and M. Gell-Mann, Proc. XVI Int. Conf. on High Energy Physics, eds. J.D. Jackson and A. Roberts (Fermilab, 1972), Vol. 2, p. 135;
H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Lett. 47B (1973) 365.

[11] T. Kinoshita (editor), *Quantum Electrodynamics*, Advanced Series on Directions in High Energy Physics, Vol. 7 (World Scientific, Singapore, 1990);
T. Kinoshita, Phys. Rev. D47 (1993) 5013.

[12] M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory*, Addison–Wesley, 1995.

[13] L.D. Faddeev and A.A. Slavnov, Gauge fields: Introduction to Quantum Theory, Benjamin/Cummings, 1980.

[14] V.B. Berestetsky, E.M. Lifshitz, and L.P. Pitaevsky, *Quantum Electrodynamics*, Pergamon, Oxford, 1982.

[15] A. Belavin, A. Polyakov, A. Schwartz and Yu. Tyupkin, *Phys. Lett.* **59** (1975) 85

[16] A. Polyakov, *Nucl. Phys.* B 120 (1977) 429

[17] G. 't Hooft, *Phys. Rev. Lett.* 37 (1976) 8;
Phys. Rev. D 14 (1976) 3432; Erratum: *ibid.* D 18 (1978) 2199

[18] C. Callan, R. Dashen and D. Gross, *Phys. Rev.* D 17 (1978) 2717

[19] E. Shuryak, *Nucl. Phys.* B 203 (1982) 93, 116, 140

[20] E.M. Ilgenfritz and M. Müller-Preussker, *Nucl. Phys.* B184 (1981) 443

[21] D. Diakonov and V. Petrov, *Nucl. Phys.* B 245 (1984) 259

[22] M. Shifman, A. Vainshtein and V. Zakharov, *Nucl. Phys.* B 147 (1979) 385

[23] D. Diakonov, *The U(1) problem and instantons*, in: Gauge Theories of the Eighties, Lecture Notes in Physics, Springer-Verlag (1983) p. 207

[24] D. Diakonov and V. Petrov, *Phys. Lett.* B 147 (1984) 351

[25] D. Diakonov and V. Petrov, *Sov. Phys. JETP* 62 (1985) 204; *Sov. Phys. JETP* 62 (1985) 431; *Nucl. Phys.* B 272 (1986) 457

[26] D. Diakonov and V. Petrov, *Spontaneous breaking of chiral symmetry in the instanton vacuum*, preprint LNPI-1153 (1986), in: Hadron matter under extreme conditions, Kiew (1986) p. 192;
D. Diakonov and V. Petrov, in: Quark Cluster Dynamics, Lecture Notes in Physics, Springer-Verlag (1992) p. 288

[27] P. Pobylitsa, *Phys. Lett.* B 226 (1989) 387

[28] D. Diakonov, V. Petrov and P. Pobylitsa, *Phys. Lett.* B 226 (1989) 372

[29] D. Diakonov, in: Skyrmions and Anomalies, World Scientific (1987) p. 27

[30] G. Bali, K. Schilling and A. Wachter, in: Confinement 95, eds. H. Toki *et al.*, World Scientific (1995) p. 82, `hep-lat/9506017`

[31] F. Karsch, E. Laermann and A. Peikert, *Nucl. Phys.* **B605**, 579 (2001), `hep-lat/0012023`

[32] D. Diakonov and V. Petrov, *Rotating chiral solitons lie on linear Regge trajectories*, preprint LNPI-1394 (1988); see also D. Diakonov, *Acta Phys. Polon.* B 25 (1994) 17.

[33] E. Shuryak, *Rev. Mod. Phys.* 65 (1993) 1

[34] V. Novikov, M. Shifman, A. Vainshtein and V. Zakharov, *Nucl. Phys.* B 191 (1981) 301

[35] E. Shuryak and J. Verbaarschot, *Nucl. Phys.* B 410 (1993) 55;

T. Schäfer, E. Shuryak and J. Verbaarschot, *Nucl. Phys.* B 412 (1994) 143; T. Schäfer and E. Shuryak, *Phys. Rev.* D50 (1994) 478

[36] M.-C. Chu, J. Grandy, S. Huang and J. Negele, *Phys. Rev. Lett.* 70 (1993) 225; *Phys. Rev.* D49 (1994) 6039;
J. Negele, *Nucl. Phys. Proc. Suppl.* 73 (1999) 92, `hep-lat/9810053`

[37] T. Schäfer and E. Shuryak, *Rev. Mod. Phys.* 70 (1998) 323, `hep-ph/9610451`

[38] L.D. Faddeev, *Looking for multi-dimensional solitons* in: Non-local Field Theories, Dubna (1976)

[39] R. Jackiw and C. Rebbi, *Phys. Rev. Lett.* 37 (1976) 172

[40] J.D. Jackson and L.B. Okun, *Rev. Mod. Phys.* 73 (2001) 663, hep-ph/0012061

[41] D. Diakonov, M. Polyakov and C. Weiss, *Nucl. Phys.* B 461 (1996) 539, hep-ph/9510232

[42] C. Bernard, *Phys. Rev.* D 19 (1979) 3013

[43] A. Hasenfratz and P. Hasenfratz, *Nucl. Phys.* B 193 (1981) 210

[44] A. Vainshtein, V. Zakharov, V. Novikov and M. Shifman, *Sov. Phys. Uspekhi* 136 (1982) 553

[45] D. Diakonov and M. Polyakov, *Nucl. Phys.* B 389 (1993) 109

[46] D. Diakonov and V. Petrov, *Phys. Rev.* D 50 (1994) 266

[47] V. Fateev, I. Frolov and A. Schwartz, *Nucl. Phys.* B 154 (1979) 1;
V. Fateev, I. Frolov and A. Schwartz, *Sov. J. Nucl. Phys.* 30 (4) (1979) 590

[48] B. Berg and M. Lüscher, *Comm. Math. Phys.* 69 (1979) 57

[49] D. Diakonov and M. Maul, *Nucl. Phys.* B 571 (2000) 91, `hep-th/9909078`

[50] P. de Forcrand, M. G. Perez, I.-O. Stamatescu, *Nucl. Phys.* B 499 (1997) 409, `hep-lat/9701012`

[51] T. DeGrand, A. Hasenfratz and T. Kovacs, *Nucl. Phys.* B 505 (1997) 417, `hep-lat/9705009`

[52] C. K. Lee and W. A. Bardeen, Nucl. Phys. B **153** (1979) 210.

[53] M. M. Musakhanov and F. C. Khanna, Phys. Lett. B **395** (1997) 298 [arXiv:hep-ph/9610418].

[54] E. D. Salvo and M. M. Musakhanov, Eur. Phys. J. C **5**, 501 (1998) [arXiv:hep-ph/9706537].

[55] M. Musakhanov, Eur. Phys. J. C **9** (1999) 235 [arXiv:hep-ph/9810295].

[56] M. Musakhanov, Nucl. Phys. A **699** (2002) 340.

[57] M. M. Musakhanov and H. C. Kim, Phys. Lett. B **572** (2003) 181 [arXiv:hep-ph/0206233].

[58] H. C. Kim, M. Musakhanov and M. Siddikov, Phys. Lett. B **608** (2005) 95 [arXiv:hep-ph/0411181].

[59] H. C. Kim, M. M. Musakhanov and M. Siddikov, Phys. Lett. B **633** (2006) 701 [arXiv:hep-ph/0508211].

[60] K. Goeke, M. M. Musakhanov and M. Siddikov, Phys. Rev. D **76** (2007) 076007 [arXiv:0707.1997 [hep-ph]].

[61] K. Goeke, H. C. Kim, M. M. Musakhanov and M. Siddikov, Phys. Rev. D **76** (2007) 116007 [arXiv:0708.3526 [hep-ph]].

[62] S. R. Coleman and E. Weinberg, Phys. Rev. D **7**(1973) 1888.

[63] R. Jackiw, Phys. Rev. D **9**(1974) 1686.

[64] E. N. Nikolov, W. Broniowski, C. V. Christov, G. Ripka and K. Goeke, Nucl. Phys. A **608** (1996) 411 [arXiv:hep-ph/9602274].

[65] P. O. Bowman, private communication $(M(m))$-dependence. See also P. O. Bowman, U. M. Heller, D. B. Leinweber, M. B. Parappilly, A. G. Williams and J. b. Zhang, Phys. Rev. D **71** (2005) 054507 [arXiv:hep-lat/0501019].

[66] P. Langacker, H. Pagels, Phys. Rev. D **8** (1973) 4595.

[67] V. Dmitrasinovic, H. J. Schulze, R. Tegen and R. H. Lemmer, Annals Phys. **238** (1995) 332

[68] J. Gasser and H. Leutwyler, Annals Phys. **158** (1984) 142.

[69] H. Leutwyler, arXiv:hep-ph/0612112.

[70] H. Leutwyler, arXiv:0706.3138 [hep-ph].

[71] C. Aubin *et al.* [MILC Collaboration], Phys. Rev. D **70** (2004) 114501 [arXiv:hep-lat/0407028].

[72] L. Del Debbio, L. Giusti, M. Luscher, R. Petronzio and N. Tantalo, JHEP **0702** (2007) 056 [arXiv:hep-lat/0610059].

[73] G. Colangelo, J. Gasser and H. Leutwyler, Nucl. Phys. B **603** (2001) 125 [arXiv:hep-ph/0103088].

[74] C. Bernard *et al.*, arXiv:hep-lat/0611024.

[75] Ph. Boucaud *et al.* [ETM Collaboration], arXiv:hep-lat/0701012.

[76] A. E. Dorokhov, Eur. Phys. J. C **42** (2005) 309 [arXiv:hep-ph/0505007].