35. Arnold V I, Shandarin S F, Zeldovich Ya B *Geophys. Astrophys. Fluid Dynamics* **20** 111 (1982)
36. Bini D, Cherubini C, Jantzen R T *Class. Quantum Grav.* **24** 5627 (2007)
37. Petrov A Z *Prostranstva Einshteina* (Einstein Spaces) (Moscow: GIFML, 1961) [Translated into English (Oxford: Pergamon Press, 1969)]
38. Belinskii V A, Khalatnikov I M *Zh. Eksp. Teor. Fiz.* **63** 1121 (1972) [*Sov. Phys. JETP* **36** 591 (1973)]
39. Damour T "Cosmological singularities, billiards and Lorentzian Kac-Moody algebras", gr-qc/0412105
40. Damour T, Nicolai H "Symmetries, singularities and the de-emergence of space", arXiv:0705.2643
41. Kamenshchik A, Moschella U, Pasquier V *Phys. Lett. B* **511** 265 (2001)

PACS numbers: **11.15.–q**, 11.30.Qc, 12.38.Aw
DOI: 10.1070/PU2008v051n06ABEH006551
DOI: 10.3367/UFNr.0178.200806k.0647

Axial anomaly in quantum electro- and chromodynamics and the structure of the vacuum in quantum chromodynamics

B L Ioffe

1. Introduction

In this report, I discuss the current state of the problem of the axial anomaly in quantum electrodynamics (QED) and quantum chromodynamics (QCD) and the relation of the axial anomaly to the structure of the vacuum in QCD. In QCD, the vacuum average of the axial anomaly is proportional to a new quantum number n, the winding number. There are an infinite number of vacuum states $|n\rangle$. The transition amplitudes between these states are amplitudes of tunnel transitions along certain paths in the space of gauge fields. I show that the axial anomaly condition implies that there are zero modes of the Dirac equation for a massless quark and that spontaneous chiral symmetry breaking occurs in QCD, which leads to the formation of a quark condensate. The axial anomaly can be represented in the form of a sum rule for the structure function in the dispersion representation of the axial – vector – vector (AVV) vertex. On the basis of this sum rule, we calculate the width of the $\pi^0 \to 2\gamma$ decay with an accuracy of 1.5%.

2. The definition of an anomaly

We suppose that the classical field-theory Lagrangian has a certain symmetry, i.e., is invariant under transformations of the fields corresponding to this symmetry. According to the Noether theorem, the symmetry corresponds to a conservation law. An anomaly is a phenomenon in which the given symmetry and the conservation law are violated as we pass to quantum theory. The reason for this violation lies in the singularity of quantum field operators at small distances, such that finding the physical quantities requires fixing not only the Lagrangian but also the renormalization procedure. (See reviews dealing with various anomalies in Refs [1 – 4].)

There are two types of anomalies, internal and external. In the first case, the gauge invariance of the classical Lagrangian is broken at the quantum level, the theory becomes unrenormalizable, and is not self-consistent. This problem can be resolved by a special choice of fields in the Lagrangian, for which all the internal anomalies cancel. (Such an approach is used in the standard model of electroweak interaction and is known as the Glashow – Illiopoulos – Maiani mechanism.) External anomalies emerge as a result of the interaction between the fields in the Lagrangian and external sources. It is these anomalies that appear in quantum electrodynamics and quantum chromodynamics; they are discussed in what follows. We show that anomalies play an important role in QED and especially in QCD. Hence, the term 'anomaly' should not mislead us — it is a normal and important ingredient of most quantum field theories.

3. Axial anomaly in QED

The Dirac equation for the electron in an external electromagnetic field $A_\mu(x)$ is

$$\mathrm{i}\gamma_\mu \frac{\partial\psi(x)}{\partial x_\mu} = m\psi(x) - e\gamma_\mu A_\mu(x)\,\psi(x)\,. \tag{1}$$

The axial current is defined as

$$j_{\mu5}(x) = \bar\psi(x)\,\gamma_\mu\gamma_5\psi(x)\,. \tag{2}$$

Its divergence calculated in classical theory, i.e., with the use of Eqn (1), is

$$\partial_\mu j_{\mu5}(x) = 2\mathrm{i}m\bar\psi(x)\,\gamma_5\psi(x) \tag{3}$$

and tends to zero as $m \to 0$. In quantum theory, the axial current must be redefined because $j_{\mu5}(x)$ is the product of two local fermionic fields, with the result that it is singular when both fields act at the same point. (Naturally, a similar statement is true for a vector current.) To achieve a meaningful approach, we split the points where the two fermionic fields act by a distance ε, such that

$$j_{\mu5}(x,\varepsilon) = \bar\psi\left(x+\frac{\varepsilon}{2}\right)\gamma_\mu\gamma_5 \times \exp\left[\mathrm{i}e\int_{x-\varepsilon/2}^{x+\varepsilon/2} \mathrm{d}y_\alpha A_\alpha(y)\right]\psi\left(x-\frac{\varepsilon}{2}\right), \tag{4}$$

and take $\varepsilon \to 0$ in the final result. The exponential factor in (4) is introduced to ensure the local gauge invariance of $j_{\mu5}(x,\varepsilon)$. The divergence of axial current (4) has the following form (we use Eqn (1) and keep only the terms that are linear in ε):

$$\partial_\mu j_{\mu5}(x,\varepsilon) = 2\mathrm{i}m\bar\psi\left(x+\frac{\varepsilon}{2}\right)\gamma_5\psi\left(x-\frac{\varepsilon}{2}\right) - \mathrm{i}e\varepsilon_\alpha\bar\psi\left(x+\frac{\varepsilon}{2}\right)\gamma_\mu\gamma_5\psi\left(x-\frac{\varepsilon}{2}\right)F_{\alpha\mu}\,, \tag{5}$$

where $F_{\alpha\mu}$ is the electromagnetic field strength. For simplicity, we assume that $F_{\mu\nu} = \text{const}$ and use the fixed-point gauge (the Fock – Schwinger gauge) $x_\mu A_\mu(x) = 0$. Then $A_\mu(x) = (1/2)\,x_\nu F_{\nu\mu}$. We calculate the vacuum average of (5). To calculate the right-hand side of (5), we use the expression for the electron propagator in an external electromagnetic field

($\not{x} = x_\mu \gamma_\mu$):

$$S(x) = \frac{\mathrm{i}}{2\pi^2}\left[\frac{\not{x}}{x^4} + \mathrm{i}\,\frac{m}{2x^2} + \frac{1}{16x^2}\, eF_{\mu\nu}(\not{x}\sigma_{\mu\nu} + \sigma_{\mu\nu}\not{x})\right], \quad (6)$$

$$\sigma_{\mu\nu} = \frac{\mathrm{i}}{2}\,(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)\,. \quad (7)$$

Vacuum averaging involves first-order corrections in e^2. Substituting Eqn (6) in Eqn (5) and ignoring the electron mass, we obtain

$$\langle 0|\,\partial_\mu j_{\mu 5}|0\rangle = \frac{e^2}{4\pi^2}\, F_{\alpha\mu}F_{\lambda\sigma}\varepsilon_{\beta\mu\lambda\sigma}\,\frac{\varepsilon_\alpha\varepsilon_\beta}{\varepsilon^2}\,. \quad (8)$$

Because there can be no preferred direction in space–time, the limit $\varepsilon \to 0$ can be achieved in a symmetric manner, and we have

$$\partial_\mu j_{\mu 5} = \frac{e^2}{8\pi^2}\, F_{\alpha\beta}\tilde{F}_{\alpha\beta}\,, \quad (9)$$

where

$$\tilde{F}_{\alpha\beta} = \frac{1}{2}\,\varepsilon_{\alpha\beta\lambda\sigma}F_{\lambda\sigma} \quad (10)$$

is the dual electromagnetic field tensor. In Eqn (9), the symbol of vacuum averaging is dropped because in the e^2-order, this equation can be considered an operator equation. Equation (9) is known as the Adler–Bell–Jackiw anomaly [5–8].

To better understand the origin of an anomaly, we consider the same problem in the momentum space. In QED, the matrix element of the transition of an axial current with a momentum q into two real or virtual photons with momenta p and p' is described by the diagrams in Fig. 1. The matrix element is

$$T_{\mu\alpha\beta}(p, p') = \Gamma_{\mu\alpha\beta}(p, p') + \Gamma_{\mu\beta\alpha}(p', p)\,, \quad (11)$$

$$\Gamma_{\mu\alpha\beta}(p, p') = -e^2 \int \frac{\mathrm{d}^4 k}{(2\pi)^4}\,\mathrm{Tr}\big[\gamma_\mu\gamma_5(\not{k} + \not{p} - m)^{-1} \times \gamma_\alpha(\not{k} - m)^{-1}\gamma_\beta(\not{k} - \not{p}' - m)^{-1}\big]\,. \quad (12)$$

Integral (12) linearly diverges. In a linearly divergent integral, the important terms are the surface terms, which emerge as a result of integrating over an infinitely remote surface in the momentum space. (This becomes especially clear when the vectors q, p, and p' are space-like and the integration contour over k_0 can be rotated to the imaginary axis, $k_0 \to \mathrm{i}k_4$, such that integration over k is carried out in Euclidean space.) The result of calculations depends on the way k is chosen: we can displace k by an arbitrary constant vector a_λ, i.e., $k_\lambda \to k_\lambda + a_\lambda$. Amplitude (11) must satisfy the conditions needed for the vector-current conservation: $p_\alpha T_{\mu\alpha\beta}(p, p') = 0$ and $p'_\beta T_{\mu\alpha\beta}(p, p') = 0$.

We try to choose the vector a_λ such that the conditions for both axial- and vector-current conservation are satisfied. We parameterize a_λ as $a_\lambda = (a+b)\,p_\lambda + bp'_\lambda$. The result of calculations shows that both conditions cannot be satisfied simultaneously: the vector-current conservation can be achieved at $a = -2$, while the axial-current conservation requires that $a = 0$ [8, 9]. The vector-current conservation is a necessary condition for the existence of QED. Hence, we select $a = -2$. The divergence of the axial current is

$$q_\mu T_{\mu\alpha\beta}(p, p') = \left[2mG(p, p') - \frac{e^2}{2\pi^2}\right]\varepsilon_{\alpha\beta\lambda\sigma}\,p_\lambda\, p'_\sigma\,. \quad (13)$$

Here, we restore the term proportional to the electron mass and define $G(p, p')$ as

$$\big\langle p, \varepsilon_\alpha;\, p', \varepsilon'_\beta\,|\,\bar{\psi}\gamma_5\psi\,|\,0\big\rangle = G(p, p')\,\varepsilon_{\alpha\beta\lambda\sigma}\,p_\lambda\,p'_\sigma\,, \quad (14)$$

with ε_α and ε'_β being the polarizations of the two photons. The fact that the axial current is not conserved, stated in Eqn (13), is equivalent to Eqn (9). Our discussion of the axial anomaly in QED was limited to terms of the order e^2. Adler and Bardeen have proved (see Refs [5, 6, 10]) that the radiative corrections caused by the photon lines connecting different points inside the triangle diagrams in Fig. 1 do not alter the anomaly equation. However, the diagram in Fig. 2 yields a nonvanishing, albeit small, correction of the order e^6 to this condition [11].

4. The axial anomaly and its relation to the structure of the vacuum in quantum chromodynamics

In QCD with massless quarks, the axial anomaly is described by a formula similar to (9):

$$\partial_\mu j^a_{\mu 5} = \frac{e^2}{8\pi^2}\, e_{\mathrm{q}}^2 N_{\mathrm{c}}\, F_{\mu\nu}\,\tilde{F}_{\mu\nu}\,. \quad (15)$$

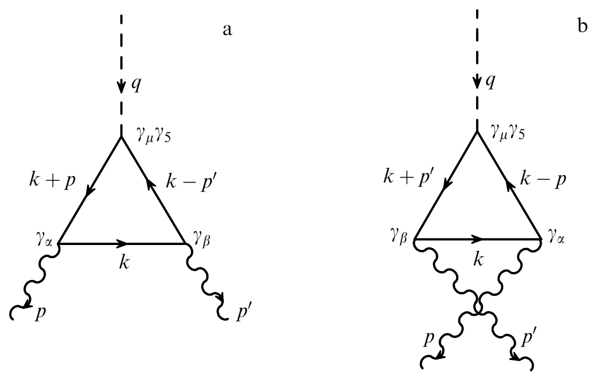

Figure 1. Feynman diagrams describing the transition of an axial current with a momentum q into two real or virtual photons with momenta p and p', $q = p + p'$: (a) the direct diagram, and (b) the crossing diagram.

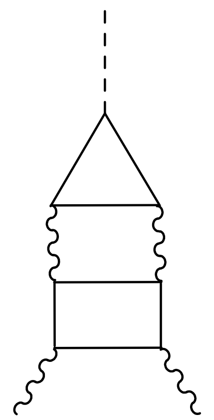

Figure 2. The e^6-order correction to the Adler–Bell–Jackiw anomaly in QED.

Here, $N_c = 3$ is the number of colors and e_q is the quark charge. (We wrote Eqn (15) for one massless quark.) There is also another anomaly in QCD, where the external fields are not electromagnetic but gluonic:

$$\partial_\mu j_{\mu 5} = \frac{\alpha_s N_c}{4\pi} G^n_{\mu\nu} \tilde{G}^n_{\mu\nu} , \tag{16}$$

where $G^n_{\mu\nu}$ is the gluonic field strength and $\tilde{G}^n_{\mu\nu}$ is its dual. Equation (16) can be considered an operator equation, and the fields $G^n_{\mu\nu}$ and $\tilde{G}^n_{\mu\nu}$ can be considered the fields of virtual gluons. In the same way as in QED, perturbative corrections to (16) begin at α_s^3 and are described by a diagram similar to the one shown in Fig. 2. In QCD, however, the coupling constant is large, with the result that the contribution provided by this diagram is not small; the contribution of diagrams obtained from the one in Fig. 2 by attaching additional quark and gluon loops are not small either. Obviously, in QCD, the octet axial current

$$j^i_{\mu 5} = \sum_q \bar{\psi}_q \gamma_\mu \gamma_5 \frac{\lambda^i}{2} \psi_q , \quad i = 1, ..., 8 \tag{17}$$

is conserved in the absence of an electromagnetic field. (Here, the λ^i are the Gell-Mann matrices, and summation is over the flavors of the light quarks, $q = u, d, s$.) The singlet axial current

$$j^{(0)}_{\mu 5} = \sum_q \bar{\psi}_q \gamma_\mu \gamma_5 \psi_q \tag{18}$$

contains the anomaly

$$\partial_\mu j^{(0)}_{\mu 5} = 3 \frac{\alpha_s N_c}{4\pi} G^n_{\mu\nu} \tilde{G}^n_{\mu\nu} . \tag{19}$$

In view of the spontaneous breaking of chiral symmetry, the pseudoscalar mesons belonging to the octet (π, K, η) are massless (in the $m_q \to 0$ approximation), while the SU(3) singlet η' is massive. In this way, the presence of an anomaly solves what is known as the U(1) problem [12].

I now discuss the important assertion that exists in QCD and relates the structure of the anomaly to the structure of the vacuum in this theory. Because we deal with the existence of degenerate vacua and tunnel (underbarrier) transitions between them, it is convenient (just as in quantum mechanics) to introduce imaginary time by setting $t = x_0 = -ix_4$; we thus operate in the Euclidean space, where $x^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$. In the Euclidean space, the action integral

$$S = \frac{1}{4} \int d^4 x G^2_{\mu\nu} \tag{20}$$

is positive. (We temporarily ignore the quark contribution.) The transition amplitudes are determined by the matrix elements of $\exp(-S)$. A theorem first proved by Belavin, Polyakov, Schwartz, and Tyupkin [13] states that

$$\frac{\alpha_s}{8\pi} \int d^4 x G^n_{\mu\nu} \tilde{G}^n_{\mu\nu} = n , \tag{21}$$

where n is an integer known as the winding number. Here, we do not prove this theorem in detail; instead, we mention its main points. The integrand in (21) can be written as the total derivative

$$G^n_{\mu\nu} \tilde{G}^n_{\mu\nu} = \partial_\mu K_\mu , \tag{22}$$

$$K_\mu = \varepsilon_{\mu\nu\gamma\delta} \left(A^n_\nu G^n_{\gamma\delta} - \frac{1}{3} f^{nmp} A^n_\nu A^m_\gamma A^p_\delta \right) . \tag{23}$$

When x^2 is large, $G_{\mu\nu}(x)$ decreases faster than $1/x^2$ (i.e., there is no physical field), and A^n_μ is a pure-gauge field. Then, we can drop the first term in the right-hand side of (23) and keep only the second term in the general expression for the gauge transformation for A^n_μ,

$$A'_\mu = U^{-1} A_\mu U + i U^{-1} \partial_\mu U \tag{24}$$

(here, U is a unitary, unimodular matrix, $U^+ = U$, $|U| = 1$). We suppose that the field A^n_μ ($n = 1, 2, 3$) belongs to the SU(2) subgroup of the color group SU(3). This subgroup plays a special role in the SU(3) group because it is isomorphic to the spatial rotation group O(3). At this point, it is convenient to introduce matrix notation for the fields A_μ:

$$A_i = \frac{1}{2} g \tau^k A^k_i , \quad k = 1, 2, 3 ; \quad i = 1, 2, 3 . \tag{25}$$

Then, according to Eqns (22) and (23), we have

$$\int d^4 x G_{\mu\nu}(x) \tilde{G}_{\mu\nu}(x) = -i \frac{4}{3} \frac{1}{g^2} \int dV \varepsilon_{ikl} \operatorname{Tr}(A_i A_k A_l) . \tag{26}$$

Substituting the second term in the right-hand side of Eqn (24) in (26), we see that the integrand in (26) is a total derivative with respect to the spatial coordinates, and therefore reduces to an integral over an infinitely remote surface. Because $|U| = 1$ on this surface, the matrix U has the form

$$U = \exp\left(2\pi n \hat{r}_a \frac{\tau^a}{2i} \right) , \tag{27}$$

where n in an integer and $\hat{r}_a$ is a unit radius vector, $\hat{r}_a = r_a / |\mathbf{r}|$. The invariance of U under spatial rotations stems from the fact that each such rotation is accompanied by a gauge transformation, a rotation in the SU(2) group. When the right-hand side of Eqn (24) is substituted in (26), we see that Eqn (21) follows from (27). Theorem (21) also follows from general mathematical considerations, because the SU(2) group is mapped onto O(3); such a map is multivalued and is determined by the number of times the O(3) group is covered. We note that the fields corresponding to different n cannot be transformed into each other by a continuous transformation. In the perturbation theory, we always deal with fields corresponding to $n = 0$. The action integral in (2) can be written as

$$S = \frac{1}{4} \int d^4 x G^n_{\mu\nu} G^n_{\mu\nu} = \frac{1}{4} \int d^4 x \left[G^n_{\mu\nu} \tilde{G}^n_{\mu\nu} + \frac{1}{2} (G^n_{\mu\nu} - \tilde{G}^n_{\mu\nu})^2 \right] . \tag{28}$$

Because the last term in (28) is positive, the minimum of the action is achieved with fields satisfying the self-duality condition

$$G^n_{\mu\nu} = \tilde{G}^n_{\mu\nu} , \tag{29}$$

$$S_{\min} = \frac{1}{4} \int d^4 x G^n_{\mu\nu} \tilde{G}^n_{\mu\nu} = \frac{8\pi^2}{g^2} |n| = \frac{2\pi}{\alpha_s} |n| . \tag{30}$$

(Negative n correspond to anti-self-dual fields $G^n_{\mu\nu} = -\tilde{G}^n_{\mu\nu}$.) The solutions of the self-duality equation (for $n = 1$), which became known as instantons, were found in [13]. It follows from Eqn (30) that in QCD in the Euclidean space, there exists an infinite number of action minima. In Minkowski space, instantons are paths of tunnel transitions (in the field space) between vacua characterized by different winding numbers but having the same energies [14–16]. By examining $n(t)$, which transforms into the winding number as $t \to \pm\infty$, it can be shown that the instanton solutions correspond to $n(t \to -\infty) = 0$ and $n(t \to \infty) = 1$ and that the transition amplitude between vacuum states is [17]

$$\left\langle \frac{\Omega_{n=1}(t \to \infty)}{\Omega_{n=0}(t \to -\infty)} \right\rangle = \exp\left(-\frac{2\pi}{\alpha_s}\right). \tag{31}$$

5. Structure of the vacuum in quantum chromodynamics

Above, we showed that in QCD, there is an infinite number of vacua with the same energies, vacua that are characterized by the values of the winding number n. We let $\Omega(n)$ denote the wave function of such a vacuum and suppose that the wave functions are normalized, $\Omega^+(n)\,\Omega(n) = 1$, and form a complete system. The ambiguity in the wave function resides in the phase factor, $\Omega(n) = \exp(\mathrm{i}\theta_n)\,\Omega'(n)$. We separate the Euclidean space into two big parts and assume that the field strength in the space between these parts is zero and the potentials are pure gauge. Then, obviously,

$$\exp(\mathrm{i}\theta_{n_1+n_2})\,\Omega(n_1+n_2) = \exp(\mathrm{i}\theta_{n_1})\,\Omega(n_1)\exp(\mathrm{i}\theta_{n_2})\,\Omega(n_2)\,. \tag{32}$$

[Here, we drop the prime on $\Omega'(n)$.] Because

$$\Omega(n_1+n_2) = \Omega(n_1)\,\Omega(n_2)\,, \tag{33}$$

we have the equation

$$\theta_{n_1+n_1} = \theta_{n_1} + \theta_{n_2}\,, \tag{34}$$

which is solved by

$$\theta_n = n\theta\,. \tag{35}$$

Thus, the vacuum wave function in QCD is a linear combination of wave functions with different winding numbers:

$$\Omega(\theta) = \sum_n \exp(\mathrm{i}n\theta)\,\Omega(n)\,. \tag{36}$$

The state $\Omega(\theta)$ is known as the θ-vacuum. The vacuum state $\Omega(\theta)$ is similar to the Bloch state of an electron in a crystal, with θ acting as momentum. But in contrast to a Bloch state, all transitions between states with different θ are forbidden for the θ-vacuum. The vacuum state $\Omega(\theta)$ can be reproduced if the term

$$L_\theta = \frac{g^2\theta}{32\pi^2}\,G_{\mu\nu}\tilde{G}_{\mu\nu}\,. \tag{37}$$

is added to the QCD Lagrangian (in Minkowski space). The presence of this term in the Lagrangian demonstrates that θ is an observable. Term (37) violates the P- and CP-invariance. However, so far all attempts to discover the violation of CP-invariance in strong interactions have failed. The strongest bound on the value of θ has been found in searches of the neutron dipole moment, $\theta < 10^{-9}$ [18].

6. Zero eigenvalues of the Dirac equation for massless quarks as a consequence of the anomaly. Spontaneous breaking of chiral symmetry in quantum chromodynamics

We consider the Dirac equation for massless quarks in QCD in Euclidean space:

$$-\mathrm{i}\gamma_\mu\nabla_\mu\psi_k = \lambda_k\psi_k\,,\quad \nabla_\mu = \partial_\mu + \mathrm{i}g\,\frac{\lambda_n}{2}\,A^n_\mu\,. \tag{38}$$

From anomaly condition (16) with $n = 1$, we have

$$\int \mathrm{d}^4x\,\mathrm{Tr}\,\langle 0|\,\partial_\mu j_{\mu 5}(x)\,|0\rangle = \frac{g^2}{16\pi^2}\int \mathrm{d}^4x\langle 0|\,G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}\,|0\rangle = 2N_c\,. \tag{39}$$

The left-hand side of Eqn (39) can be written as an operator as follows:

$$\begin{aligned}\int \mathrm{d}^4x\,\mathrm{Tr}\,\langle 0|\,\partial_\mu j_{\mu 5}(x)\,|0\rangle &= -\int \mathrm{d}^4x\partial_\mu\,\mathrm{Tr}\,\langle 0|\,\mathrm{i}\,\nabla\!\!\!/^{-1}(x,x)\,\gamma_\mu\gamma_5\,|0\rangle \\ &= -\int \mathrm{d}^4x\nabla_\mu\,\mathrm{Tr}\left[\sum_k \frac{\psi_k(x)\,\psi^+_k(x)}{\lambda_k}\,\gamma_\mu\gamma_5\right] \\ &= -\int \mathrm{d}^4x\,\mathrm{Tr}\left[\sum_k \frac{\psi_k(x)\,\psi^+_k(x)}{\lambda_k}\,2\lambda_k\gamma_5\right].\end{aligned} \tag{40}$$

States with nonzero λ_k contribute nothing to (40) because each such state $\psi_k(x)$ corresponds to the state $\gamma_5\psi_k(x)$ with the eigenvalue $-\lambda_k$, and the two states are orthogonal. Thus, only the zero modes contribute, and hence we have

$$2\int \mathrm{d}^4x\,\mathrm{Tr}\left[\gamma_5\psi_0(x)\,\psi^+_0(x)\right] = -2N_c\,. \tag{41}$$

This implies that in the case where $n = 1$, i.e., in the instanton field, the zero mode is right-handed: the quark spin is directed along the quark momentum, $\gamma_5\psi_0 = -\psi_0$. (Actually, for a quark in the instanton field, only one right-handed zero mode exists, because spin is correlated with color and the factor N_c in the right-hand side of Eqn (41) disappears.) At $n = -1$, the resulting equation differs from (41) only in sign, i.e., a left-handed zero mode exists in an anti-instanton field. In the general case, we have the Atiyah–Singer theorem [19], according to which

$$n = n_L - n_R\,, \tag{42}$$

where n_L and n_R are the respective numbers of left- and right-hand zero modes. It follows from (41) that in an instanton field, the zero mode violates the chiral symmetry of the Lagrangian, i.e., the invariance under the transformations $\psi \to \gamma_5\psi$. (We note that in passing from the Euclidean metric to Minkowski space, the function ψ^+ is replaced by $\bar{\psi}$.) Thus, the presence of instantons is an indication that a quark condensate exists in the QCD vacuum:

$$\langle 0|\,\bar{\psi}\psi\,|0\rangle \neq 0\,, \tag{43}$$

which breaks the chiral symmetry of the Lagrangian. (Unfortunately, it is impossible to calculate the quark condensate on the basis of the instanton approach because this approach is meaningful only when the distances are small, while the condensate forms over large distances.)

The winding number n corresponds to the topological current operator

$$Q_5(x) = \frac{\alpha_s}{8\pi} \, G^n_{\mu\nu}(x) \, \tilde{G}^n_{\mu\nu}(x) \, . \tag{44}$$

It was found in [20] that the vacuum correlator of topological currents

$$\zeta(q^2) = \mathrm{i} \int \mathrm{d}^4 x \, \exp(\mathrm{i}qx) \Big\langle 0 \,|\, T\{Q_5(x) \, , \, Q_5(0)\} \,|\, \Big\rangle \tag{45}$$

vanishes at $q^2 = 0$ if the theory contains at least one massless quark. Later, it was proved in [21] that in the limit as $N_c \to \infty$, the relation

$$\zeta(0) = \langle 0 \,|\, \bar{q}q \,|\, 0 \rangle \left(\sum_i^{N_f} \frac{1}{m_i} \right)^{-1} \tag{46}$$

holds. In the cases of two and three massless quarks, the validity of Eqn (46) was proved in Ref. [22], where the limit $N_c \to \infty$ was not used. The concept of topological current turned out to be highly effective in QCD: it has been used to establish the spin composition of the proton [23], to establish a relation between the spin structure functions for large and small Q^2 [24, 25], and to determine the axial coupling constants for the nucleon [26].

7. The sum rule for the axial anomaly in quantum chromodynamics

We consider the general representation of the transition amplitude of the axial current into two photons with momenta p and p' in terms of the structure functions (form factors) without kinematic singularities, $T_{\mu\alpha\beta}(p, p')$ [27]. We limit ourselves to the case where $p^2 = p'^2$. Then [28, 29]

$$T_{\mu\alpha\beta}(p, p') = F_1(q^2, p^2) \, q_\mu \varepsilon_{\alpha\beta\rho\sigma} p_\rho p'_\sigma - \frac{1}{2} F_2(q^2, p^2)$$
$$\times \left[\varepsilon_{\mu\alpha\beta\sigma}(p - p')_\sigma - \frac{p_\alpha}{p^2} \, \varepsilon_{\mu\beta\rho\sigma} p_\rho p'_\sigma + \frac{p'_\beta}{p^2} \, \varepsilon_{\mu\alpha\rho\sigma} p_\rho p'_\sigma \right] . \tag{47}$$

The anomaly condition in QCD reduces to

$$F_2(q^2, p^2) + q^2 F_1(q^2, p^2)$$
$$= 2 \sum_q m_q G(q^2, p^2) - \frac{e^2}{2\pi^2} \sum_q e_q^2 N_c \, . \tag{48}$$

Because $T_{\mu\alpha\beta}(p, p')$ is nonsingular at $p^2 = 0$, we have $F_2(q^2, 0) = 0$. The functions $F_1(q^2, p^2)$, $F_2(q^2, p^2)$, and $G(q^2, p^2)$ can be described by dispersion relations in q^2 with no subtractions. Using these relations, we can prove the sum rule

$$\int_{4m^2}^{\infty} \mathrm{Im} \, F_1(t, p^2) \, \mathrm{d}t = \frac{e^2}{2\pi^2} \sum e_q^2 N_c \, , \tag{49}$$

where m^2 is the smallest of quark masses. The sum rule in (49) was proved in [30] for $p^2 < 0$, $m = 0$, in [28] for $p^2 = p'^2$, and in [31] in the general cases where $p^2 \neq p'^2$. We note that (49) also holds for massive quarks. We consider the most interesting case where the axial current is the third component of the isovector current:

$$j^{(3)}_{\mu 5} = \bar{u} \gamma_\mu \gamma_5 u - \bar{d} \gamma_\mu \gamma_5 d \, . \tag{50}$$

We ignore the masses of the u- and d-quarks and assume that $p^2 = p'^2 = 0$. Combining (47) and (48), we obtain

$$T_{\mu\alpha\beta}(p, p') = -\frac{2\alpha}{\pi} N_c \, \frac{q_\mu}{q^2} \, (e_u^2 - e_d^2) \, \varepsilon_{\alpha\beta\lambda\sigma} p_\lambda p'_\sigma \, . \tag{51}$$

It follows from (51) that the transition of the isovector axial current into two photons occurs through an intermediate massless state. Such a state (in the limit m_u, $m_d \to 0$) is the π^0-meson (Fig. 3). Combining the fact that $\langle 0 \,|\, j^{(3)}_{\mu 5} \,|\, \pi^0 \rangle = \sqrt{2} \, \mathrm{i} f_\pi q_\mu$ and the anomaly condition, we can find the matrix element of the $\pi^0 \to 2\gamma$ decay,

$$M(\pi^0 \to 2\gamma) = A \varepsilon_{\alpha\beta\lambda\sigma} \varepsilon_{1\alpha} \varepsilon_{2\beta} p_\lambda p'_\sigma \, , \tag{52}$$

find the constant A, and calculate the width of $\pi^0 \to 2\gamma$ as

$$\Gamma(\pi^0 \to 2\gamma) = \frac{\alpha^2}{32\pi^3} \, \frac{m_\pi^3}{f_\pi^2} \, . \tag{53}$$

This result was first obtained in [32]. Under the assumption that $f_{\pi_0} = f_{\pi^+} = 130.7$ MeV, we obtain $\Gamma(\pi^0 \to 2\gamma)_{\text{theory}} = 7.73$ eV from (53). It is difficult to estimate the accuracy of the prediction, but apparently it varies between 5 and 10%. The experimental value of this quantity averaged over all existing measurements (data for the year 2006) is $\Gamma(\pi^0 \to 2\gamma) = 7.8 \pm 0.6$ eV [33]. To achieve better accuracy for the theoretical prediction, we must (a) insert f_{π^0} instead of f_{π^+} in (53), and (b) allow the contribution of excited states (in addition to π^0) to the sum rule (49) for the isovector current at $p^2 = 0$. This program was implemented in Ref. [34], where it was shown that the difference $\Delta f_\pi = f_{\pi^0} - f_{\pi^+}$ is small: $\Delta f_\pi / f_\pi \approx -1.0 \times 10^{-3}$. Among the excited states, only the η-meson contributes significantly. Its contribution is determined by the value of the $\pi^0 - \eta$ mixing angle [35, 36] and the width $\Gamma(\eta \to 2\gamma) = 510$ eV [33]. It was found in Ref. [34] that $\Gamma(\pi^0 \to 2\gamma)_{\text{theory}} = 7.93 \pm 1.5\%$. The most recent measurements in [37] yield $\Gamma(\pi^0 \to 2\gamma)_{\text{exp}} = 7.93 \pm 2\%$, $\pm 2.1\%$, i.e., the experimental data are in extremely good agreement with the theoretical predictions.

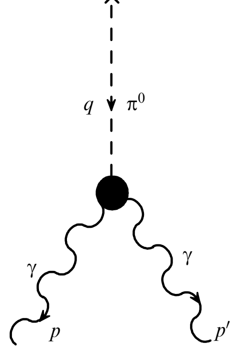

Figure 3. The diagram describing the transition of the isovector axial current (denoted by X) into two photons).

It would seem that Eqn (51) suggests that the existence of a massless (in the limit of massless u- and d-quarks) Goldstone π^0-meson is a consequence of the axial anomaly described by the triangle diagrams in Fig. 1. This is not the case, however. A direct calculation of $\mathrm{Im}\,F_1(q^2, p^2)$ (it is to this function that the intermediate π^0-meson contributes) for $p^2 \neq 0$ shows [9, 28] that in this case, $\mathrm{Im}\,F_1(q^2, p^2)$ is a regular function of q that tends to a constant as $q^2 \to 0$ and has no singularities of the $\delta(q^2)$ type, in contrast to the case $p^2 = 0$ described above. Thus, the amplitude $T_{\mu\alpha\beta}(p, p')$ corresponding to the transition of the axial current to two virtual photons and calculated according to the diagrams in Fig. 1 has no pole in q^2 at $q^2 = 0$. On the other hand, based on a chiral effective theory (e.g., see Ref. [38]), we can state that the transition amplitude of the axial current to two virtual photons must contain the contribution provided by the intermediate massless π^0-meson (see Fig. 3). As shown in [6], the introduction of gluon lines into the diagrams in Fig. 1 does not change the expression for the anomaly. (Actually, this was shown in [6] to be true for QED, but there is no difference between QCD and QED in this aspect.) Thus, from examining the case where $p^2 \neq 0$, we conclude that the appearance of a massless π^0-meson in the dispersion representation of the AVV form factor is not caused by an anomaly. The presence of massless Goldstone mesons (π, K, η) stems from the spontaneous breaking of chiral symmetry in the QCD vacuum. That there is a singularity at $q^2 = 0$ in the amplitude $T_{\mu\alpha\beta}(p, p')$ when $p^2 = 0$ is sometimes interpreted as the double nature of the anomaly, the ultraviolet and the infrared (e.g., see Ref. [3]). I believe that in view of the absence of such a singularity when $p^2 \neq 0$, this interpretation is faulty: the nature of an anomaly in QED and QCD stems from ultraviolet divergences, the singularity in the amplitudes at small distances. (In this respect, QED and QCD differ dramatically from the two-dimensional Schwinger model, in which the origin of an anomaly is truly double (see Ref.[3]).)

For the eighth component of the octet current, the transition amplitude of the axial current to two real photons, $F_1(q^2, 0)$, has a pole at $q^2 = 0$ if $m_{\mathrm{u}} = m_{\mathrm{d}} = m_{\mathrm{s}} = 0$. It is only natural to associate this pole with the η-meson. However, a relation for $\Gamma(\eta \to 2\gamma)$ similar to (53) differs dramatically from the experimental result. A possible explanation of such a discrepancy is the strong nonperturbative interaction of the type of instantons in a pseudoscalar channel mixing η- and η'-mesons [39]. In the case of a singlet axial current, the amplitude $j_{\mu 5}^{(0)} \to 2\gamma$ contains diagrams of the type shown in Fig. 2 (with virtual gluons instead of photons), their extensions, and nonperturbative contributions. Hence, we cannot expect reliable predictions concerning the width of $\eta' \to 2\gamma$ based on anomalies.

’t Hooft hypothesized [40] that the singularities of the amplitudes calculated in QCD on the quark–gluon basis should reproduce themselves in calculations on the hadron basis. Obviously, this is true if both perturbative and nonperturbative interactions are taken into account. However, as a rule, we know nothing about the nonperturbative interactions. In the cases discussed above (except for the decay of π^0 into two real photons), ’t Hooft’s hypothesis does not hold [9].

8. Conclusion

1. An anomaly is an important and necessary element of quantum field theory.
2. An anomaly emerges because the amplitudes of quantum field theory contain ultraviolet singularities, in view of which it is necessary to augment the Lagrangian by renormalization conditions.
3. An anomaly in QCD is related to the appearance of a new quantum number, the winding number.
4. The vacuum in QCD is a linear combination of an infinite number of vacua with different winding numbers.
5. Transitions between vacua with different winding numbers are tunnel transitions occurring along classical paths in the field space, self-dual solutions of QCD equations, or instantons.
6. The axial anomaly in QCD results in the appearance of zero modes in the Dirac equations for light quarks and points to the existence of spontaneous breaking of chiral symmetry in the QCD vacuum, the existence of a quark condensate.
7. The axial anomaly predicts the width of the $\pi^0 \to 2\gamma$ decay with a high accuracy ($\sim 2\%$), a result corroborated by experiments.

References

1. Treiman S B et al. *Current Algebra and Anomalies* (Princeton Series in Physics) (Princeton, NJ: Princeton Univ. Press, 1985)
2. Collins J C *Renormalization* (Cambridge: Cambridge Univ. Press, 1984)
3. Shifman M A *Phys. Rep.* **209** 341 (1991)
4. Peskin M E, Schroeder D V *An Introduction to Quantum Field Theories* (Reading, Mass.: Addison-Wesley, 1995)
5. Adler S L *Phys. Rev.* **177** 2426 (1969)
6. Adler S L, Bardeen W A *Phys. Rev.* **182** 1517 (1969)
7. Bell J, Jackiw R *Nuovo Cimento* **51** 47 (1969)
8. Jackiw R “Field theoretical investigations in current algebra”, in Treiman S B et al. *Current Algebra and Anomalies* (Princeton Series in Physics) (Princeton, NJ: Princeton Univ. Press, 1985) pp. 81–210
9. Ioffe B L *Int. J. Mod. Phys. A* **21** 6249 (2006)
10. Adler S L “Anomalies to all orders”, in *50 Years of Yang-Mills Theory* (Ed. G ’t Hooft) (Singapore: World Scientific, 2005) pp. 187–228; hep-th/0405040
11. Anselm A A, Iogansen A A *Pis’ma Zh. Eksp. Teor. Fiz.* **49** 185 (1989) [*JETP Lett.* **49** 214 (1989)]
12. Weinberg S *Phys. Rev. D* **11** 3583 (1975)
13. Belavin A A et al. *Phys. Lett. B* **59** 85 (1975)
14. Gribov V N, unpublished
15. Jackiw R, Rebbi C *Phys. Rev. Lett.* **37** 172 (1976)
16. Callan C G (Jr), Dashen R F, Gross D J *Phys. Lett. B* **63** 334 (1976)
17. Bitar K M, Chang S-J *Phys. Rev. D* **17** 486 (1978)
18. Altarev I S et al. *Phys. Lett. B* **276** 242 (1992)
19. Atiyah M F, Singer I M *Ann. Math.* **87** 484 (1968); **93** 119 (1971)
20. Crewther R J *Phys. Lett. B* **70** 349 (1977)
21. Di Vecchia P, Veneziano G *Nucl. Phys. B* **171** 253 (1980)
22. Ioffe B L *Yad. Fiz.* **62** 2226 (1999) [*Phys. At. Nucl.* **62** 2052 (1999)]
23. Ioffe B L, Oganesian A G *Phys. Rev. D* **57** R6590 (1998)
24. Burkert V D, Ioffe B L *Phys. Lett. B* **296** 223 (1992)
25. Burkert V D, Ioffe B L *Zh. Eksp. Teor. Fiz.* **105** 1153 (1994) [*JETP* **78** 619 (1994)]
26. Ioffe B L *Survey High Energy Phys.* **14** 89 (1999); hep-ph/9804238
27. Eletsky V L, Ioffe B L, Kogan Ya I *Phys. Lett. B* **122** 423 (1983)
28. Hořejši J *Phys. Rev. D* **32** 1029 (1985)
29. Bass S D et al. *Zh. Russ. Fiz. Obshch* (J. Moscow Phys. Soc.) **1** 317 (1991)
30. Frishman Y et al. *Nucl. Phys. B* **177** 157 (1981)
31. Veretin O L, Teryaev O V *Yad. Fiz.* **58** 2266 (1995) [*Phys. At. Nucl.* **58** 2150 (1995)]
32. Dolgov A D, Zakharov V I *Nucl. Phys. B* **27** 525 (1971)
33. Yao W-M et al. (Particle Data Group) “Review of Particle Physics” *J. Phys. G: Nucl. Part. Phys.* **33** 1 (2006)
34. Ioffe B L, Oganesian A G *Phys. Lett. B* **647** 389 (2007)
35. Ioffe B L *Yad. Fiz.* **29** 1611 (1979) [*Sov. J. Nucl. Phys.* **20** 827 (1979)]
36. Gross D J, Treiman S B, Wilczek F *Phys. Rev. D* **19** 2188 (1979)
37. de Jager K *Prog. Part. Nucl. Phys.* (Available online 28 December 2007); arXiv:0801.4520

38. Ioffe B L *Usp. Fiz. Nauk* **171** 1273 (2001) [*Phys. Usp.* **44** 1211 (2001)]
39. Geshkenbein B V, Ioffe B L *Nucl. Phys. B* **166** 340 (1980)
40. 't Hooft G, in *Recent Developments in Gauge Theories* (NATO Advanced Study Inst. Series, Ser. B, Vol. 59, Eds G 't Hooft et al.) (New York: Plenum Press, 1980) p. 241

PACS number: **03.30.+p**
DOI: 10.1070/PU2008v051n06ABEH006552
DOI: 10.3367/UFNr.0178.200806l.0663

The theory of relativity and the Pythagorean theorem

L B Okun

1. Introduction

The report "Energy and mass in the works of Einstein, Landau and Feynman" that I was preparing for the Session of the Division of Physical Sciences of the Russian Academy of Sciences (DPS RAS) on the occasion of the 100th anniversary of Lev Davidovich Landau's birth was to consist of two parts, one on history and the other on physics. The history part was absorbed into the article "Einstein's formula: $E_0 = mc^2$. 'Isn't the Lord laughing?'" that appeared in the May issue of *Uspekhi Fizicheskikh Nauk* [*Physics-Uspekhi*] journal [1]. The physics part is published in the present article. It is devoted to various, so to speak, technical aspects of the theory, such as the dimensional analysis and fundamental constants c and $\hbar$; the kinematics of a single particle in the entire velocity range from 0 to c; systems of two or more free particles; and the interactions between particles: electromagnetic, gravitational, etc. The text uses the slides of the talk at the session of the Section of Nuclear Physics of the DPS RAS in November 2007 at the Institute for Theoretical and Experimental Physics (ITEP). My goal was to present the main formulas of the theory of relativity in the simplest possible way, using mostly the Pythagorean theorem.

2. Relativity

The advanced standpoint. The history of the concept of mass in physics runs to many centuries and is very interesting, but I leave it aside here. Instead, this will be an attempt to look at mass from an advanced standpoint. I borrowed the words from the famous title of Felix Klein's *Elementary Mathematics from an Advanced Standpoint* (traditionally translated into Russian incorrectly as *Elementary Mathematics from the Standpoint of Higher Mathematics*. See V G Boltyanskii's foreword to the 4th Russian edition). The advanced modern standpoint based on principles of symmetry in general and on the theory of relativity in particular makes it possible to avoid inevitable terminological confusion and paradoxes.

The principle of relativity. Ever since the time of Galileo and Newton, the concept of relativity has been connected with the impossibility of detecting, by means of any experiments, a translational (uniform and rectilinear) motion of closed space (for instance, inside a ship) while remaining within this space. At the turn of XIX and XX centuries Poincaré gave to this idea the name 'the principle of relativity'. [1] In 1905 Einstein generalized this principle to the case of the existence of the limiting velocity of propagation of signals. (The finite velocity of propagation of light has been discovered by Römer already in 1676). Planck called the theory constructed in this way 'Einstein's theory of relativity'.

Mechanics and optics. Newton tried to construct a unified theory uniting the theory of motion of massive objects (mechanics) and the theory of propagation of light (optics). In fact, it became possible to create the unified theory of particles of massive matter and of light only in the XXth century. It was established on the road to this vantage ground of truth that light is also a sort of matter, just like the massive stuff, but that its particles are massless. This interpretation of particles of light — photons — continues to face resistance from many students of physics, and even more from physics teachers.

3. Dimensions

Units in which $c = 1$. The maximum possible velocity is known as the speed of light and is denoted by c. When dealing with formulas of the theory of relativity it is convenient to use a system of units in which c is chosen as a unit of velocity. Since $c/c = 1$, using this system means that we set $c = 1$ in all formulas, thus simplifying them greatly. If time is measured in seconds, then distance in this system of units should be measured in light seconds: one light second equals 3×10^{10} cm.

Poincaré and c. One of the creators of the theory of relativity, Henri Poincaré, when discussing in 1904 the fact that c is found in every equation of electrodynamics, compared the situation with the geocentric theory of Ptolemy's epicycles in which every relation between motions of celestial bodies included the terrestrial year. Poincaré expressed his hope that the future Copernicus would rid electrodynamics of c [3]. However, Einstein showed already in 1905 that c was to play the key role as the limit for the velocity of signal propagation.

Two system of units: SI and $c = 1$. The unit of velocity in the International System of Units SI, 1 m s^{-1}, is forced on us by convenience arguments and by standardization of manufacturing and commerce but not by the laws of Nature. In contrast to this, c as a unit of velocity is imposed by nature itself when we wish to consider fundamental processes of Nature.

Dimensional factors. Consider some physical quantity a. Let us denote by $[a]$ the dimension of the quantity a. The dimension of a definitely changes if it is multiplied by any power of the universal constant c but its physical meaning remains unaffected. In what follows I explain why this is so.

Velocity, momentum, energy, mass. The dimensions of momentum, mass, and velocity of a particle are usually related by the formula $[\mathbf{p}] = [m][\mathbf{v}]$ while the dimensions of energy, mass, and velocity are related by the formula $[E] = [m][\mathbf{v}^2]$.

Let us introduce dimensionless velocity $\mathbf{v}/c$ and from now on denote this ratio as $\mathbf{v}$. Likewise, referring to momentum $\mathbf{p}$ we actually mean the ratio $\mathbf{p}/c$. When speaking of energy, we

[1] This sentence was added by the Author in the English proof.