Anomalies and low-energy theorems of quantum chromodynamics[1)]

M. A. Shifman

Institute of Theoretical and Experimental Physics, Moscow
Usp. Fiz. Nauk **157**, 561–598 (1989)

We discuss the scale and chiral anomalies in quantum chromodynamics and their implications for the theory of hadrons. In the first part the physical meaning of the anomaly is demonstrated. To this end the simplest gauge model is considered—two-dimensional Schwinger model. In this model it is extremely easy to explain what properties of the theory are responsible for the quantum anomaly and the physical nature of the phenomenon is elucidated. The second part is devoted to derivation of the anomaly relations within QCD. The subtle question of the multiloop corrections is discussed. The third part presents applications. Starting from the chiral and scale anomalies we obtain low-energy theorems which lead, in turn, to predictions for observable processes. In particular, a proof based on first principles is given of the existence of massless pions, the amplitudes are calculated for the conversion of gluon operators into pions which can be measured experimentally. Other practical problems are also considered, the solution of which turns out to be possible due to anomalies.

INTRODUCTION

In this review we will discuss quantum anomalies and the consequences stemming from them for the theory of hadrons.

The term "quantum anomalies" in field theory has a concrete narrow meaning. Let us consider an action possessing certain invariance at the classical level. If this invariance can not be preserved at the quantum level, i.e., taking account of the quantum corrections, such a phenomenon is called a "quantum anomaly."

The first encounter of the theorists with quantum anomalies occurred long before the QCD era (the famous $\pi^0\to 2\gamma$ puzzle, see, e.g., the review in Ref. 1). After the triumph of the gauge theories—the Glashow-Weinberg-Salam model in 1972 and QCD in 1973—the status of the problem drastically changed: from a relatively local issue it turned into an important and universal theoretical construction occupying one of the central places in modern theory.

There are two types of anomalies; let us call them internal and external. In the first case the symmetry we deal with is the gauge symmetry—the gauge invariance of the classical action is violated at the quantum level. In other words, when quantum corrections–"loops"–are taken into account, the current with which gauge bosons interact ceases to be conserved.

External anomalies also result in current non-conservation. In this case, however, the anomalous current is not connected with the gauge bosons and corresponds to "external" global symmetries of the classical action.

The presence of internal anomalies is usually considered to be a disaster for the theory. Such theories can not be consistently quantized, they are non-renormalizable and self-contradictory. (Let us note, though, a series of recent articles[2] presenting an attempt to circumvent the problems. We will not deal with these articles since the development here is far from completion.) The standard strategy consists of canceling all internal anomalies by means of a special choice of the fields in the original Lagrangian. For instance, in the Glashow-Weinberg-Salam model the internal triangle anomaly cancels provided that the quarks and leptons from the given generation possess quite definite values of the hypercharge—namely, the standard values—and all the quarks and leptons in the generation take part in the cancellation. The requirement of the absence of internal anomalies imposes contraints on the structure of the gauge models. The constraints can serve (and do serve) as a guideline in model-building. Suffice it to recall that within the superstring approach, the present-day candidate for the "theory of everything," the cancellation of the internal anomalies leaves only two possible options for the gauge group, SU_{32} and $E_8\times E_8$.[3]

We will not dwell on these issues. Both topics—quantization of the theories with internal anomalies and cancellation of the internal anomalies in the course of model-building—call for a special discussion.

In this review we will concentrate on external anomalies. The main purpose is to explain the physical meaning of the phenomenon in the most simple and transparent form and to demonstrate the wide-ranging consequences for the theory of hadrons stemming from the existence of two anomalies in QCD: the chiral anomaly and the scale anomaly. The attention which will be paid to pedagogical aspects and special emphasis on practical applications—both these are points that distinguish the discussion below from numerous excellent recent reviews on anomalies available in the literature. In particular, I would like to mention the review paper of Ref. 4 where technical and mathematical details are fully discussed. One can also find there a detailed list of references.

I begin with the simplest known example: two-dimensional quantum electrodynamics (QED). The example is very instructive and nicely illustrates the answers to such questions as

—What properties of the theory are responsible for the occurrence of the quantum anomalies?

—What is the physical nature of the phenomenon?

The crucial importance of the following two points will be clearly seen: (i) the presence of an infinite number of degrees of freedom in field theory and the necessity of introducing an ultraviolet cut-off; (ii) the clash between two classical symmetries in which only one will turn out to be a "victor" at the quantum level; the one which gains the victory will be preserved at a price of sacrificing the second, less "fortunate" symmetry. It is this latter symmetry that is the quantum anomaly.

The second part of the review is devoted to chiral and

 0038-5670/89/040289-21$01.80

conformal anomalies in QCD. Here I present different derivations of the anomalies and discuss some subtle points which are usually skipped in the literature.

Finally, the third part is devoted to the low-energy theorems following from the anomalous relations and to applications in hadron physics. I will show here how one can use the anomalies for predicting such nontrivial quantities as, for instance, the gluon-ion conversion amplitudes.

1. CHIRAL ANOMALY IN THE SCHWINGER MODEL

1.1. Schwinger model on a circle

Two-dimensional QED with the massless Dirac fermion seems to be the simplest gauge model. The Lagrangian is

$$\mathcal{L}=\frac{1}{4e_0^2}F_{\mu\nu}F_{\mu\nu}+\bar{\psi}i\hat{D}\psi, \tag{1.1}$$

where $F_{\mu\nu}$ is the photon field strength tensor

$$F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\nu, \tag{1.2}$$

e_0 is the gauge coupling constant having the dimension of mass for $D=2$. Moreover, D_μ is the covariant derivative

$$iD_\mu=i\partial_\mu+A_\mu, \tag{1.3}$$

and ψ is the two-component spinor field.

Gamma matrices can be chosen in the following way:

$$\begin{aligned}&\gamma^0=\sigma_2,\ \gamma^1=i\sigma_1,\ \gamma^5=\sigma_3\ \text{(in Minkowski space)},\\ &\gamma^1=\sigma_1,\ \gamma^2=\sigma_2,\gamma^5=\sigma_3\ \ \text{(in Euclidean space)}.\end{aligned} \tag{1.4}$$

Correspondingly, the spinor

$$\psi_L=\begin{pmatrix}\psi_1\\ 0\end{pmatrix}$$

will be called left-handed ($\gamma^5\psi_L=\psi_L$), while the spinor

$$\psi_R=\begin{pmatrix}0\\ \psi_2\end{pmatrix}$$

will be called right-handed ($\gamma^5\psi_R=-\psi_R$).

In spite of the considerable simplification compared to the four-dimensional QED, the dynamics of the model (1.1) is still too complicated for our purposes. Indeed, the set of asymptotic states in this model drastically differs from the fields in the Lagrangian. In the two-dimensional theory the photon, as is well-known, has no transverse degrees of freedom and essentially reduces to the Coulomb interaction. The latter, however, grows linearly with the distance. The linear growth of the Coulomb potential results in the confinement of the charged fermions in the Schwinger model irrespectively of the value of the coupling constant e_0. The model (1.1) was even used as a prototype for describing the color confinement in QCD (see, e.g., Ref. 5).

In order to simplify the situation further let us do the following. Consider the system described by the Lagrangian (1.1) in a finite spatial domain of length L. If e_0L is small, $e_0L\ll 1$, the Coulomb interaction never becomes strong and one can actually treat it as a small perturbation. In particular, in the first approximation its effect can be neglected altogether.

We impose periodic boundary conditions on the field A_μ and antiperiodic ones on ψ. Thus, the problem to be considered below is the Schwinger model on the circle. Notice that the antiperiodic boundary conditions are imposed on the fermion field for convenience only. As will be seen, any other boundary condition—periodic, for instance—would do as well; nothing would be changed except minor technical details.

Thus,

$$\begin{aligned}&A\left(x=-\frac{L}{2},\ t\right)=A\left(x=\frac{L}{2},\ t\right),\\ &\psi\left(x=\frac{L}{2},\ t\right)=-\psi\left(x=-\frac{L}{2},\ t\right).\end{aligned} \tag{1.5}$$

Eq. (1.5) implies that the fields A and ψ can be expanded in the Fourier modes, $\exp[ikx2\pi/L]$ for the bosons and $\exp\{i[k+(1/2)]x\cdot 2\pi/L\}$ for the fermions. Now, let us recall the fact that the Lagrangian (1.1) is invariant under the local gauge transformations:

$$\psi\to\psi e^{i\alpha(x,\,t)},\quad A_\mu\to A_\mu+\partial_\mu\alpha(x,\,t). \tag{1.6}$$

It is quite evident that all modes for the field A_1 except the zero mode (i.e., $k=0$) can be "gauged away." Indeed, the term of the type $a(t)\exp(ikx)$ in A_1 is gauged away with the aid of the gauge function $\alpha(x,t)=-(ik)^{-1}a(t)\exp(ikx)$. The latter is periodic on the circle, as it should be, and does not violate the conditions (1.5).

Thus, in the most general case we can treat A_1 as an x-independent constant.

This is not the end of the story, however, since the possibilities provided by gauge invariance are not exhausted yet. There exists another class of admissible gauge transformations with the gauge function which is not periodic in x,

$$\alpha=\frac{2\pi}{L}nx, \tag{1.7}$$

where n is an integer ($n=\pm 1,\pm 2,\ldots$). In spite of the nonperiodicity, such a choice of the gauge function is also compatible with the conditions (1.5). This fact is readily verifiable: since $\partial\alpha/x=\text{const}$ and $\partial\alpha/\partial t=0$ the periodicity for A_μ is not violated; the analogous assertion is also valid for the phase factor $\exp(i\alpha)$—the difference of phases at the endpoints of the intervale $x\in[-L/2,L/2]$ is equal to $2\pi n$.

As a result, we arrive at the conclusion that the variable A_1 (recall that in the sense of x-dependence A_1 is a constant) should not be considered in the whole interval $(-\infty,\infty)$. The points $A_1,A\pm 2\pi/L,A_1\pm 4\pi/L$, etc. are gauge equivalent and must be identified. The variable A_1 is an independent variable only in the interval $[0,2\pi/L)$; going beyond these limits we find ourselves in the gauge image of the original interval. Following the commonly accepted terminology we may say that A_1 lives on the circle of length $2\pi/L$.

Moreover, the gauge invariance of the theory is closely interrelated with the conservation law for the electric charge. Indeed, the Lagrangian (1.1) with finite L admits multiplication of the fermion field by a constant phase,

$$\psi\to e^{i\alpha}\psi,\quad \psi^+\to\psi^+e^{-i\alpha}.$$

Using the standard line of reasoning one easily derives from this phase invariance the conservation of the electric current

$$j_\mu=\bar{\psi}\gamma_\mu\psi,\quad \dot{Q}(t)=0,$$

where $Q = \int dx\, j_0(x,t)$. (The vanishing of the current divergence stems also directly from the equations of motion.) We note that the classical Lagrangian (1.1) is invariant under one more rotation, the global axial transformation

$$\psi \to e^{i\alpha\gamma^5}\psi, \quad \psi^+ \to \psi^+ e^{-i\alpha\gamma^5},$$

which multiplies the left- and right-handed fermions by the opposite phases ($\gamma^5 = \sigma_3$). If the axial charge of the left-handed fermion $Q_5 = +1$, for the right-handed fermion $Q_5 = -1$. At the classical level the axial current

$$j_{\mu 5} = \bar{\psi}\gamma_\mu\gamma_5\psi$$

is conserved just in the same way as the electromagnetic one. The conservation of Q and Q_5 is equivalent to the conservation of the number of the left-handed and right-handed fermions separately. The fact is quite obvious for any Born graph. Indeed, in all such graphs the fermion lines are continuous, the photon emission does not change their chirality, and the number of ingoing fermion legs is equal to that of the outgoing legs. In the exact answer, however, only the sum of the chiral charges is conserved, only one out of two classical symmetries survives the quantization of the theory.

As will be seen below, the characteristic excitation frequencies for A_1 are of order e while those associated with the fermionic degrees of freedom are of order L^{-1}. Since $eL \ll 1$ the variable A_1 is adiabatic with respect to the fermionic degrees of freedom. In the next section we will analyze in more detail the fermion sector assuming temporarily that A_1 is a fixed (time-independent) quantity.

1.2. Dirac sea. The vacuum wave function

Following the standard prescription of the adiabatic approximation we freeze the photon field A_μ and consider it as "external." In the accepted gauge the A_1 component reduces to a constant. As for A_0 we can put $A_0 \approx 0$ (the fact that A_0 is actually non-vanishing results in negligible corrections).

The difference between these two components lies in the fact that the fluctuations of A_0 are small, while this is not the case for A_1. The wave function is not localized in A_1 in the vicinity of $A_1 \approx 0$. It is just this phenomenon—delocalization of the A_1 wave function and the possibility of penetration to large values of A_1—that will lead to observable manifestations of the chiral anomaly. The $\mu = 0$ component of the photon field is responsible for the Coulomb interaction between the charges; the corresponding effect is of order $eL \ll 1$ and does not show up in the leading approximation to which we will limit ourselves in the present section.

In the two-dimensional electrodynamics the Dirac equation determining the fermion energy levels has the form

$$\left[i\frac{\partial}{\partial t} + \sigma_3\left(i\frac{\partial}{\partial x} - A_1\right)\right]\psi = 0. \qquad (1.8)$$

For the k th stationary state $\psi \sim \exp(-iE_k t)\psi_k(x)$, and the energy of the k th state is

$$E_k\psi_k = -\sigma_3\left(i\frac{\partial}{\partial x} - A_1\right)\psi_k(x). \qquad (1.9)$$

Furthermore, the eigenfunctions are proportional to

$$\psi_k \sim \exp\left[i\left(k+\frac{1}{2}\right)\cdot\frac{2\pi}{L}x\right] \quad (k = 0, \pm 1, \pm 2, \ldots). \qquad (1.10)$$

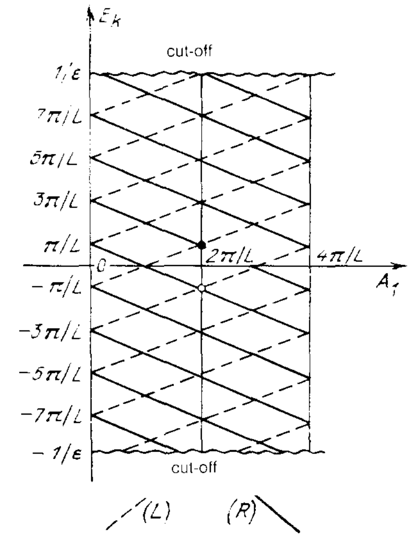

FIG. 1.

The extra term $(1/2)(2\pi/L)x$ in the exponent ensures the antiperiodic boundary conditions, see Eq. (1.5). As a result, we conclude that the energy of the k th level for the left-handed fermions (the definition is given following Eq. (1.4)) is

$$E_{k(L)} = \left(k+\frac{1}{2}\right)\frac{2\pi}{L} + A_1; \qquad (1.11a)$$

while for the right-handed fermions

$$E_{k(R)} = -\left(k+\frac{1}{2}\right)\cdot\frac{2\pi}{L} - A_1. \qquad (1.11b)$$

The energy level structure dependence on A_1 is displayed in Fig. 1. The dashed lines show the behavior of $E_{k(L)}$ and the solid lines $E_{k(R)}$. At $A_1 = 0$ the energy levels for the left-handed and right-handed fermions are degenerate.

If A_1 increases, the degeneracy is lifted and the levels are split. At the point $A_1 = 2\pi/L$ the structure of the energy levels is precisely the same as for $A_1 = 0$; the degeneracy takes place again. The identity of the points $A_1 = 0$ and $A_1 = 2\pi/L$ is the remnant of the gauge invariance of the original theory (see the discussion above).

We note that the identity is achieved in a nontrivial way: in passing from $A_1 = 0$ to $A_1 = 2\pi/L$ a restructuring of the fermion levels takes place. All left-handed levels are shifted upwards by one interval while all right-handed levels are shifted downwards by the same one interval. This phenomenon, the restructuring of the fermion levels, lies at the basis of the chiral anomaly in the model at hand, as will become clear shortly.

Let us proceed now from the one-particle Dirac equation to field theory. The first task is the construction of the ground state, the vacuum. To this end, following the well-known Dirac prescription, we fill up all levels lying in the Dirac sea, leaving all positive-energy levels empty. The following notations will be used below for filled and empty levels with a given k:

$$|1_{L,R}, k\rangle, \quad |0_{L,R}, k\rangle,$$

respectively. The subscript L(R) indicates that we are dealing with left-handed (right-handed) fermions.

Recall that A_1 is a slowly varying adiabatic variable; the corresponding quantum mechanics will be considered later. At first, the value of A_1 is fixed in the vicinity of zero, $A_1 \approx 0$. Then the fermion wave function of the vacuum, as seen from Fig. 1, reduces to

$$\Psi_{\text{ferm}\cdot\text{vac}} = \Big(\prod_{k=-1,-2,\ldots} |1_L, k\rangle\Big)\Big(\prod_{k=0,1,2,\ldots} |0_L, k\rangle\Big) \times \Big(\prod_{k=0,1,2,\ldots} |1_R, k\rangle\Big)\Big(\prod_{k=-1,-2,\ldots} |0_R, k\rangle\Big). \tag{1.12}$$

The Dirac sea, or all negative-energy levels, are completely filled.

Now, let A_1 increase adiabatically from 0 up to $2\pi/L$. The same figure shows that at $A_1 = 2\pi/L$ the wave function (1.12) describes the state which, from the standpoint of the normally filled Dirac sea, contains one left-handed particle and one right-handed hole (small circles on Fig. 1).

Do the quantum numbers of the fermion sector change in the process of the transition from $A_1 = 0$ to $A_1 = 2\pi/L$? Answering this question in the most naive manner we would say that the appearance of the particle and a hole does not change the electric charge since the electric charges of the particle and the hole are obviously opposite. In other words, the electromagnetic current is conserved.

On the other hand, the axial charges of the left-handed particle and the right-handed hole are the same ($Q_5 = 1$) and, hence, in the transition at hand

$$\Delta Q_5 = 2. \tag{1.13}$$

Eq. (1.13) can be rewritten as follows:

$$\Delta Q_5 = \frac{L}{\pi}\,\Delta A_1$$

Dividing by Δt, the transition time, we get

$$\dot{Q}_5 = \frac{L}{\pi}\,\dot{A}_1. \tag{1.14}$$

which implies, in turn, that the conserved quantity has the form

$$\int \Big(j_{05} - \frac{1}{\pi} A_1\Big)\,\mathrm{d}x. \tag{1.15}$$

The current corresponding to the charge (1.15) is, obviously,

$$\begin{gathered}\tilde{j}_{\mu 5} = j_{\mu 5} - \frac{1}{\pi}\varepsilon_{\mu\nu}A_\nu, \quad \partial_\mu \tilde{j}_{\mu 5} = 0,\\ \partial_\mu j_{\mu 5} = \frac{1}{\pi}\varepsilon_{\mu\nu}\partial_\mu A_\nu,\end{gathered} \tag{1.16}$$

where $\varepsilon_{\mu\nu}$ is the antisymmetric tensor, $\varepsilon_{12} = -\varepsilon_{21} = 1$. The second equality in (1.16) represents the famous axial anomaly in the Schwinger model. We succeeded in deriving it by the "hand-waving" arguments, by inspecting the picture of motion of the fermion levels in the external field $A_1(t)$. It turns out that in this language the chiral anomaly presents an extremely simple and widely known phenomenon: the crossing of the zero point in the energy scale by this or that level (or by a group of levels). The presence of the infinite number of levels and the Dirac "multiparticle" interpretation, according to which the emergence of the filled level from the sea means the appearance of the particle while the submergence of the empty level into the sea is equivalent to production of a hole, are the most essential elements of the whole construction. With the finite number of levels, when there is no place for such an interpretation, there can be no quantum anomaly.

I would like also to draw the reader's attention to a somewhat different (although intimately related with the previous) aspect of the picture. The fermion levels move parallel to each other through the bulk of the Dirac sea. Therefore, the disappearance of the levels beyond the zero-energy mark occurs simultaneously with the disappearance of their "copies" beyond the ultraviolet cutoff, which is always implicitly present in the field theory (below we will introduce it explicitly). Because of this fact the heuristic derivation of the anomaly given in this section and a more standard treatment based on the ultraviolet regularization are actually the same. Often it turns out more convenient to trace just the crossing of the ultraviolet cutoff by the levels from the Dirac sea; and beyond the toy models, in the real theories like QCD, the latter approach becomes an absolute necessity, not a question of convenience, due to the notorious "infrared slavery." The connection between the ultraviolet and infrared interpretations of the anomaly is discussed in more detail is subsecs. 1.3, 1.5. The interested reader is referred to the original work[45] where all subtle points are exhaustively analyzed.

1.3. Ultraviolet regularization

In spite of the transparent character of this heuristic derivation almost every one of the "evident" points above can be questioned by the careful reader. Indeed, why is the wave function (1.12) the appropriate choice? In what sense is the energy of this state minimal taking into account the fact that, according to (1.11),

$$E \sim -\sum_{m=0}^{\infty}\Big(m + \frac{1}{2}\Big)\cdot\frac{2\pi}{L},$$

and the sum is ill-defined (the series is divergent)?

Moreover, it is usually asserted that the quantum anomalies are due to the necessity of the ultraviolet regularization of the theory. If so, why speak of the Dirac sea and the crossing of the zero energy point by the fermion levels?

Surprising though it is, all these questions are connected with each other. Probably, it will be most convenient to start with the last one.

Now I will explain that although the ultraviolet regularization was not even mentioned thus far, actually, it is the key element; more than that, the derivation sketched above tacitly assumes quite a specific regularization.

The fermion levels stretch in the energy scale up to indefinitely large energies (more exactly, $|E|$). The wave function (1.12) describing the fermion sector at $A_1 \approx 0$ contains, in particular, the direct product of the infinitely large number of the filled states $|1_R, k\rangle$, $|1_L, k\rangle$ with negative energy. It is clear that such an object—the infinite product—is ill-defined, and one can not do without some kind of regularization in calculating physical quantities. The contribution corresponding to large energies (momenta) should be somehow cut off.

At first sight, it seems that it would be sufficient simply to throw away the terms with $|k| > |k|_{\max}$ ($|k_{\max}$ is a fixed number independent of A_1). This is a regularization, of course, but, clearly enough, the recipe will inevitably lead to a violation of gauge invariance and electric charge non-conservation. Indeed, in gauge theories the momentum p always appears only in the combination $(p+A)$, and any gauge invariant cut off should respect this property. In other words, we have the right to limit $p+A$, not p (or, what is the same, k).

In order to preserve the gauge invariance it is possible and convenient to use the regularization called in the literature the Schwinger, or ε splitting. The regularization will provide a more solid mathematical basis to the heuristic derivation presented above.

Instead of the original currents

$$j_\mu = \bar{\psi}(t,x)\gamma_\mu\psi(t,x),\quad j_{\mu 5} = \bar{\psi}(t,x)\gamma_\mu\gamma_5\psi(t,x) \quad (1.17)$$

we introduce the regularized objects

$$\begin{aligned} j_\mu^{\mathrm{Reg}} &= \bar{\psi}(t,x+\varepsilon)\gamma_\mu\psi(t,x)\exp\left(-i\int_x^{x+\varepsilon} A_1\,\mathrm{d}x\right),\\ j_{\mu 5}^{\mathrm{Reg}} &= \bar{\psi}(t,x+\varepsilon)\gamma_\mu\gamma_5\psi(t,x)\exp\left(-i\int_x^{x+\varepsilon} A_1\,\mathrm{d}x\right). \end{aligned} \quad (1.18)$$

It is implied that $\varepsilon\to 0$ in the final answer for the physical quantities. At the intermediate stages, however, all computations are performed with fixed ε. The exponential factor in (1.18) ensures the gauge invariance of the "split" currents. Without this factor multiplying $\psi(t,x)$ by an x-dependent phase, $\psi(t,x)\to\exp(i^{\alpha}(x))\psi(t,x)$, yields

$$\begin{aligned} &\psi_\alpha^+(t,x+\varepsilon)\psi_\beta(t,x)\to\\ &\to\exp(-i\alpha(x+\varepsilon)+i\alpha(x))\psi_\alpha^+(t,x+\varepsilon)\psi_\beta(t,x). \end{aligned} \quad (1.19)$$

The gauge transformation of A_1 $(A_1\to A_1-\partial\alpha/\partial x)$ compensates for the phase factor in Eq. (1.19).

Now there is no difficulty to calculate the electric and axial charges of the state (1.12) "scientifically." If

$$Q=\int j_0^{\mathrm{Reg}}(t,x)\,\mathrm{d}x,\quad Q_5=\int j_{05}^{\mathrm{Reg}}(t,x)\,\mathrm{d}x, \quad (1.20)$$

then for the vacuum wave function we, evidently, get

$$Q=Q_{\mathrm{L}}+Q_{\mathrm{R}},\quad Q_5=Q_{\mathrm{L}}-Q_{\mathrm{R}}, \quad (1.21)$$

$$Q_{\mathrm{L}}=\sum_k \exp\left\{-i\varepsilon\left[\left(k+\tfrac{1}{2}\right)\cdot\frac{2\pi}{L}+A_1\right]\right\}, \quad (1.22)$$

$$Q_{\mathrm{R}}=\sum_{k'} \exp\left\{-i\varepsilon\left[\left(k'+\tfrac{1}{2}\right)\cdot\frac{2\pi}{L}+A_1\right]\right\}. \quad (1.23)$$

In the limit $\varepsilon\to 0$ both charges, Q_{L} and Q_{R}, turn into the sum of units—each unit represents one filled level from the Dirac sea. Eqs. (1.22) and (1.23) once again demonstrate the gauge invariance of the accepted regularization. Indeed, the cut-off suppresses the states with $|p_1+A_1|\geqslant\varepsilon^{-1}$. If it were not for the phase factor in Eq. (1.18) the suppressing function would not contain the desired combination, $p+A$.

We hasten to add here that although superficially Eqs. (1.22), (1.23) do not differ from each other, actually they do not coincide because the summation runs over different values of k,k'. Just what the particular values are is easy to establish from Fig. 1 (see also Eq. (1.12)).

Let $|A_1|<\pi/L$. Then in the "left-handed" sea the filled levels have $k=-1,-2,\ldots$. In the right-handed" sea the filled levels correspond to $k=0,1,2,\ldots$. Thus, if $|A_1|<\pi/L$ we have

$$\begin{aligned} Q_{\mathrm{L}} &= \sum_{k=-1}^{-\infty}\exp\left\{-i\varepsilon\left[\left(k+\tfrac{1}{2}\right)\cdot\frac{2\pi}{L}+A_1\right]\right\},\\ Q_{\mathrm{R}} &= \sum_{k=0}^{\infty}\exp\left\{-i\varepsilon\left[\left(k+\tfrac{1}{2}\right)\cdot\frac{2\pi}{L}+A_1\right]\right\}. \end{aligned} \quad (1.24)$$

Performing the summation and expanding in ε we arrive at

$$\begin{aligned} (Q_{\mathrm{L}})_{\mathrm{vac}} &= \frac{1}{-i\varepsilon(2\pi/L)}+\frac{L}{2\pi}A_1+O(\varepsilon),\\ (Q_{\mathrm{R}})_{\mathrm{vac}} &= \frac{1}{i\varepsilon(2\pi/L)}-\frac{L}{2\pi}A_1+O(\varepsilon). \end{aligned} \quad (1.25)$$

We pause here and summarize the results. Eqs. (1.25) show that under our choice of the vacuum wave function $\Psi_{\text{ferm.vac.}}$ the charge of the vacuum vanishes, $Q=Q_{\mathrm{L}}+Q_{\mathrm{R}}=0$; there is no time dependence, the charge is conserved. The axial charge consists of two terms: the first term represents an infinitely large constant, and the second one gives a linear A_1 dependence. In the transition $(A_1\approx 0)\to(A_1\approx 2\pi/L)$ the axial charge changes by two units.

These conclusions are not new for us. We have found just the same from the illustrative picture described above in which the electric and axial charges of the Dirac sea are determined intuitively. Now we learned how to sum up the infinite series of units, $\Sigma_k 1$, the charges of the "left-handed" and "right-handed" seas, by virtue of the well-defined procedure which automatically cuts off the levels with $|p_1+A_1|\gtrsim\varepsilon^{-1}$.

The procedure suggests an alternative language for describing the axial charge nonconservation in the transition $(A_1\approx 0)\to(A_1\approx 2\pi/L)$. Previously we thought that the nonconservation is due to the level crossing of the zero-energy point. It is equally correct to say—as we see now—that the nonconservation is explained by the following: one right-handed level from the sea leaves the "fiducial domain" via the lower boundary (the cutoff $-\varepsilon^{-1}$) and one new left-handed level appears in the sea through the same boundary (Fig. 1). Both phenomena, though—the crossing of the zero energy point and the departure (arrival) of the levels via the ultraviolet cut off—occur simultaneously and represent, actually, two different faces of one and the same anomaly, which admits both the infrared and ultraviolet interpretations. The connection between the ultraviolet and infrared aspects of the quantum anomaly lies at the basis of the so called 't Hooft consistency conditions, to be discussed in Sec. 2.4.

One last remark concerning the axial charge. Instead of Eq. (1.18) one could regularize the axial charge in a different way, so that $\partial_\mu j_{\mu 5}=0$ and $\Delta Q_5=0$. (A nice exercise for the reader!) Under such a regularization, however, the expression for the axial current would not be gauge invariant. Specifically, the conserved axial current, apart from Eq. (1.18), includes an extra term $(-1/\pi)\varepsilon_{\mu\nu}A_\nu$ (cf. Eq. (1.16)). As has been already mentioned there is no regularization ensuring simultaneously the gauge invariance and conservation of $j_{\mu 5}$.

1.4. Theta vacuum

Now we leave the issue of charges and proceed to calculation of the fermion-sea energy, the problem which could not be solved at the naive level, without regularization. Fortunately, we already have all the necessary elements.

The fermion part of the Hamiltonian (cf. Eq. (1.8))

$$H = -\psi^+(t, x)\,\sigma_3\left(i\frac{\partial}{\partial x} - A_1\right)\psi(t, x)$$

reduces after the ε splitting to

$$H^{\rm Reg} = -\psi^+(t, x+\varepsilon)\,\sigma_3\left(i\frac{\partial}{\partial x} - A_1\right)\psi(t, x)\,\exp\left(-i\int_x^{x+\varepsilon} A_1\,\mathrm{d}x\right). \tag{1.26}$$

This formula implies, in turn, the following regularized expressions for the energies of the "left-handed" and "right-handed" seas:

$$E_{\rm L} = \sum_{k=-1}^{-\infty} E_{k({\rm L})}\exp(-i\varepsilon E_{k({\rm L})}),$$
$$E_{\rm R} = \sum_{k=0}^{\infty} E_{k({\rm R})}\exp(i\varepsilon E_{k({\rm R})}), \tag{1.27}$$

where the energies of the individual levels $E_{k\,({\rm L,R})}$ are given in Eq. (1.11) and the summation runs over all the negative energy levels. The concrete values of the summation indices in Eq. (1.27) correspond to $|A_1| < \pi/L$. Expressions (1.27) have an absolutely obvious meaning: in the limit $\varepsilon\to 0$ they simply reduce to the sum of the energies of all filled fermion levels from the Dirac sea. The additional exponential factors guarantee the convergence of the sums.

Furthermore we notice that $E_{\rm L}$ and $E_{\rm R}$ can be obtained by differentiating the expressions (1.24) for $Q_{\rm L,R}$ with respect to ε. The latter—the geometrical progression—are trivially summable. Expanding in ε we get

$$E_{\rm sea} = E_{\rm L} + E_{\rm R} = \frac{L}{2\pi}\left(A_1^2 - \frac{\pi^2}{L^2}\right) + O(\varepsilon) + (\text{a constant independent of } A_1). \tag{1.28}$$

Two remarks are in order here. First, it is instructive to check that the Born–Oppenheimer approximation, accepted from the very beginning, is indeed justified. In other words, let us verify that the dynamics of the variable A_1 is slow in the scale characteristic of the fermion sector.

The effective Lagrangian determining the quantum mechanics of A_1 is

$$\mathcal{L} = \frac{L}{2e_0^2}\dot{A}_1^2 - \frac{L}{2\pi}A_1^2. \tag{1.29}$$

This is the ordinary harmonic oscillator with the wave function of the ground state

$$\Psi_0(A_1) = \left(\frac{L}{e_0\pi^{3/2}}\right)^{1/4}\exp\left(-\frac{LA_1^2}{2\pi^{1/2}e_0}\right) \tag{1.30}$$

and the level splitting

$$\omega_{\rm A} = \frac{e_0}{\pi^{1/2}}. \tag{1.31}$$

The characteristic frequencies in the fermion sector are $\omega_{\rm ferm} \sim L^{-1}$. Hence, $\omega_A/\omega_{\rm ferm} \sim Le_0 \ll 1$, q.e.d.

The second remark concerns the structure of the total vacuum wave function. We have convinced ourselves that

$$\Psi_{\rm vac} = \Psi_{\rm ferm\cdot vac}\Psi_0(A_1) \tag{1.32}$$

is the eigenstate of the Hamiltonian of the Schwinger model on the circle in the Born-Oppenheimer approximation. The wave function (1.32) is quite satisfactory from the point of view of the "small" gauge transformations, i.e., those continuously deformable to the trivial (unit) transformation (more exactly, Eq. (1.32) refers to the specific gauge in which the gauge degrees of freedom associated with A_1 are eliminated and A_1 is independent of x). This wave function, however, is not invariant under the "large" gauge transformations $A_1 \to A_1 + (2\pi/L)k$, $k = 0 \pm 1, = 2, \ldots$.

The essence of the situation becomes clear if we return to Fig. 1. When A_1 performs small and slow oscillations in the vicinity of zero, the Dirac sea is filled in such a way as is shown in Eq. (1.12). But A_1 can oscillate as well in the vicinity of the gauge equivalent point $A_1 = 2\pi/L$. In this case, if we do *not* restructure the fermion sector and leave it just as in Eq. (1.12), then the configuration of Eq. (1.12) is obviously *not* the vacuum—it corresponds to one particle plus one hole. The assertion is confirmed, in particular, by the plot showing the Dirac sea energy as a function of A_1 (Fig. 2). In order to get the configuration with the lower energy it is necessary to fill the fermion levels as follows

$$\prod_{k=-2,\,-3,\,\ldots} |1_{\rm L}, k\rangle \prod_{k=-1,\,0,\,1,\,\ldots} |1_{\rm R}, k\rangle$$

(the empty levels are not shown explicitly, cf. Eq. (1.12)).

Thus, the Hilbert space is naturally split into distinct sectors corresponding to different structure of the fermion sea. The wave function of the ground state in the nth sector has the form

$$\Psi_n = \left(\prod_{k=-1-n}^{-\infty} |1_{\rm L}, k\rangle\right)\left(\prod_{k=-n}^{\infty} |1_{\rm R}, k\rangle\right)\Psi_0\left(A_1 - \frac{2\pi}{L}n\right)$$
$$(n = 0, \pm 1, \pm 2, \ldots). \tag{1.33}$$

The organization of the fermion sea is correlated with the position of the "center of oscillation" of A_1. It is quite evident that Ψ_n and $\Psi_{n'}$ are strictly orthogonal to each other due to the fermion factors if $n \neq n'$.

Is it possible to construct a vacuum wave function invariant under the "large" gauge transformations $A_1 \to A_1 + (2\pi/L)k$ (with the simultaneous renumbering of the fermion levels)? The answer is yes. Moreover, such a wave function is not unique. It depends on a new hidden parameter θ which is often called in the literature the vacuum angle. Consider the linear combination

$$\Psi_{\theta\,\rm vac} = \sum_n e^{in\theta}\Psi_n. \tag{1.34}$$

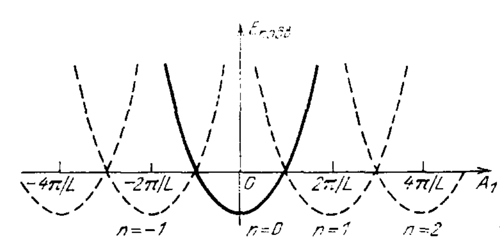

FIG. 2.

This linear combination is also an eigenfunction of the Hamiltonian with the lowest energy, just in the same way as Ψ_n. But unlike Ψ_n the "large" gauge transformations leave $\Psi_{\theta\text{vac}}$ essentially intact. More exactly, under $A_1 \to A_1 + (2\pi/L)$ the wave function (1.34) is multiplied by $\exp(i\theta)$. This overall phase of the wave function is unobservable; all physical quantities resulting from averaging over the θ-vacuum are invariant under the gauge transformations.

Summarizing, we have now become acquainted with one more notion, the vacuum angle θ, the θ-vacuum, which is absolutely transparent in the Schwinger model on the circle and has a direct analog in QCD. Notice that the presence of the vacuum angle θ in the wave function is imitated in the Lagrangian language by adding the so called topological density to the Lagrangian (1.1). In the Schwinger model the topological density is

$$\Delta\mathscr{L}_\theta = \frac{\theta}{2\pi}\varepsilon_{\mu\nu}F_{\mu\nu} = \frac{\theta}{\pi}\varepsilon_{\mu\nu}\partial_\mu A_\nu. \qquad (1.35)$$

The corresponding extra term in the action is the integral of the total derivative: it does not affect the equations of motion and gives a vanishing contribution for any topologically trivial configuration $A(x,t)$. The topological density $\Delta\mathscr{L}_\theta$ shows up only if

$$\int (A_1(x, t = +\infty) - A_1(x, t = -\infty))\,\mathrm{d}x = 2\pi k$$
$$(|k| = 1, 2, \ldots).$$

The topological properties are mentioned here not by chance. It is very instructive to discuss the topological aspect of the theoretical construction under consideration, the more so because in this point as well there is a direct parallel with QCD.

The model at hand possesses the U(1) gauge invariance. An element of the U(1) group, as is well known, can be written as $\exp(i\alpha)$. Using the gauge freedom one can reduce the fields $A_1(x,t)$ or $\psi(x,t)$ at a given moment of time to a standard form by choosing an appropriate gauge function $\alpha(x)$. (For instance, the standard form of A_1 is $A_1 = \text{const.}$)

Moreover, under our boundary conditions the variable x represents the circle of length L and, consequently, we deal here with the (continuous) mappings of the circle into the gauge group U(1). The set of the mappings can be divided in classes. The mathematical formula expressing the fact that the mappings are decomposed into classes is

$$\pi_1(U(1)) = \mathbb{Z}. \qquad (1.36)$$

The meaning of Eq. (1.36) is very simple. Inside each class all mappings, by definition, can be reduced to each other by continuous deformations. On the other hand, no continuous deformations transform mappings from one class into those belonging to another class.

When the mappings of the circle on U(1) are considered, the difference between the classes is especially transparent. Actually, in this case the circle is mapped on another circle because $\exp(i\alpha)$ topologically is nothing other than the circle (Fig. 3). Assume that we started from a certain point, went around the circle a once and returned to the starting point. In doing so we simultaneously went around the circle b. 0, ± 1, $= 2$, etc., times. (Minus corresponds to circulation in the opposite direction.) The number of circuits around the circle b defines the class of the mapping. It is

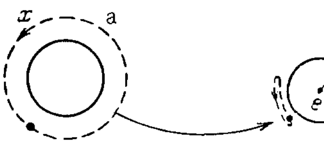

FIG. 3.

quite clear that all mappings with a given number of circuits are continuously deformable into each other. On the contrary, different numbers of circuits guarantee that continuous deformation is impossible. The letter $\mathbb{Z}$ in Eq. (1.36) denotes the set of integers and shows that the set of different mapping classes is isomorphic to the set of integers; each class is characterized by an integer having the meaning of the number of circuits. The mappings corresponding to zero circuit number are called topologically trivial, the others are topologically nontrivial.

This information is sufficient to establish the existence of the vacuum sectors labeled by n ($n = 0, \pm 1, \pm 2, \ldots$) for which $(A_\mu)_{\text{vac}} \propto \partial_\mu \alpha_{(n)}$, without any explicit construction like (1.33); ($\alpha_{(n)}$ belongs to the nth class). The necessity of introducing the vacuum angle θ also stems from this information.

Finally, the last issue to be discussed in connection with the Schwinger model. Sometimes the question is raised as to why the vacuum wave function can not be chosen in the form (1.33) with fixed n. The gauge invariance under the "small" (topologically trivial) transformations is preserved which automatically implies electric charge conservation. What is lost is only the invariance under the "large" (topologically nontrivial) transformations; it seems that there is nothing bad in that. Then, why is it absolutely necessary to pass to $\Psi_{\theta\text{vac}} = \Sigma_n e^{in\theta}\Psi_n$?

The point is that Ψ_n taken as the vacuum wave function violates clusterization—one of the basic properties in field theory which can be traced back to causality and unitarity of the theory. The following is understood by clusterization: the vacuum expectation value of the T product of several local operators must be reducible to the sum over intermediate states including the vacuum intermediate state plus excitations over the *given vacuum*. Violation of clusterization can be demonstrated explicitly. Consider the two-point function

$$\mathscr{A}(t) = \langle \Psi_n | T\{\mathscr{O}^+(t)\,\mathscr{O}(0)\} | \Psi_n \rangle, \qquad (1.37)$$

where

$$\mathscr{O}(t) = \int \bar{\psi}(x, t)(1 + \gamma_5)\psi(x, t)\,\mathrm{d}x.$$

The operator $\mathscr{O}$ changes the axial charge of the state by two units (adds a particle and a hole to the Dirac sea), $\mathscr{O}^+$ returns it back, and, as a result, $\mathscr{A}(t) \neq 0$. Moreover, if $t \to \infty$ in the euclidean domain $\mathscr{A}(t) \to \text{const.}$ (For a concrete calculation see, e.g., Ref. 6 based on the bozonization method. In this work the limit $L \to \infty$ is considered but all relevant expressions can be readily rewritten for finite L.) The fact that $\mathscr{A}(t)$ tends to a nonvanishing constant at $t \to \infty$ means, according to clusterization, that the operators $\bar{\psi}(1 \pm \gamma^5)\psi$ acquire a nonvanishing vacuum expectation value.

On the other hand, if $|\text{vac}\rangle = |\Psi_n\rangle$ then

$\langle\bar{\psi}(1\pm\gamma^5)\psi\rangle = 0$ for a trivial reason. Indeed, the operator $\bar{\psi}(1\pm\gamma^5)\psi$ acting on Ψ_n produces an electron and a hole, and the corresponding state is, obviously, orthogonal to Ψ_n^*.

The clusterization property is restored if from Ψ_n one passes to the θ-vacuum (1.34). In this case there emerges the non-zero nondiagonal expectation value

$$\langle\Psi^*_{n\pm1}\mid\bar{\psi}(1\pm\gamma_5)\psi\mid\Psi_n\rangle \backsim L^{-1}e^{-\pi^{3/2}/Le_0}e^{i\theta}. \qquad (1.38)$$

If the line of reasoning appealing to the necessity of clusterization will seem too academic to the reader, it might be instructive to consider another argument, somewhat connected, by the way, with Eq. (1.37) and the subsequent discussion). Let us ask the question: what will happen if instead of the massless Schwinger model we will consider the model with a small mass, i.e., introduce an extra mass term $\Delta\mathscr{L}_m = -m\bar{\psi}\psi$ in the Lagrangian (1.1)? Naturally, all physical quantities obtained in the massless model will be shifted. It is equally natural to require, however, the shifts to be small for small m, so that there would be no change in the limit $m\to0$. Otherwise, we would encounter an unstable situation while we would like to have the mass term as a small perturbation.

But in the presence of degenerate states (and the states Ψ_n with different n are degenerate) any perturbation is potentially dangerous and can lead to large effects. Just such a disaster occurs, in particular, if $\Delta\mathscr{L}_m$, acting on the vacuum, yields "another vacuum." In other words, if the operator $\Delta\mathscr{L}_m$ is nondiagonal.

If we prescribe the states like Ψ_n to be the vacuum then $\Delta\mathscr{L}_m$ will by no means be diagonal as it follows from the discussion after Eq. (1.37). This we cannot accept. On the other hand, the mass term is certainly diagonalized in the basis of the wave functions (1.34),

$$\langle\Psi_{\theta'\mathrm{vac}}\mid\Delta\mathscr{L}_m\mid\Psi_{\theta\mathrm{vac}}\rangle = 0.$$

1.5. Two faces of the anomaly

In conclusion, let us discuss the connection between the picture presented above and the more standard derivation of the chiral anomaly in the Schwinger model. We have already emphasized the double nature of the anomaly which shows up as the infrared effect in the current and the ultraviolet effect in the divergence of the current. The line of reasoning accepted thus far puts more emphasis on the infrared aspect of the problem—the finite "box" served as a natural infrared regularization. The same result for $\partial_\mu j_{\mu5}$ as in Eq. (1.16) could be obtained with no reference to the infrared regularization.

A conventional treatment of the issue deals directly with $\partial_\mu j_{\mu5}$. Then we need to bother only about the ultraviolet regularization, and, in particular, the theory can be considered in the infinite space since the finiteness of L does not affect the result coming from the short distances.

The method of the ultraviolet regularization commonly used is due to Pauli and Villars. In the model at hand it reduces to the following. In addition to the original massless fermions in the Lagrangian, heavy regulator fermions are introduced with the mass M_0 ($M_0\to\infty$) and the opposite metric. The latter means that each loop of the regulator fermions is supplied by an extra minus sign relatively to the normal fermion loop. The interaction of the regulator fermions with the photons is assumed to be just the same as for the original fermions, and the only difference is the mass. Then the role of the Pauli-Villars fermions in the low-energy processes ($E\ll M_0$) is the introduction of the ultraviolet cutoff in the formally divergent integrals corresponding to the fermion loops. Such a regularization procedure, clearly, automatically guarantees the gauge invariance and the electromagnetic current conservation.

In the model regularized according to Pauli and Villars the axial current has the form

$$j_{\mu5} = \bar{\psi}\gamma_\mu\gamma_5\psi + \bar{R}\gamma_\mu\gamma_5R, \qquad (1.39)$$

where R is the fermion regulator. In calculating the current divergence one can now use the naive equations of motion. Then

$$\partial_\mu j_{\mu5} = 2iM_0\bar{R}\gamma_5R.$$

The divergence is non-vanishing (the axial current is not conserved!), but, as was expected, $\partial_\mu j_{\mu5}$ contains only the regulator anomalous term.

The last step is contracting the regulator fields in the loop in order to convert $M_0\to\bar{R}\gamma_5R$ into the "normal" light fields in limit $M_0\to\infty$. The relevant diagrams are displayed in Fig. 4 where the solid line denotes the standard heavy fermion propagator $(p\gamma - M_0)^{-1}$. The graph 4*a* does not depend on the external field. The corresponding contribution to $\partial_\mu j_{\mu5}$ represents a number which can be set equal to zero. The graph 4*c*, with two photon legs, and all others having more legs die off in the limit $M_0\to\infty$. The only surviving graph is that of Fig. 4b. Calculation of this diagram is trivial, and yields

$$2iM_0\bar{R}\gamma_5R \to \frac{1}{\pi}\varepsilon_{\mu\nu}\partial_\mu A_\nu.$$

As a result, we reproduce the anomalous relation $\partial_\mu j_{\mu5} = (1/\pi)\varepsilon_{\mu\nu}\partial_\mu A_\nu$ obtained previously by a different method.

2. ANOMALIES IN QUANTUM CHROMODYNAMICS

2.1. Classical symmetries

Before proceeding to quantum anomalies in QCD let us first list the symmetries of the classical action. We will assume that the theory contains n_f massless quarks and temporarily forget about heavy quarks which are not essential in the given context. In the actual situation $n_f = 2$ ($q =$ u,d) or $n_f = 3$($q =$ u,d,s). The corrections due to the small u-, d-, s-quark masses can (and will) be considered separately where necessary.

In the limit $m_q = 0$—in the literature it is often referred to as the chiral limit—the classical QCD action is

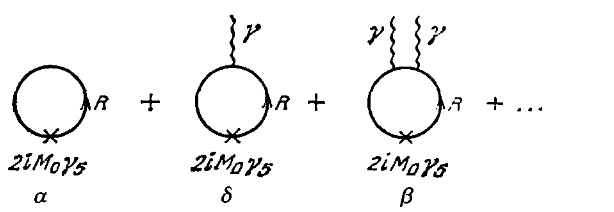

FIG. 4.

$$S=\int \mathcal{L}(x)\,d^4x,\quad \mathcal{L}=-\frac{1}{4}G^a_{\mu\nu}G^a_{\mu\nu}+\sum_q(\bar{q}_L i\hat{D}q_L+\bar{q}_R i\hat{D}q_R),$$
$$q_{L,R}=\frac{1}{2}(1\pm\gamma_5)q \tag{2.1}$$

where q is the quark field and $G^a_{\mu\nu}$ is the gluon field strength tensor. The action S is invariant under the following global transformations:

(i) Rotations of quarks of different flavors,

$$q\to Uq,\quad q=\begin{pmatrix}u\\d\\s\end{pmatrix},$$

where U is a 3 by 3 matrix, an arbitrary element of the SU(3) group. Since the left-handed and right-handed quarks enter the Lagrangian (2.1) as separate terms independent chiral transformations are allowed,

$$q_L\to Uq_L,\quad q_R\to U'q_R,$$

where U and U' are generally speaking different SU(3) matrices. Thus, the QCD Lagrangian actually possesses $SU(3)_L\times SU(3)_R$ symmetry called the chiral flavor invariance.

Sometimes it turns out to be more convenient to form linear combinations from the generators of the $SU(3)_L\times SU(3)_R$ so as to pass from the chiral rotations to the vector and axial transformations. In other words it is equally correct to say that the symmetry group of the classical action (2.1) is $SU(3)_V\times SU(3)_A$.

(ii) The phase $U(1)$ transformations of two types:

$$q_L\to q_L e^{i\alpha},\quad q_R\to q_R e^{i\alpha} \tag{2.2}$$

$$q_L\to q_L e^{i\beta},\quad q_R\to q_R e^{-i\beta}. \tag{2.3}$$

The physical meaning of Eqs. (2.2) and (2.3) is obvious: any Born diagram conserves the number of the left-handed and right-handed fermions separately.

(iii) The scale transformations, i.e., dilatations of the coordinates with the simultaneous rescaling of the quark and gluon fields in accordance with their normal dimension,

$$A_\mu(x)\to\lambda A_\mu(\lambda x),\quad q(x)\to\lambda^{3/2}q(\lambda x). \tag{2.4}$$

The scale invariance obviously stems from the fact that the classical action (2.1) contains no dimensionful constants. This invariance is a part of a larger conformal group. Unfortunately, discussion of the conformal symmetry is beyond the scope of the present review.

At the quantum level the fate of the above symmetries is different. The currents generating the $SU(3)_V\times SU(3)_A$ transformations are conserved even with the ultraviolet regularization switched on. They are anomaly-free.[2)] The vector SU(3) symmetry corresponding to the current

$$j^a_{\mu V}=j^a_{\mu L}+j^a_{\mu R}\quad (a=1,\ \ldots,\ 8),$$

is realized linearly. As far as the axial SU(3) subgroup

$$j^a_{\mu A}=j^a_{\mu L}-j^a_{\mu R},$$

is concerned here the spontaneous breaking of the symmetry takes place, and the given SU(3) subgroup is realized nonlinearly, with participation of the octet of the massless Goldstone bosons (π,η,K). Further discussion of the $SU(3)_A$ and its spontaneous breaking is given in Sec. 2.4.

Moreover, the vector U(1) invariance $q\to qe^{i\alpha}$ also stays a valid anomaly-free symmetry at the quantum level. This symmetry is responsible for the fact that the quark number is constant in any QCD process.

Finally, the current generating the axial U(1) transformations and the dilatation current are not conserved at the quantum level due to the anomalies. Below we dwell on this issue.

2.2. The axial and scale (dilatation) anomalies

As was explained above, the concrete form of the anomalous relations can be established without going beyond perturbation theory provided an appropriate ultraviolet regularization is chosen. We begin this subsection by a warning referring to a subtle point which causes a lot of confusion in the current literature. Two languages are used for description of the same anomalous relations, and many authors do not even realize the distinction between the languages.

Within the first approach one establishes an operator relation, say, between the divergence of the axial current and $G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}$ (see Eq. (2.9)). Both the axial current $j_{\mu A}$ and the product

$$G_{\mu\nu}\tilde{G}_{\mu\nu}\equiv\frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}G^a_{\mu\nu}G^a_{\rho\sigma}$$

are treated within this procedure as Heisenberg operators of the quantum field theory: $j_{\mu A}$ as the quark operator, $G_{\mu\nu},\tilde{G}_{\mu\nu}$ as the gluon one. In order to convert the operator relations into the observable amplitudes it is necessary to make one more step: calculate, according to the general rules, the matrix elements of the operators figuring in the right-hand and left-hand sides of the anomalous equality. (The words "observable amplitudes" are used here metaphorically. In the present section we will speak only about quark and gluon scattering amplitudes while in reality, of course, only hadronic amplitudes are observable. The same remark refers to the expression "matrix elements.")

Within the second approach one operates directly with the matrix elements. More exactly, usually one fixes an external (or background) gluonic field and determines $\partial_\mu j_{\mu A}$ in this field in some way. In the absence of an external field $\partial_\mu j_{\mu A}=0$. The fact of existence of the anomaly implies that $\partial_\mu j_{\mu A}\neq 0$ and $\partial_\mu j_{\mu A}(x)$ is locally expressible in terms of the external field at the same point, $G_{\mu\nu}(x)$. The analysis of the anomaly in the Schwinger model (Sec. 1) has been undertaken just in their vein.

Although the same letters are used in both cases—perhaps, the confusion is due to this tradition—it is quite evident that the Heisenberg operator at the point x and the expression for the background field at the same point are by no means identical objects. Certainly, in the leading order

$$\langle G_{\mu\nu}\tilde{G}_{\mu\nu}\rangle=(G_{\mu\nu}\tilde{G}_{\mu\nu})_{\rm ext},\quad \langle G_{\mu\nu}G_{\mu\nu}\rangle=(G_{\mu\nu}G_{\mu\nu})_{\rm ext}, \tag{2.5}$$

where $\langle\ldots\rangle$ denotes in the case at hand averaging over the external gluon field while the subscript "ext" marks the external field. In the next-to-leading order, however, the right-hand side of Eq. (2.5) acquires α_s corrections, generally speaking. Therefore if the anomalies are discussed beyond the leading order it is absolutely necessary to specify what particular relations are considered: the operator relations or those for the matrix elements. Only in the one-loop approximation both versions superficially coincide.

In order to derive the low-energy theorems applicable to physical hadronic processes (see Sec. 3) we will need the quantum anomalies in operator form. In the remainder of the paper the term "anomaly" will mean the operator anomalous relation.

Let us proceed now to calculations and begin with the axial anomaly since it is simpler in the technical sense and a close example has been already analyzed in the Schwinger model.

The current generating the axial U(1) transformation is

$$j_{\mu A}=\sum_{q=u,\,d,\,s}\bar{q}\gamma_\mu\gamma_5 q. \tag{2.6}$$

Differentiating naively and invoking the equation of motion $\hat{D}q=0$ we get

$$\partial_\mu j_{\mu A}=\sum_q(\bar{q}\hat{D}\gamma_5 q-\bar{q}\gamma_5\hat{D}q)=0,$$

The lesson obtained in the Schwinger model teaches us, however, that this is not the whole story and conservation of the axial current will evaporate after switching on the ultraviolet regularization. In the Schwinger model on the circle ($eL\ll 1$) we deal with the weak coupling regime and, therefore, can choose any of the alternative lines of reasoning based either on the infrared or on the ultraviolet approaches. In QCD it is rather meaningless to speak about the infrared behavior of quarks. To make the calculation of the anomaly reliable we must avoid infrared formulations and invoke only the Green's functions at short distances. As a result, we are forced to shift the emphasis from the analysis of the current $j_{\mu A}$ and concentrate directly on $\partial_\mu j_{\mu A}$.

Using one of the variants of the ultraviolet regularization, the well-tried ε-splitting, we write

$$j^{R}_{\mu A}=\sum_q\bar{q}(x+\varepsilon)\gamma_\mu\gamma_5\left(\exp\int_{x-\varepsilon}^{x+\varepsilon}igA_\rho(y)\,dy\right)q(x-\varepsilon). \tag{2.7}$$

Then the problem reduces to evaluating the quark loop in the background gluon field. Using the equations of motion and expanding the exponential in the braces up to terms of the first order in ε, we get

$$\partial_\mu j^{R}_{\mu A}=\sum_q\{\bar{q}(x+\varepsilon)[-ig\hat{A}(x+\varepsilon)\gamma_5-\gamma_5 ig\hat{A}(x-\varepsilon)+ig\gamma_\mu\gamma_5\varepsilon_\beta G_{\mu\beta}(0)]\,q(x-\varepsilon)\}.$$

The third term in the square brackets contains the gluon field strength tensor and results from differentiation of the exponential factor. For convenience we have imposed the Fock-Schwinger gauge condition $y_\mu A^a_\mu(y)=0$ on the external field (see the review paper of Ref. 7). (This gauge condition is not obligatory, of course.) In this gauge $A_\mu(y)=1/2\,y_\rho G_{\rho\mu}(0)+\dots$. As usual, $A_\mu=1/2t^aA^a_\mu$, where t^a are the color Gell-Mann matrices.

Contracting the quark lines in the loop we arrive at the following expression

$$\begin{aligned}\partial_\mu j^{R}_{\mu A}&=-ign_f\,TR_{\mathrm{color+Lorentz}}\,(-2i\varepsilon_\rho G_{\rho\mu}(0)\gamma_\mu\gamma_5 S(x-\varepsilon,x+\varepsilon))\\&=-n_f\frac{g^2}{2}G^a_{\rho\mu}(0)\tilde{G}^a_{\alpha\varphi}(0)\frac{\varepsilon_\rho\varepsilon_\alpha}{\varepsilon^2}\frac{1}{8\pi^2}TR_{\mathrm{Lorentz}}\gamma_\mu\gamma_5\gamma_\varphi\gamma_5\\&=\frac{n_f\alpha_s}{4\pi}(G^a_{\alpha\beta}\tilde{G}^a_{\alpha\beta})_{\mathrm{ext}};\end{aligned} \tag{2.8}$$

Here $S(y,x)$ is the massless quark propagator in the background field. The expression for this propagator as a series in the background field in the Fock-Schwinger gauge in known (Ref. 7). The terms vanishing in the limit $\varepsilon\to 0$ are neglected in Eq. (2.8). Notice that the quarks propagate only at very short distances $\sim\varepsilon\to 0$.

Eq. (2.8) gives the answer for the axial anomaly in the one-loop approximation

$$\begin{aligned}\partial_\mu j_{\mu A}&=\frac{n_f\alpha_s}{4\pi}G^a_{\alpha\beta}\tilde{G}^{a\alpha\beta},\\ \tilde{G}^{\alpha\beta}&=\frac{1}{2}\varepsilon^{\alpha\beta\rho\sigma}G_{\rho\sigma}.\end{aligned} \tag{2.9}$$

This result is readily reproducible within the Pauli-Villars regularization (an exercise for the reader). Higher orders in α_s will be discussed in Sec. 2.3.

We now proceed to the dilatation anomaly. First of all it is instructive to check that the scale transformation is generated by the current

$$j_{\nu D}=x^\mu\theta_{\mu\nu}, \tag{2.10}$$

where $\theta_{\mu\nu}$ is the QCD energy-momentum tensor (symmetric and conserved):

$$\begin{aligned}\theta_{\mu\nu}=&-G^a_{\mu\alpha}G^a_{\nu\alpha}\\&+\frac{1}{4}g_{\mu\nu}G^a_{\alpha\beta}G^a_{\alpha\beta}+\frac{i}{4}\sum_q[\bar{q}(\gamma_\mu D_\nu+\gamma_\nu D_\mu)q\\&-\bar{q}(\gamma_\mu\overleftarrow{D}_\nu+\gamma_\nu\overleftarrow{D}_\mu)q].\end{aligned} \tag{2.11}$$

The corresponding charge can be represented as

$$D\equiv\int j_{0D}\,d^3x=tH+\tilde{D},\quad \tilde{D}=\int x^i\theta_{i0}\,d^3x, \tag{2.12}$$

where H is the Hamiltonian and

$$[\tilde{D},p_i]=ip_i,$$
$$[\tilde{D},\exp ip_ix^i]=-(x^ip_i)\exp ip_ix^i\qquad(i=1,2,3).$$

Let $\mathcal{O}(x)$ be an arbitrary colorless operator. Then

$$\begin{aligned}[D,\mathcal{O}(x)]&=t[H,\mathcal{O}(x)]+[\tilde{D},\mathcal{O}(x)]\\&=-it\frac{\partial}{\partial t}\mathcal{O}(x)+[\tilde{D},\mathcal{O}(x)].\end{aligned} \tag{2.13}$$

Moreover, the commutator $[\tilde{D},\mathcal{O}(x)]$ contains two terms, the first one corresponding to rescaling of x_i and the second one to rescaling of the operator $\mathcal{O}$. Indeed,

$$\begin{aligned}[\tilde{D},\mathcal{O}(t,x^i)]&=[\tilde{D},e^{ip_ix^i}\mathcal{O}(t,0)e^{-ip_ix^i}]\\&=[\tilde{D},e^{ip_ix^i}]\mathcal{O}(t,0)e^{-ip_ix^i}+e^{ip_ix^i}\mathcal{O}(t,0)[\tilde{D},e^{-ip_ix^i}]\\&\quad+e^{ip_ix^i}[\tilde{D},\mathcal{O}(t,0)]e^{-ip_ix^i}=-ix^i\partial_i\mathcal{O}(t,x^i)\\&\quad+e^{ip_ix^i}[\tilde{D},\mathcal{O}(t,0)]e^{-ip_ix^i}.\end{aligned} \tag{2.14}$$

It is easy to understand that the commutator $[\tilde{D},\mathcal{O}(t,0)]$ should be proportional to $\mathcal{O}(t,0)$ and, hence,

$$[D,\mathcal{O}(x)]=-i\left(x^\mu\frac{\partial}{\partial x^\mu}\mathcal{O}(x)+d\mathcal{O}(x)\right), \tag{2.15}$$

where d is a dimensionless number determined by a particular form of the operator $\mathcal{O}$.

Only the general properties of the quantum field theory have been used thus far. The numerical value of the coefficient d depends on the specific structure of the theory. In

QCD d is equal to the normal dimension of the operator $\mathscr{O}$. (For instance, if $\mathscr{O}(x) = \mathscr{L}(x)$ where $\mathscr{L}$ is the Lagrangian, $d = 4$).

Eq. (2.15) expresses in a mathematical language the change of the units of mass and length. This explains the origin of the name, the scale transformations.

Notice that the reservation about the colorless structure of the operator $\mathscr{O}$ is not superfluous. For the colored operators, say, q or A^a_μ, the commutator with D, as is seen from the direct computation, does not reduce to the form (2.15). The reader should not be puzzled by this fact. The complication is due to gauge fixing, and additional terms can be eliminated by an appropriate gauge transformations.

Using the classical equations of motion we get

$$\partial^\nu j_{\nu D} = \theta^\mu_\mu = 0 \tag{2.16}$$

(cf. Eq. (2.11)). Classically the trace of the energy-momentum tensor vanishes provided that $m_q = 0$. At the quantum level $\theta_{\mu\nu}$ is no more traceless, and the trace θ^μ_μ is given by the dilatation (scale) anomaly.

The fact that the scale invariance (2.4) is lost in loops is obvious. Indeed, the invariance (2.4) takes place because there are no dimensional parameters in the classical action (2.1). Already in one loop, however, such a parameter inevitably appears in the effective action, the ultraviolet cut-off M_0.

The M_0 dependence of the effective action is known beforehand. It is very convenient to use this information, seemingly, the shortest way to calculate θ^μ_μ. In the one-loop approximation the effective action can be written somewhat symbolically as

$$S_{\text{eff}} = -\frac{1}{4}\int\left(\frac{1}{g_0^2} - \frac{b}{16\pi^2}\ln M_0^2 x^2_{\text{char}}\right)(G^a_{\mu\nu}G^a_{\mu\nu})_{\text{ext}}\, d^4x + \ldots, \tag{2.17}$$

where $b = (11N_c/3) - (2N_f/3)$ is the first coefficient in the Gell-Mann-Low function and the dots stand for the fermion terms. We temporarily rescaled the gluon field, $gA \to A$, so that the coupling constant figures only as an overall factor in front of G^2.

The variation of the effective action under the transformation (2.4) is

$$\delta S_{\text{eff}} = -\left(\frac{1}{32\pi^2}\, bG^2_{\mu\nu}\right)\ln\lambda, \tag{2.18}$$

implying that

$$\partial^\mu j_{\mu D} = -\frac{1}{32\pi^2}\, bG^2. \tag{2.19}$$

Returning to the standard normalization of the gluon field we finally get

$$\theta^\mu_\mu = -\frac{b\alpha_s}{8\pi}\, G^a_{\mu\nu}G^a_{\mu\nu} + O(\alpha_s^2). \tag{2.20}$$

2.3. Multiloop corrections

Although in practical applications it is quite sufficient to limit oneself to the one-loop expressions for the anomalies (2.9), (2.20) the question of the higher-order corrections, even though it is academic, still deserves a brief discussion.

Till recently it was generally believed that the question was totally solved. Namely, the axial anomaly, by the Adler-Bardeen theorem,[8] is purely one-loop (i.e., there are no corrections to Eq. (2.9)) while the scale anomaly contains the complete QCD β function in the right-hand side,

$$\beta(\alpha_s(\mu)) = \frac{d\alpha_s(\mu)}{d\ln\mu}.$$

In other words, $(-b\alpha_s/8\pi)$ in Eq. (2.20) is substituted by $\beta(\alpha_s)/4\alpha_s$ if higher order corrections are taken into account.[9]

Later it became clear, however, that the situation is far from being so simple and calls for additional study.[10] Surprising though it is the flaw in the standard arguments has been first revealed not within QCD but in a more complex model, supersymmetric Yang-Mills gauge theory. The minimal model of such a type includes gluons and gluinos, Majorana fermions in the adjoint representation of the color group. The axial current $j_{\mu A}$ (the so-called R current) and the energy-momentum tensor $\theta_{\mu\nu}$ in supersymmetric theories appear in one supermultiplet[11] and, consequently, the coefficients in the chiral and dilatation anomalies can not be different–one-loop for $\partial^\mu j_{\mu A}$ and multiloop for θ^μ_μ.

Being unable to discuss here the multiloop corrections in detail I note only that the generally accepted treatment is based on the confusion mentioned in the beginning of Subsec. 2.2. The standard derivation of the Adler-Bardeen theorem seems to be valid only provided that the axial anomaly is interpreted as an operator equality. At the same time the relation

$$\theta^\mu_\mu = \frac{\beta(\alpha_s)}{4\alpha_s}\,(G^a_\mu G^a_{\mu\nu})_{\text{ext}} \tag{2.21}$$

holds only for the matrix elements (cf. the derivation presented at the end of Subsec. 2.2).

It is quite natural to try to reduce both anomalies to a unified form, preferably to the operator form. The $\partial^\mu j_{\mu A}$ is completely specified by the one-loop approximation (2.9), at least, within a certain ultraviolet regularization.[12] As far as θ^μ_μ is concerned, in this case at the moment we, strictly speaking, do not know even the two-loop coefficient in front of the operator $G^2_{\mu\nu}$, to say nothing about higher order corrections. Calculations existing in the literature should be reanalyzed anew in order to separate the genuinely ultraviolet contributions from those containing an infrared part; the latter should be interpreted as a matrix element.

The remainder of the subsection elucidates this assertion and in principle can be omitted by the reader with no detriment to the subsequent material.

Calculating the effective action is beyond any doubt the most convenient method for determination of the scale anomaly. The term "effective action" is applied to two somewhat different objects. The first object is the sum of all vacuum loops in the given external field. The functional Γ obtained in this way and depending on the external field is often called the effective action, although another name would be more exact: the generator of the one-particle irreducible (1PI) vertices. The other object is the effective action a la Wilson,[13] $S_W(\mu)$, which differs from Γ in the respect that S_W takes into account only the contribution of the loop momenta $p > \mu$ and excludes the infrared domain $p \lesssim \mu$. While Γ represents a c-number functional of the external field, S_W is the quantum-field operator; it can be considered as the original action with respect to the low-frequency fluctuations with the wavelengths $s\mu$. The functional Γ results from taking the matrix element of $\exp(iS_W(\mu))$.

The standard line of reasoning which leads to Eq. (2.21) is based on the analysis of Γ. Just this functional defines the effective coupling constant $g^2(\mu)$,

$$\Gamma = -\frac{1}{4}\frac{1}{g^2(\mu)}(G_{\mu\nu}^2)_{\rm ext} + \dots,$$

and differentiating $g^2(\mu)$ with respect to $\ln \mu$ yields the complete β function. If, instead, we would like to get θ_μ^μ in operator form it is necessary to work with S_W, not Γ. In other words, in calculating the two-loop vacuum graphs in the external field one should carefully separate—and discard—possible infrared contributions.

In QCD this has not been done yet. The question whether the higher order coefficients in the operator scale anomaly do or do not coincide with the higher order coefficients in the β function remains open. But in a simpler model, scalar electrodynamics, it is definitely proven[10] that such a coincidence does not take place. Indeed, in the scalar QED (cf. Eq. (2.25))

$$S_W(\mu) = \int d^4x \left\{ -\frac{1}{32\pi^2}\left[\frac{2\pi}{\alpha_0} + \frac{1}{3}\ln\frac{M_0}{\mu} + \frac{\alpha_0}{2\pi}\left(1-\frac{\xi}{2}\right)\ln\frac{M_0}{\mu}\right] F_{\mu\nu}F_{\mu\nu} + Z D_\mu\varphi^* D_\mu\varphi \right\}, \tag{2.22}$$

where $\alpha = e^2/4\pi$, $F_{\mu\nu}$ is the photon field tensor, φ is the complex scalar massless field and Z is its renormalization constant

$$Z = 1 - \left(1+\frac{\xi}{2}\right)\frac{\alpha_0}{\pi}\ln\frac{M_0}{\mu}. \tag{2.23}$$

Here M_0 is the ultraviolet cut off and ξ is the gauge parameter in the photon propagator ($D_{\mu\nu} = e^2[-g_{\mu\nu} + (\xi k_\mu k_\nu / k^2)]/k^2$).

Eq. (2.22) demonstrates that the two-loop contribution in the operator $F_{\mu\nu}^{2}$ does not coincide with that in the β function ($\beta^{(2)} = 1/\pi$). More than that, it is not even gauge invariant!

The explicit gauge invariance is recovered after proceeding from S_W to Γ, i.e., after calculating the photon matrix element of S_W. Formally the operator $\int d^4x D_\mu\varphi^* D_\mu\varphi$ vanishes in virtue of the equations of motion. One can convince oneself, however,[10] that in the external photon field

$$\langle D_\mu\varphi^* D_\mu\varphi\rangle = \frac{1}{64\pi^2}F_{\mu\nu}^2 \tag{2.24}$$

because of the infrared contribution.

Combining Eqs. (2.22)–(2.24) we arrive at

$$\Gamma = \int\left[-\frac{1}{32\pi^2}\left(\frac{2\pi}{\alpha_0} + \frac{1}{3}\ln\frac{M_0}{\mu} + \frac{\alpha}{\pi}\ln\frac{M_0}{\mu}\right)F_{\mu\nu}^2\, d^4x + \dots\right] \tag{2.25}$$

(cf. Eq. (2.22)). The two-loop coefficient in front of $F_{\mu\nu}^2$ in Γ corresponds, as it should, to the two-loop coefficient in the β function.

2.4. Anomalies and external currents

If in addition to "pure" QCD one introduces the interaction of quarks with photons or with other "external" currents then, apart from the axial and scale anomalies (2.9), (2.20), new anomalies appear and, moreover, some new terms appear in the old anomalous relations for $\partial^\mu j_{\mu A}$ and θ_μ^μ. For simplicity let us consider only the photons leaving aside the possibility of other external currents.

Consider one of the flavor-octet axial currents

$$a_\mu^3 = \bar{u}\gamma_\mu\gamma_5 u - \bar{d}\gamma_\mu\gamma_5 d \tag{2.26}$$

with the π^0 quantum numbers. It is quite clear that in pure QCD $\partial^\mu a_\mu^3 = 0$ (see Subsec. 2.2). With the photons switched on, however, the divergence of a_μ^3 is non-vanishing. From Fig. 5 it is easy to understand that

$$\partial^\mu a_\mu^3 = \frac{\alpha}{2\pi}N_c(Q_u^2 - Q_d^2)F_{\mu\nu}\tilde{F}_{\mu\nu}, \tag{2.27}$$

where $N_c = 3$ is the number of colors and $Q_u = 2/3$, $Q_d = -1/3$ are the u-, d-quark electric charges. The product $F_{\mu\nu}\tilde{F}_{\mu\nu}$ in the right-hand side of Eq. (2.27) is an operator with respect to the photons. If we were interested in the amplitude of the transition into two photons with the momenta $k^{(1)}$, $k^{(2)}$ and polarization vectors $\varepsilon^{(1)}$, $\varepsilon^{(2)}$ we would have

$$F\tilde{F} \to (-2)F^{(1)}\tilde{F}^{(2)}, \quad F_{\mu\nu}^{(i)} = k_\mu^{(i)}\varepsilon_\nu^{(i)} - k_\nu^{(i)}\varepsilon_\mu^{(i)}.$$

The expression (2.27) has far-reaching consequences. Let us ask the question: "what is the amplitude of the transition of the current a_μ^3—just the current itself, not $\partial^\mu a_\mu^3$—into two real photons?" In order to find the answer one should first of all carry out a kinematical analysis, i.e., write down all kinematical structures for the amplitude $a_\mu^3 \to 2\gamma$ with the additional conditions $k^{(1)2} = k^{(2)2} = 0$, plus gauge invariance with respect to both photons, plus Eq. (2.27). Omitting the details of this simple exercise we give the answer:

$$\langle 0 | a_\mu^3 | 2\gamma\rangle = -\frac{iq_\mu}{q^2}\frac{4\alpha}{\pi}N_c(Q_u^2 - Q_d^2)\,\varepsilon^{\mu\nu\alpha\beta}k_\mu^{(1)}\varepsilon_\nu^{(1)}k_\alpha^{(2)}\varepsilon_\beta^{(2)} + \text{a term transfers to } q, \tag{2.28}$$

where $q = k^{(1)} + k^{(2)}$.

The most striking point is the presence of the pole factor q_μ/q^2 in the right-hand side of Eq. (2.28).[14,15] It is worth emphasizing that Eq. (2.28) is the exact result of QCD valid, in particular, in the limit $q^2 \to 0$. Technically, the singularity in a_μ^3 at $q^3 \to 0$ is the consequence of the massless quark propagation in the triangle graph of Fig. 5. In the color confining theory, however, it is absolutely meaningless to speak about quarks as physics 1 degrees of freedom in the spectrum. Then, how can one understand the singularity of the amplitude $\langle 0|a_\mu^3|2\gamma\rangle$ in physical terms?

Only one explanation is possible. The physical spectrum should include massless colorless composite particles coupled both to photons and a_μ^3. The corresponding contri-

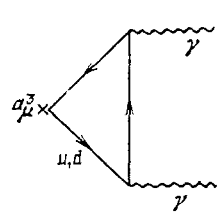

FIG. 5.

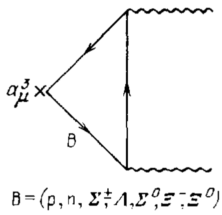

FIG. 6.

bution should saturate the matrix element $\langle 0|a_\mu^3|2\gamma\rangle$, or, more exactly, its singular part.

Thus, starting from massless quarks and assuming color confinement we succeeded in proving the existence of massless hadrons. The proof is based on the requirement of matching the singular contributions in a_μ^3 at the level of quarks and at the level of hadrons. This elegant idea belongs to G. 't Hooft[15] and is called the 't Hooft consistency condition.

Let us try to find out what particular hadrons are massless[3)]. In principle, there are two alternative variants, both leading to the pole singularity in a_μ^3:

(i) massless baryons;

(ii) massless pion.

Consider first the scenario (i) $M_B = 0$. In the theory with u, d, s quarks the spin 1/2 baryons form the well-known octet

$$B = (p,\ n,\ \Sigma^\pm,\ \Lambda,\ \Sigma^0,\ \Xi^-,\ \Xi^0). \tag{2.29}$$

The scenario (i)—a very important point—implies that both, the vector and axial SU(3) symmetries are realized linearly[4)] and, hence, the baryon-photon coupling constants and the constants $\langle B|a_\mu^3|B\rangle$ at zero momentum transfer are fixed (for instance, $\langle\Sigma^+|a_\mu^3|\Sigma^+\rangle = 2\bar{\Sigma}\gamma_\mu\gamma_5\Sigma$). Calculating the triangle diagram of Fig. 6 (more exactly, its pole part) we find that the baryon octet does not contribute to the pole in a_μ^3 due to the cancellations: the proton is canceled by Ξ^- and Σ^- by Σ^+. The other baryons from (2.29) decouple from the photon. The immediate conclusion is: scenario (i) is not compatible with the anomalous relation (2.27).

Thus, we are forced to conclude that it is scenario (ii) that is actually realized, and the massless hadron saturating Eq. (2.28) is the pion. The 't Hooft consistency condition then appears trivial (Fig. 7) and requires only a certain relation between the $\pi^0 \to a_\mu^3$ amplitude and the $\pi^0 \to 2\gamma$ coupling constant. The corresponding result has been known for a long time. It is nothing other than the famous PCAC prediction for $A(\pi^0 \to 2\gamma)$ in terms of f_π, the $\pi \to \mu\nu$ coupling constant ($\langle 0|a_\mu^3|\pi^0\rangle = \sqrt{2} i f_\pi q_\mu$ where q_μ is the pion momentum; for an excellent review of PCAC and a representative list of references see Ref. 18).

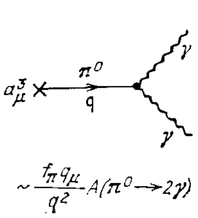

FIG. 7.

The existence of the massless pion coupled to the axial current a_μ^3 means the spontaneous breaking of the axial $SU(3)_A$ symmetry, or more exactly, the non-linear Goldstone realization of the $SU(3)_A$.

Summarizing, the analysis of the anomalies in QCD with the external currents allows one to get important additional information. In particular, we are able to get such a profound result as the spontaneous breaking of the flavor $SU(3)_A$ group and the presence of the octet of the Goldstone mesons in QCD.

A brief digression of historical nature is in order here. The Goldstone nature of the π, η, K mesons has been exploited in hadronic physics long before the QCD era. The soft pion technique created in the sixties on the basis of the hypothesis of approximate chiral invariance of strong interactions (in those days that was a brilliant hypothesis indeed) generated a series of elegant predictions in low-energy pion physics. The Goldberger-Treiman and Adler-Weisberger relations (see, e.g., Ref. 18) seem to be the most typical. The consequences of PCAC proved to be extremely useful in diverse applications. The agreement of the PCAC predictions with the data served as conclusive evidence in favor of the approximate chiral invariance of strong interactions. The latter, in turn, was the key argument used by the fathers of QCD[19] in order to introduce the vector coupling of gluons with quarks. Somewhat later the same ideas enabled one to determine the masses of the current u, d, s quarks[20] and fix the quark condensate.

Returning to the anomalies we recall that the introduction of photons (and other "external" currents) apart from creating the new anomalies modifies the old ones in $\partial^\mu j_{\mu A}$ and θ_μ^μ adding some new terms there. We postpone the discussion of the corresponding effects till Subsec. 3.3.

3. LOW-ENERGY QCD THEOREMS. APPLICATIONS TO HADRON PHYSICS

Low-energy theorems in field theory have been invented almost as long ago as field theory itself. Suffice it to recall the Low theorems[21] for the photon bremsstrahlung in the low-frequency limit.

As a rule, the low-energy theorems represent relations among amplitudes, or n-point Green's functions with different number of legs (particles). The relations between the amplitudes emerge as a reflection of some symmetry, exact or anomalous, existing in the theory. For instance, the Low theorems mentioned above are the consequence of gauge invariance of electromagnetic interactions.

The search for symmetries and constraints they impose on the observable quantities is of special importance in the theory of hadrons. The low-energy theorems, or the Ward identities stemming from the symmetry properties of the theory and independent of the unknown details of the confinement mechanism yield information on physical processes inaccessible by other existing methods. In purely theoretical aspect they serve as a reference point, a basis for further theoretical constructions.

Apart from the classical PCAC theorems[5)] in QCD there emerge additional predictions both for the Goldstone meson emission and for some other processes. The first predictions were obtained shortly after the invention of QCD (e.g., Ref. 22). The systematic analysis of the subject dates back to Refs. 23–26. Many applications based on relations (2.9), (2.20) have been worked out recently. Some of them will be discussed below.

The most nontrivial results refer to the amplitudes describing the conversion of gluons into hadrons and photons at large distances. The information obtained in this way supplements our understanding of the hadron structure and that of the QCD vacuum.

This section is devoted to derivation of the basic low-energy theorems specific to QCD. Their virtues will be demonstrated in a few instructive examples.

3.1. The scale Ward identities

If $m_{u,d,s} = 0$ the only mass parameter of the theory that appears at the quantum level, is

$$\mu = M_0 e^{-8\pi^2/bg_0^2}, \tag{3.1}$$

where as usual M_0 is the ultraviolet cutoff and g_0 is the corresponding coupling constant, $g_0 = g(M_0)$. Eq. (3.1) is a manifestation of the anomaly in the trace of the energy-momentum tensor,

$$\sigma(x) \equiv \theta^\mu_\mu(x) = -\frac{b\alpha_s}{8\pi} G^a_{\mu\nu} G^a_{\mu\nu} \tag{3.2}$$

derived in Sec. 2 (see Eq. (2.20)).

The classical scale invariance of the action, being lost in loops, still does not vanish without a trace (both, in the direct and mathematical meanings of these words). The remnant of this symmetry is a set of the low-energy theorems of the type

$$\lim_{q\to 0} i \int e^{iqx}\, dx \langle T\{\mathcal{O}(x), \sigma(0)\}\rangle_{\text{connected}} = -d_n \langle \mathcal{O}\rangle, \tag{3.3}$$

where $\mathcal{O}$ is an arbitrary local operator built from gluons and/or u,d,s quarks, d_n is its normal dimension. Actually, Eq. (3.3) is valid up to $\mathcal{O}(m_q)$ terms which are implied but will not be written out explicitly except in cases where it is necessary. To make the expressions more concise the subscript "connected" will be omitted in what follows.

A few important concrete examples are

$$i \int \langle T\{\sigma(x)\sigma(0)\}\rangle\, dx = -4\langle\sigma\rangle \tag{3.4}$$

and

$$i \int \langle T\{\bar{q}(x) q(x), \sigma(0)\}\rangle\, dx = -3\langle\bar{q}q\rangle. \tag{3.5}$$

Relations between different Green's functions involving $\sigma(x)$ exist not only for the vanishing external momenta. More exactly, it is not necessary to have *all* external momenta equal to zero. Those which appear in "foreign" operators $\mathcal{O}$ can be arbitrary.[27]

The derivations of these formulae are all of one type, and we will consider the simplest case, Eq. (3.3). First, rescale the gluon field,

$$\overline{G^a_{\mu\nu}} = g_0 G^a_{\mu\nu}.$$

Then the g_0 dependence in the Lagrangian, reduces to $-(1/4g_0^2)\overline{G}^2$, and consequently

$$i \int \langle T\{\mathcal{O}(x), \overline{G^2(0)}\}\rangle\, dx \equiv -\frac{d}{d(1/4g_0^2)} \langle\mathcal{O}\rangle. \tag{3.6}$$

The fact that the only mass parameter of the theory is given by Eq. (3.1) implies

$$\langle\mathcal{O}\rangle = \text{const}\cdot (M_0 e^{-8\pi^2/bg_0^2})^{d_n}. \tag{3.7}$$

Differentiating Eq. (3.7) we arrive at Eq. (3.3).

The simple derivation sketched above actually should be supplemented by some regularization since in most cases $\langle\mathcal{O}\rangle$ is divergent in perturbation theory (a rare exception is the operator $\bar{q}q$ in the chiral limit $m_q \to 0$). Ref. 23 shows that a careful regularization procedure involving a sufficient number of regulator fields does not alter our conclusions.

Let us proceed now to phenomenological consequences of the scale Ward identities.

In a world with no light quarks the lightest (and, hence, stable) particle would be scalar gluonium, σ_g. Study of its properties and, in particular, scattering amplitudes would be an important direction in "experimental" physics.

It is highly probable that the σ_g mass is small compared to the characteristic scale in this channel ($\sim 20\ \text{GeV}^2$).[23] From this point of view the situation resembles that with pions in the real world, $m_\pi^2 \ll 1\ \text{GeV}^2$. Of course, the smallness of m_π^2 is parametrical while the suppression of $m^2_{\sigma_g}$ seems to be of a numerical character. In spite of this fact one can formulate the following problem: construct the low-energy (effective) Lagrangian describing the interaction of σ_g and realizing the scale Ward identities (3.3)–(3.5) in the same way that the pion chiral Lagrangian realizes the chiral Ward identities at the tree level.

The solution has been given in Ref. 27 (more complicated cases including electromagnetism are considered in Ref. 28). The very fact that the solution can be found is far from being trivial and crucially depends on the sign of the vacuum energy: the solution is stable only provided that $\varepsilon_{\text{vac}} < 0$.

Thus, our task is to find an effective theory for the field $\sigma(x)$ [see Eq. (3.2)] in which $\theta^\mu_\mu = \sigma$. Recalling that θ^μ_μ is connected with dilatations it is quite natural to call $\sigma(x)$ the dilaton field. We note that the normal dimension of $\sigma(x)$ is equal to four. Therefore, if the kinetic term in the Lagrangian is

$$\mathcal{L}_{\text{kin}} = \text{const}\cdot(\partial_\mu\sigma)^2 \sigma^{-3/2}, \tag{3.8}$$

then the integral $\int d^4x \mathcal{L}_{\text{kin}}$ is scale-invariant and, hence, $\mathcal{L}_{\text{kin}}$ does not contribute to θ^μ_μ. We can concentrate on the potential term $\mathcal{L}_{\text{pot}}(\sigma)$ and choose it so as to ensure the desired equality $\theta^\mu_\mu = \sigma$.

The variation of σ under the scale transformation $x \to x(1+\varepsilon)$ is $\sigma \to \sigma(1-4\varepsilon)$. The corresponding variation of the action

$$\Delta_{\text{scale}} \int d^4x\, \mathcal{L}_{\text{pot}}(\sigma) = \int \left(4\mathcal{L}_{\text{pot}} - 4\sigma \frac{\delta\mathcal{L}_{\text{pot}}}{\delta\sigma}\right) d^4x, \tag{3.9}$$

according to the general rules is equal to $-\int d^4x \theta^\mu_\mu(x)$.

As a result we get

$$4\mathcal{L}_{\text{pot}} - 4\sigma \frac{\delta\mathcal{L}_{\text{pot}}}{\delta\sigma} = -\sigma, \tag{3.10}$$

The solution of this simple equation is

$$\mathcal{L}_{\text{pot}} = \frac{1}{4}\sigma(\ln\sigma + \text{const}). \tag{3.11}$$

The constants figuring in Eq. (3.8), (3.11) can be expressed

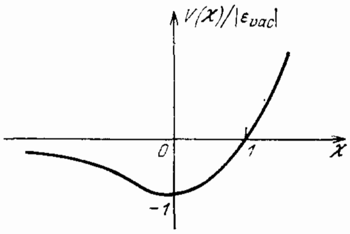

FIG. 8.

in terms of the mass of σ_g and the vacuum energy, m and ε_{vac}, respectively. It is convenient to redefine the field $\sigma = 4\varepsilon_{vac} \exp \chi$; then the final answer for the effective Lagrangian takes the form[6]

$$\mathcal{L}_{eff} = -\frac{\varepsilon_{vac}}{m^2}\frac{1}{2}(\partial_\mu \chi)^2 e^{\chi/2} + \varepsilon_{vac}(\chi - 1)\, e^{\chi}. \qquad (3.12)$$

Recall that $\varepsilon_{vac} < 0$ (Fig. 8).

Thus far we have actually used only the information coded in Eq. (3.4) and other analogous Ward identities. The theorems (3.3) determine the form of any local *colorless* operator $\mathcal{O}$ (in pure gluodynamics) in terms of a single constant,

$$\mathcal{O} = \langle \mathcal{O} \rangle_{vac} \exp\left(\frac{d_n}{4}\chi\right). \qquad (3.13)$$

In particular,

$$\sigma = \langle \sigma_{vac} \rangle e^{\chi} = 4\varepsilon_{vac} e^{\chi}.$$

One of the consequences of Eq. (3.13) is

$$\langle \mathcal{O}_1 \mathcal{O}_2 \rangle = \langle \mathcal{O}_1 \rangle \langle \mathcal{O}_2 \rangle.$$

The factorization property will not seem surprising if one recalls that the solution for the effective scalar field theory constructed in this section is valid in the tree approximation, i.e., in the leading order in $1/N_c$. At $N_c \to \infty$ the vacuum expectation values of the colorless operators do indeed factorize. It is instructive to check that the Lagrangian (3.12) satisfies the general requirements of the multicolor QCD[29,30]:

$$\mathcal{L}_{eff} = \frac{1}{2}(\partial_\mu \varphi)^2 - \frac{1}{2} m^2 \varphi^2 + \frac{1}{4}\frac{m}{|\varepsilon_{vac}|^{1/2}}(\partial_\mu \varphi)^2 \varphi + \dots$$
$$- \frac{2}{3!}\frac{m^3}{|\varepsilon_{vac}|^{1/2}}\varphi^3 + \dots,$$

where we have changed the normalization of the field x ($x \to \tilde{x}$) in order to ensure the standard normalization of the kinetic term and expanded the exponentials. Since $m \sim N_c^0$ and $\varepsilon_{vac} \sim N_c^2$ the behavior of the interaction vertices is in full accord with the general rules: the cubic constant αN_c^{-1}, the quartic constant αN_c^{-2}, etc.

The Lagrangian (3.12) exhaustively describes the interaction of the scalar gluonium at energies $\sim m_{\sigma_g}$. Unfortunately, for obvious reasons, we can not undertake a direct experimental test of the consequences stemming from Eq. (3.12).

The best we can get is a very rough indirect estimate which simultaneously shows the possible size of the effects induced by the light quarks. Assume that relations analogous to (3.13) are valid not only in pure gluodynamics but in the theory with quarks as well, both for gluon and quark operators. Given these hypotheses—and this automatically excludes the scalar quarkonium contribution and σ_q-σ_g mixing—we readily find the nucleon matrix element of $\bar{u}u + \bar{d}d$. At the tree level

$$\langle N | \bar{u}u + \bar{d}d | N \rangle = \langle \bar{u}u + \bar{d}d \rangle \cdot \frac{3}{4} \langle N | \chi | N \rangle$$
$$= \langle \bar{u}u + \bar{d}d \rangle \frac{3}{4} \frac{\langle N | \theta_{\mu\mu} | N \rangle}{\langle \theta_{\mu\mu} \rangle}$$
$$= \frac{3}{4} \langle \bar{u}u + \bar{d}d \rangle \frac{m_N}{4\varepsilon_{vac}} \bar{\Psi}_N \Psi_N, \qquad (3.14)$$

where Ψ_N is the nucleon wave function, $\langle N|\theta^\mu_\mu|N\rangle = m_N \bar{\Psi}_N \Psi_N$ and $\langle \theta^\mu_\mu \rangle = 4\varepsilon_{vac}$. On the other hand, the matrix element $\langle N|\bar{u}u + \bar{d}d|N\rangle$ is more or less known empirically, and amounts roughiy to $3\bar{\Psi}_N \Psi_N$. The theoretical estimate (3.14) yields $\sim 1.8\bar{\Psi}_N \Psi_N$ provided the standard values of the quark and gluon condensates are substituted. We note that the signs of the empirical and theoretical estimates coincide while the absolute values differ by a factor ~ 2. We conclude that the results for the scalar gluonium obtained in pure gluodynamics are valid in the real world up to a factor of 2. On the other hand, the effect due to switching on the light quarks is of order unity.

3.2. Correlation functions of the topological density

The operator

$$K = \frac{g^2}{32\pi^2} G^a_{\mu\nu} \tilde{G}^a_{\mu\nu} \qquad (3.15)$$

can be naturally called the density of the topological charge (cf. Eq. (1.35)). If CP invariance is not violated it is evident that $\langle K \rangle = 0$. Of physical interest are the correlation functions of K with other CP-odd quantities.

Below we will show that[26]

$$i \int \langle T\{K(x),\, K(0)\}\rangle\, dx \equiv \Delta_K, \qquad (3.16)$$

$$i \int \left\langle T\left\{ \sum_{q=u,\,d} m_q \bar{q}(x)\, i\gamma_5 q(x),\, K(0)\right\}\right\rangle dx = -2\Delta_K, \qquad (3.17)$$

where

$$\Delta_K = -\frac{1}{8} f_\pi^2 m_\pi^2 \frac{4 m_u m_d}{(m_u + m_d)^2} = O(m_q). \qquad (3.18)$$

For simplicity it is assumed that there are only two light quarks, u and d. Let us draw the reader's attention to the fact that the two-point function (3.16) vanishes for massless quarks. That is why in this subsection, unlike the previous ones, we do not put $m_q = 0$.

Eq. (3.18) is trivially generalizable to the case of three light quarks.

The starting point for the derivation is the anomalous relation for a^0_μ, the SU(2) analog of the current $j_{\mu A}$ introduced in Eq. (2.6):

$$\partial_\mu a^0_\mu = 2(m_u \bar{u} i\gamma_5 u + m_d \bar{d} i\gamma_5 d) + 4K,$$
$$a^0_\mu = \bar{u}\gamma_\mu\gamma_5 u + \bar{d}\gamma_\mu\gamma_5 d. \qquad (3.19)$$

It is convenient to introduce the following auxiliary two-point functions

$$\pi_{\mu\nu}(q) = i \int e^{iqx} \langle T\{a^0_\mu(x),\, a^0_\nu(0)\}\rangle\, dx,$$
$$\tilde{\pi}_\mu(q) = i \int e^{iqx} \langle T\{a^0_\mu(x),\, K(0)\}\rangle\, dx. \qquad (3.20)$$

The fact that there are no physical massless states coupled to a_μ^0 implies

$$q_\mu q_\nu \pi_{\mu\nu}(q) \to 0, \quad q_\mu \tilde{\pi}_\mu(q) \to 0 \quad \text{при} \quad q \to 0, \tag{3.21}$$

($m_\pi \neq 0$ since the u-, d-quark masses are taken into account explicitly).

Integrating by parts and using Eq. (3.19) we reduce $q_\mu q_\nu \pi_{\mu\nu}$ and $q_\mu \tilde{\pi}_\mu$ to the T products of the divergences plus a contact term coming from the equal-time commutator $[\bar{q}(x)\gamma_5 q(x), a_0^0(0)]\delta(x_0)$:

$$0 = (q_\mu q_\nu \pi_{\mu\nu})_{q=0} = \langle 4(m_u \bar{u}u + m_d \bar{d}d)\rangle + 16\{K, K\} + 16\{\sum_{q=u,\,d} m_q \bar{q} i\gamma_5 q,\ K\} + 4\{\sum_{q=u,\,d} m_q \bar{q} i\gamma_5 q,\ \sum_{q=u,\,d} m_q \bar{q} i\gamma_5 q\}, \tag{3.22}$$

$$0 = (q_\mu \tilde{\pi}_\mu)_{q=0} = 4\{K, K\} + 2\{\sum_{q=u,\,d} m_q \bar{q} i\gamma_5 q,\ K\}.$$

The braces here stand for $\{\ldots\} \to \langle i\int dx T\{\ldots\}\rangle$, cf. Eq. (3.16).

Invoking the well-known PCAC relation $(m_u + m_d)\langle \bar{u}u + \bar{d}d\rangle = -f_\pi^2 m_\pi^2$ and eliminating the "superfluous" correlator from Eq. (3.22) we arrive at

$$16\{K, K\} = -2f_\pi^2 m_\pi^2 + O(m_{u,\,d}^2) + 4\{\sum_{q=u,\,d} m_q \bar{q} i\gamma_5 q,\ \sum_{q=u,\,d} m_q \bar{q} i\gamma_5 q\}. \tag{3.23}$$

The T product in the right-hand side is superficially proportional to $m_{u,d}^2$ while only terms linear in m_q are to be kept in the accepted approximation. Such terms appear, however, due to the intermediate states with $m^2 \sim m_{u,d}$. It is important that there is only one such state-pion-with the quark content $\bar{u}u, \bar{d}d$. As is known,[31] the singlet state $\sim(\bar{u}u + \bar{d}d)$ does not become massless in the chiral limit because of the gluon anomaly in the singlet axial current.

Taking into account that the pion residue is fixed by PCAC,

$$\langle 0 | \bar{u} i\gamma_5 u | \pi^0\rangle = -\langle 0 | \bar{d} i\gamma_5 d | \pi^0\rangle = \frac{f_\pi m_\pi^2}{\sqrt{2}(m_u + m_d)}. \tag{3.24}$$

and substituting Eq. (3.24) in Eq. (3.23) we immediately get the theorem quoted above (see Eq. (3.16)–(3.18)). We note that in the particular case $m_u = m_d$ Eq. (3.18) has been first obtained in Ref. 22 in another context.

It is in order to make here two remarks. First of all, it is worth emphasizing the connection between the theorem (3.16)–(3.18) and the U(1) problem. If the mass of the singlet pseudoscalar meson were vanishing in the chiral limit, as one would expect naively, the pseudoscalar singlet should be included in Eq. (3.23) along with the pion and this would completely kill $-2f_\pi^2 m_\pi^2$ in the right-hand side. Then $\{K,K\}$ would vanish, at least in the linear in m_q approximation. Thus, the effect is of the qualitative character:

$$\begin{aligned} &\text{valid U(1)-symmetry} \to i\int \langle T\{K(x),\ K(0)\}\rangle\, d^4x = 0,\\ &\text{no U(1)-symmetry} \to i\int \langle T\{K(x),\ K(0)\}\rangle\, d^4x \neq 0. \end{aligned} \tag{3.25}$$

The second remark concerns the three quark generalization of the theorem. Repeating the derivation with the s quark included one can easily convince oneself that the only change is the expression for Δ_K, namely,

$$(\Delta_K)_{u,\,d,\,s} = -\frac{3}{8} f_\pi^2 (m_\pi^2 + m_\eta^2) \times \frac{m_u m_d m_s}{(m_u + m_d + m_s)(m_u m_d + m_u m_s + m_d m_s)}. \tag{3.26}$$

In the limit $m_s/m_{u,d} \gg 1$ Eq. (3.26) reduces to (3.18). (The PCAC relation $m_\pi^2/m_\eta^2 = [3(m_u + m_d)]/(4m_s)$ is helpful in verifying the latter assertion.)

The application of the low energy theorem (3.16) which immediately comes to mind is the calculation of the mass of the so-called phantom axion.[32,26] The details of this construction are irrelevant for us here. It is only important that the phantom axion is a pseudoscalar particle with zero classical mass and an effective interaction of the type

$$\mathscr{L}_{\text{eff}}^{(a)} = \frac{g^2}{32\pi^2 \sqrt{2}\varphi_0} a(x) G_{\mu\nu}^a(x) \tilde{G}_{\mu\nu}^a(x), \tag{3.27}$$

where φ_0 is a large parameter of dimension of mass and $a(x)$ is the axion field. The axion mass is generated at the quantum level due to the fact that the correlation function (3.16) is not zero. Combining Eqs. (3.27), (3.16), (3.18) we easily get

$$m_a = \left(\frac{1}{2|\varphi_0|^2}|\Delta_K|\right)^{1/2} = \frac{f_\pi m_\pi}{4|\varphi_0|}\left[\frac{4m_u m_d}{(m_u + m_d)^2}\right]^{1/2}. \tag{3.28}$$

3.3. Conversion of gluons into pions and photons

Any QCD expert unfamiliar with the low-energy theorems would say that the matrix elements such as

$$\langle 0 | \alpha_s G^2 | 2\pi\rangle,\ \langle 0 | \alpha_s G^2 | 2\gamma\rangle,\ \langle 0 | \alpha_s G\tilde{G} | 2\gamma\rangle \tag{3.29}$$

can not be reliably calculated in the present-day theory since they are determined by large distance dynamics. The last part of this assertion is true, of course, and still, surprising though it is, the gluon conversion amplitudes (3.29) are unambiguously calculable in the limit of small momenta of the final particles. For instance,

$$\left\langle 0 \left| -\frac{b\alpha_s}{8\pi} G_{\mu\nu}^a G_{\mu\nu}^a \right| \pi^+(p_1)\pi^-(p_2)\right\rangle = (p_1 + p_2)^2 + O(p^4). \tag{3.30}$$

The proof is based on the consideration of the matrix element $\langle \pi^+(p_1)\pi^-(p_2)|\theta_{\mu\nu}|0\rangle$ where $\theta_{\mu\nu}$ is the complete regularized energy-momentum tensor in QCD. The most general expression for $\langle \pi^+\pi^-|\theta_{\mu\nu}|0\rangle$ is

$$\langle \pi^+(p_1)\pi^-(p_2) | \theta_{\mu\nu} | 0\rangle = A r_\mu r_\nu + B q_\mu q_\nu + C g_{\mu\nu}, \tag{3.31}$$

where $r = p_1 - p_2$, $q = p_1 + p_2$, and A, B, C are some scalar functions of the four-momenta. For soft pions, however, these functions reduce to constants fixed by the following requirements (in the chiral limit):

(i) symmetry and conservation of $\theta_{\mu\nu}$;

(ii) normalization condition $\langle \pi^+(p)|\theta_{\mu\nu}|\pi^-(p)\rangle = 2p_\mu p_\nu$;

(iii) neutrality of $\theta_{\mu\nu}$ with respect to the axial charge, $[\theta_{\mu\nu}(x), Q_5^+] = 0$.

The points (i) and (ii) are obvious by themselves. The point (iii) is slightly less trivial. Recall, however, that in the chiral limit the axial charge Q_5^+ is a conserved operator and, hence,

$$[Q_5, \theta_{00}(x)] = i\partial_\mu a_\mu(x) = 0.$$

The vanishing of the commutator $[\theta_{\mu\nu}, Q_5^+]$ can be also checked directly since explicit expression for $\theta_{\mu\nu}$ and Q_5^+ in terms of the fundamental fields are known.

It is convenient to rewrite the condition (iii) in the following equivalent form:

B') $\langle \pi^+\pi^- \mid \theta_{\mu\nu} \mid 0\rangle \to 0$, valid p_1 or $p_2 \to 0$.

The equivalence of (iii) and (iii') becomes clear if one takes into account that $\langle \pi^+\pi^- | \theta_{\mu\nu}|0\rangle_{p_1\to 0} = if_\pi^{-1}\langle \pi^+ | [\theta_{\mu\nu}, Q_5^+]|0\rangle$.

Combining (i), (ii), and (iii') we arrive at the following result

$$\langle \pi^+\pi^- \mid \theta_{\mu\nu} \mid 0\rangle = \frac{1}{2} r_\mu r_\nu - \frac{1}{2} q_\mu q_\nu + \frac{1}{2} q^2 g_{\mu\nu}, \qquad (3.32)$$

$$\langle \pi^+\pi^- \mid \theta_{\mu\mu} \mid 0\rangle = q^2. \qquad (3.33)$$

Using the scale anomaly (2.20) we reproduce the theorem (3.30).[7)]

The left-hand side of Eq. (3.30) is proportional to α_s while the right-hand side does not contain such a factor. This "disbalance" holds for the conversion of gluons into photons as well (see below) and results in far-reaching consequences.

Without any connection with hadron physics it is worth noting that the ideas presented above have been exploited in Ref. 33 treating the propagation of Goldstone bosons in the gravitational background. Eq. (3.33) actually means that the corresponding equation does not contain the term proportional to the scalar curvature, the so called ξ term.

If the pion matrix element (3.29) measures the coupling of gluons (in the 0^+ channel) with the "real" quarks, the photon amplitudes probe the virtual quark loops since there is no direct photon-gluon interaction. Below we will show that

$$\left\langle 0 \left| \frac{\alpha_s}{4\pi} G^a_{\alpha\beta}\tilde{G}^a_{\alpha\beta} \right| \gamma(k_1)\gamma(k_2)\right\rangle = \frac{\alpha}{\pi} N_c \langle Q_q^2\rangle F^{(1)}_{\mu\nu}\tilde{F}^{(2)}_{\mu\nu}, \qquad (3.34)$$

where $F^{(i)}_{\mu\nu} = k^{(i)}_\mu \varepsilon^{(i)}_\nu - k^{(i)}_\nu \varepsilon^{(i)}_\mu$.

Eq. (3.34) is valid in the chiral limit and, moreover, assumes that the momentum transfer $(k_1 + k_2)^2$ is small compared to the characteristic hadron scale ~ 1 GeV.

The proof of this relation is very simple. In order to derive the theorem, Eq. (3.34), multiply the amplitude

$$\langle 0 \mid \sum_q \bar{q}\gamma_\mu\gamma_5 q \mid \gamma(k_1)\gamma(k_2)\rangle$$

by $q_\mu \equiv (k_1 + k_2)_\mu$ and let $q \to 0$. On the other hand, since there are no massless particles in the singlet channel even in the limit $m_q = 0$, we must get zero. On the other hand, the explicit computation of the divergence of the axial current yields

$$n_f \frac{\alpha_s}{4\pi} G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} - N_c \frac{\alpha}{\pi} F^{(1)}_{\mu\nu}\tilde{F}^{(2)}_{\mu\nu} \sum_q Q_q^2, \qquad (3.35)$$

where both the gluon and photon parts of the chiral anomaly are taken into account.

The low-energy theorem for the matrix element is very peculiar and interesting. For a detailed discussion see Ref. 48.

Concluding the subsection we give without any details the expressions for the conversion of the operator $G\tilde{G}$ into π and η[34–36]

$$\left\langle 0 \left| \frac{3\alpha_s}{4\pi} G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} \right| \pi^0\right\rangle = \frac{3}{\sqrt{2}} \frac{m_d - m_u}{m_d + m_u} f_\pi m_\pi^2,$$
$$\left\langle 0 \left| \frac{3\alpha_s}{4\pi} G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} \right| \eta\right\rangle = \sqrt{\frac{3}{2}} f_\pi m_\eta^2. \qquad (3.36)$$

3.4. Correlation function of "mixed" currents

An interesting low-energy theorem has been obtained in Ref. 46. It refers to the correlation function of the currents

$$J^{ij}_\mu = g\bar{q}^i(x)\gamma_\alpha G^a_{\alpha\mu}(x) t^a q^j(x),$$

where t^a stands for the Gell-Mann matrices, i is the flavor index, $q^1 = u, q^2 = d, q^3 = s$ (it is assumed that $i \neq j$; for coinciding i, j one should keep only the flavor-non-singlet part). The current J^{ij}_μ is rather unusual, it contains both the quark and gluon fields.

The result of Ref. 46 can be summarized as follows:

$$\Pi^{ij}_{\mu\nu} = i \int d^4x \langle T\{J^{ij}_\mu(x), J^{ij}_\nu(0)\}\rangle$$
$$= \frac{g^2}{8} g_{\mu\nu} \left\langle \Big(\sum_{q=u,d,s} \bar{q}\gamma_\beta t^a q\Big) \ (\bar{q}^i\gamma_\beta t^a q^i + \bar{q}^j\gamma_\beta t^a q^j)\right\rangle.$$

The left-hand side in this equality is a two-point function at zero momentum transfer while the right-hand side is the vacuum expectation value of a local operator. If one leaves the chiral limit and switches on the quark mass, there arise $O(m_q^2)$ corrections on the right-hand side.

The derivation of this relation conceptually does not differ from that for other relations of this kind. The starting observation[46] is the fact that the current J^{ij}_μ reduces to the divergence of an operator resembling the quark piece of the energy-momentum tensor (but flavor-non-singlet): Namely,

$$J^{ij}_\mu = \partial_\alpha \theta^{ij}_{\mu\nu},$$

where

$$\theta^{ij}_{\mu\alpha} = \bar{q}^i\gamma_\mu i D_\alpha q^j.$$

Moreover, the authors of Ref. 46 analyze the two-point function $\langle T\{\theta^{ij}_{\mu\alpha}(x), J^{ij}_\nu(0)\}\rangle$. After differentiating it with respect to x they get, in the limit of zero momentum transfer on the one hand, zero (there are no massless particles in this channel) and, on the other hand, the correlation function of interest plus the equal-time commutator $[\theta^{ij}_{\mu\alpha}, J^{ij}_\nu]$ plus the Schwinger term. Computing the commutator and the Schwinger term they arrive at the theorem quoted above. The theorem can be (and has been) used for describing deep inelastic lepton-hadron scattering and for predicting the masses of hybrid mesons within the QCD sum rules.

In the following sections we will discuss further some situations in which the low-energy theorems help to solve concrete problems of hadron physics.

3.5. Effective coupling of the Higgs boson with nucleons and pions

If a light Higgs boson H (with mass $\lesssim 1$ GeV) existed in nature—and this, generally speaking, is not ruled out—its production cross section and decay properties would be largely determined by the $H\bar{N}N$ and $H\pi\pi$ constants. Even if there is no such boson calculation of these constants, it is a wonderful exercise which will bring pleasure to any theorist.

Thus, consider the Higgs boson H whose interaction with quarks is as follows:

$$\mathscr{L}_{\text{int}} = -\frac{1}{\eta} H \sum_q m_q \bar{q} q, \tag{3.37}$$

where $\eta = (G_F \sqrt{2})^{-1/2}$ is the vacuum expectation value in the Glashow-Weinberg-Salam standard model. We will treat η, however, as an independent parameter.

The first question we address is: "What is the form of the $H\bar{N}N$ vertex at small momentum transfers?" The answer has been given in Ref. 37.

Since nucleons are basically built from u and d quarks it is tempting to say that the dominant role in (3.37) belongs to the terms $m_u \bar{u}u + m_d \bar{d}d$. Recall, however, that the masses of the current u and d quarks—and it specifically is the current quarks that enter Eq. (3.37)—are very small. If one parametrizes the $H\bar{N}N$ vertex as $(-\lambda/\eta)H\bar{N}N$, where λ is a constant of dimension of mass, then naively

$$(\lambda)_{u,d} \approx 2m_u + m_d \approx 15 \quad \text{MeV for a proton}, \tag{3.38}$$

and this is quite negligible. In what follows we assume that $m_u = m_d = 0$.

Rather paradoxically, the $H\bar{N}N$ coupling is determined essentially by the Higgs interaction with heavy quarks, i.e., those whose mass $m_Q \gtrsim M_H/2$. Since there are no such quarks inside the nucleon (more exactly, their admixture is supposed to be small) the heavy quarks show up only in virtual loops, thus inducing the effective Lagrangian of the gluon-boson interaction (Fig. 9). The form of this Lagrangian is fixed either by a direct computation of the diagram given in Fig. 9 or, equivalently, by differentiating the known loop of Fig. 10 with respect to m_Q (for details see Ref. 38). We have

$$\mathscr{L}_{\text{eff}} = \frac{1}{\eta}\frac{1}{12} H n_Q \frac{\alpha_s}{\pi} G^a_{\mu\nu} G^a_{\mu\nu}, \tag{3.39}$$

where n_Q is the number of heavy quarks. Thus, the problem is reduced to the calculation of the matrix element $\langle N|(\alpha_s/\pi)G^2|N\rangle$ at zero momentum transfer. To perform the calculation we invoke the anomalous relation of the type (2.20). If heavy quarks *are* present in the theory *but we are interested in low-energy processes*

$$\theta_{\mu\mu} = -\frac{9\alpha_s}{8\pi} G^a_{\mu\nu} G^a_{\mu\nu} + O(m_Q^{-2}).$$

Let us draw the reader's attention to the fact that at $m_Q \to \infty$ heavy quarks decouple, and the coefficient in the Gell-Mann-Low function is determined only by gluons and three

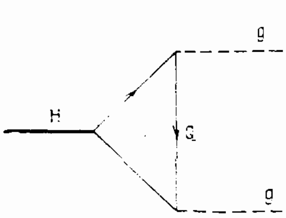

FIG. 9.

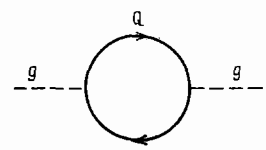

FIG. 10.

massless quarks (the constant b in Eq. (2.20) is replaced by 9), in full accord with the decoupling theorem of Ref. 39.

On the other hand, from the definition of the energy-momentum tensor we, evidently, get

$$\langle N | \theta_{\mu\mu} | N \rangle = m_N \bar{\Psi}_N \Psi_N. \tag{3.40}$$

The nucleon mass results from averaging the purely gluonic operator G^2(!) over the nucleon state. Combining Eqs. (3.40) and (3.39) we find that for every heavy quark

$$\lambda_Q = \frac{2}{27} m_N \approx +70 \text{ MeV} \tag{3.41}$$

a factor of ~ 5 larger than $(\lambda)_{u,d}$.

The same ideas are applicable in the problem of the Higgs-boson-pion interaction.[40] Just as in the nucleon case the direct interaction with the light quarks can be neglected, and the induced vertex (3.39) plays the dominant role.

The conversion of gluons into a pair of Goldstone mesons ($\pi\pi$, $\eta\eta$, $K\bar{K}$) is described by Eq. (3.30). Accounting for the fact that in the $H \to \pi^+\pi^-$ decay $(p_1 + p_2)^2 = M_H^2$, Eqs. (3.30) and (3.39) imply

$$A(H \to \pi^+\pi^-) = -\frac{1}{\eta}\frac{2n_Q}{27} M_H^2. \tag{3.42}$$

This striking result demonstrates that for a 1 GeV Higgs the transition into $\pi^+\pi^- + \pi^0\pi^0$ is the dominant decay mode, a fact discovered quite recently.[40] Previously theorists believed that such a Higgs particle, if it exists, should predominantly decay into $\mu^+\mu^-$ or $K\bar{K}$. For three heavy flavors (c,b,t) Eq. (3.42) yields

$$\frac{\Gamma(H \to \pi\pi)}{\Gamma(H \to \mu^+\mu^-)} \approx \frac{1}{27}\frac{M_H^2}{m_\mu^2} \sim 10. \tag{3.43}$$

A detailed discussion including analysis of possible corrections to Eq. (3.42) is given in Ref. 48.

Leaving aside for a while the main topic it is worth noting that the same Ref. 40 suggests an elegant general method for calculating the bremsstrahlung of soft Higgs bosons in any process with the light quarks and/or gluons. Indeed, for the vanishing Higgs boson momentum the H field in Eq. (3.39) reduces to a constant, and this Lagrangian, being added to the original QCD Lagrangian

$$\mathscr{L}_{\text{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^a_{\mu\nu},$$

simply changes the coupling constant α_s. What is the effective shift in α_s? The simplest way to answer this question lies in the rescaling of the fields which has been already used in deriving Eq. (3.4). In terms of $\bar{G}^a_{\mu\nu}$

$$\mathscr{L}_{\text{QCD}} + \mathscr{L}_{\text{eff}} = -\frac{1}{4}\left(\frac{1}{g^2} - \frac{H}{\eta}\frac{n_Q}{12\pi^2}\right) \overline{G^a_{\mu\nu}}\, \overline{G^a_{\mu\nu}}.$$

This implies

$$\delta\alpha_s = \frac{H}{\eta} n_Q \frac{\alpha_s^2}{3\pi}. \tag{3.44}$$

Therefore, if the α_s dependence of the amplitude of some "basic" process $a \to b$ is known one can immediately determine the amplitude of the "daughter" process $a \to b$ + soft H. Namely,

$$A(a \to b + \text{soft H}) = \frac{n_Q}{\eta} \frac{\alpha_s^2}{3\pi} \frac{\partial}{\partial \alpha_s} A(a \to b). \tag{3.45}$$

(It is worth emphasizing once more that real heavy quarks should not participate in the "basic" process; otherwise the gluon mechanism considered here will be overshadowed by the direct coupling $H\bar{Q}Q$.)

For the perturbative part of the amplitude associated with the hard gluons $A(a \to b) \sim \alpha_s^k$, and according to (3.45) the Higgs boson bremsstrahlung is suppressed by an extra power of α_s. The soft Higgs bosons are shaken off much more efficiently if the transition $a \to b$ is generated by nonperturbative effects. Indeed, in this case $A(a \to b) \sim \Lambda_{\text{QCD}}^d \sim \exp(-2\pi d/b\alpha_s)$ (cf. Eq. (3.7)). Then the amplitude of the soft Higgs bremsstrahlung is

$$A(a \to b + \text{H})_{p_\text{H} \to 0} = d \cdot \frac{2}{27} n_Q \eta^{-1} A(a \to b). \tag{3.46}$$

3.6. The pion spectrum in the decays $\psi' \to J/\psi\pi\pi, \Upsilon' \to \Upsilon\pi\pi$

The qualitative experimental situation in these decays is as follows. It is known that the pions are emitted predominantly in the S wave. The contribution of the D-wave pions does not exceed a few percent in the total decay rate. Besides that, the effective amplitude

$$A_{\text{eff}}(q^2) = \left[\frac{d\Gamma}{dq^2}(\Phi.\text{o.})^{-1}\right]^{1/2} = \left[(\text{S-wave})^2 + (\overline{\text{D-wave}})^2\right]^{1/2} \tag{3.47}$$

is well approximated by a linear dependence of the type

$$A_{\text{eff}} = C(q^2 - \delta). \tag{3.48}$$

(The bar in Eq. (3.47) denotes averaging over the pion phase space, C and δ in Eq. (3.48) are numerical constants, $q = p_1 + p_2$).

The questions addressed to the theory are:

—Why is the relative contribution of the D wave so small?

—Could one have predicted the linear law (3.48) beforehand?

—What does the theory say about C and δ?

The starting point of the theoretical analysis is the so called multipole expansion[41–43], which reflects the fact that we are dealing here with a two-step process. At the first step the gluons are emitted by the heavy quarks (quasi) locally[8)] and at the second step conversion of gluons into pions at large distances occurs. As a result, the amplitude factorizes:

$$A(n_i{}^3S_1 \to n_f{}^3S_1 + \pi\pi)$$
$$= \langle A_{if}\rangle \langle 0 \mid \alpha_s \mathbf{E}^a\mathbf{E}^a \mid \pi\pi\rangle + \text{higher multiple}, \tag{3.49}$$

where $\mathbf{E}^a$ is the chromoelectric field and the coefficient $\langle A_{if}\rangle$ contains information only about heavy quarkonium. One can find it using a model or phenomenologically, as is done in Ref. 24 where $\langle A_{if}\rangle_{J/\psi}$ is extracted from $\psi' \to J/\psi\eta$.

Now we are interested not in the heavy quarkonium structure but in the gluon dynamics at large distances. Therefore, we will treat $\langle A_{if}\rangle$ as a given number and will concentrate on the pion matrix element. Since our aim is pedagogical we will work in the chiral limit, $m_\pi^2 = 0$.

A trivial but important observation is

$$\alpha_s \mathbf{E}^a\mathbf{E}^a = \frac{\alpha_s}{2}(\mathbf{E}^a\mathbf{E}^a + \mathbf{H}^a\mathbf{H}^a) + \frac{\alpha_s}{2}(\mathbf{E}^a\mathbf{E}^a - \mathbf{H}^a\mathbf{H}^a) = \alpha_s\theta_{00}^G + \frac{2\pi}{b}\theta_{\mu\mu}, \tag{3.50}$$

where $\mathbf{H}^a$ is the chromomagnetic field and θ_{00}^G is the 00 component of the gluon part of the energy-momentum tensor.

Using the theorem (2.20) we get

$$\langle \pi^+\pi^- \mid \alpha_s \mathbf{E}^a\mathbf{E}^a \mid 0\rangle = \frac{2\pi}{b} q^2 + \langle \pi^+\pi^- \mid \alpha_s(\mu)\,\theta_{00}^G(\mu) \mid 0\rangle. \tag{3.51}$$

In the right-hand side we had to indicate the normalization point of the operator $\alpha_s\,\theta_{00}^G$ since it is not renormalization-group-invariant; μ is of order of the inverse quarkonium radius.

Repeating almost literally the arguments leading to Eqs. (3.32), (3.33) one can easily convince oneself that

$$\langle \pi^+(p_1)\,\pi^-(p_2) \mid \theta_{\mu\nu}^G \mid 0\rangle = \rho_\pi^G(\mu)\left(\frac{1}{2} r_\mu r_\nu - \frac{1}{2} q_\mu q_\nu\right) + \left(\frac{11}{36} + \frac{1}{4}\rho_\pi^G(\mu)\right) q^2 g_{\mu\nu}, \tag{3.52}$$

where ρ_π^G is the gluon share of the pion momentum which is more or less known empirically, $\rho_\pi^G(\mu \sim 1\ \text{GeV}) \simeq 0.4$.

At this stage it is convenient to rewrite the $\psi' \to J/\psi\pi\pi$ amplitude (for definiteness we will speak about this decay) in the Lorentz-invariant form:

$$A(\psi' \to J/\psi\pi^+\pi^-) = \frac{-\langle A_{1\text{f}}\rangle}{M_{\psi'}^2}(\varepsilon'_\alpha\varepsilon_\alpha) \times \left(\frac{2\pi}{b} q^2\mathcal{P}^2 + \langle \pi^+\pi^- \mid \alpha_s(\mu)\,\theta_{\beta\gamma}^G(\mu) \mid 0\rangle \mathcal{P}_\beta\mathcal{P}_\gamma\right),$$
$$\mathcal{P} \equiv \mathcal{P}(\psi'). \tag{3.53}$$

In the ψ' rest frame Eq. (3.53) reduces to Eq. (3.49).

Omitting a number of simple-arithmetical operations, we quote the final answer[25] for the amplitude:

$$A(\psi' \to J/\psi\pi^+\pi^-) = -\langle A_{1\text{f}}\rangle\langle\varepsilon'_\alpha\varepsilon_\alpha\rangle \times \frac{2\pi}{b}\left\{\left[q^2 - \varkappa\frac{(q\mathcal{P})}{\mathcal{P}^2}\right] + \frac{3}{2}\varkappa\left[r_\beta r_\gamma - \frac{1}{3}(q_\beta q_\gamma - q^2 g_{\beta\gamma})\right]\frac{\mathcal{P}_\beta\mathcal{P}_\gamma}{\mathcal{P}^2}\right\}, \tag{3.54}$$

$$\varkappa = \frac{b}{6\pi}\alpha_s(\mu)\,\rho_\pi^G(\mu).$$

The term in the first square brackets represents a pure S wave, while that in the second square brackets represents the D wave. (This is clearly seen in the $\pi^+\pi^-$ center of mass system.) Some nonessential $O(\alpha_s)$ corrections to the S wave amplitude are omitted in Eq. (3.54).

Thus, the $\psi' \to J/\psi\pi\pi$ amplitude is completely specified. Our theoretical prediction (3.54) possesses the following properties:

(i) The D wave contribution in A_{eff} (see Eq. (3.47)) relatively to that of the S wave is equal to $\varkappa^2/5$. Although numerically $\varkappa$ is rather uncertain—the normalization point μ is rather low and constitutes several hundred MeV for ψ' and about 1 GeV for Υ—still a very strong suppression of the

D wave is undeniable. Indeed, if $\alpha_s(\mu)\sim 0.7$, $\rho_\pi^G(\mu)\sim 0.4$, then

$$\varkappa_{\psi'}\sim 0{,}15-0{,}2 \tag{3.55}$$

and $\varkappa^2/5\sim 10^{-2}$. This fact implies that the gluons "like" to convert into the quark states in the 0^+ channel but are reluctant to do that in the 2^+ channel where the mixing parameter is smal, $\sim\alpha_s/\pi$.

(ii) Neglecting now the D wave, which is perfectly justified, we can write down the amplitude as follows:

$$A(\psi'\to J/\psi\pi\pi)=-\langle A_{1t}\rangle(\varepsilon'_\alpha\varepsilon_\alpha)\cdot\frac{2\pi}{b}[q^2-\varkappa(\Delta M)^2]; \tag{3.56}$$

where $\Delta M=M_{\psi'}-M_{J/\psi}$. This result nicely reproduces the empirical linear q^2 dependence (3.48). The intercept in q^2 in Eq. (3.56) lies in the vicinity of 0.1 GeV^2 if $\varkappa=0.2$ ($\delta_{\psi'\,\mathrm{exp}}$ is also approximately equal to 0.1 GeV^2).

Inspired by this success in the charmonium family we can attempt to find the pion spectrum intercept δ_Υ for $\Upsilon'\to\Upsilon\pi\pi$. The accuracy of both the multipole and α_s expansions is expected to be much better here because of the smaller size of Υ. The μ dependence of ρ_π^G—is known to be very weak and the main change is due to a falling off of α_s by a factor of two. Moreover, $(\Delta M)_\psi\approx(\Delta M)_\gamma$ and hence the theory predicts that $\varkappa_\Upsilon\sim 0.5\,\varkappa_\psi$ and $\delta_\Upsilon\sim 0.5\,\delta_\psi$.

This prediction probing rather subtle aspects of QCD has been confirmed experimentally.

4. CONCLUSIONS

In this paper I have given a pedagogical review of the anomalies and their role in quantum chromodynamics. I had no intention to discuss numerous new developments of the issue, growing up rapidly especially in connection with string theory. My task was to explain the simple physical meaning of the phenomenon and to destroy a mysterious attitude to the subject that emerged partially due to the successful and deep mathematization. I wanted also to consider possible applications in hadron physics. Actually all basic applications worked out in the literature are presented here except a single topic, namely, the so-called U(1) problem (the η' problem; for some marginal remarks see Sec. 3.2). The interested reader is referred to the excellent reviews of Refs. 44 and 47.

[1)] To be published as a chapter in the volume "Vacuum structure and QCD sum rules."

[2)] Let us emphasize that the assertion about the absence of anomalies in the flavor-octet current $j^a_{\mu V}, j^a_{\mu A}$ refers to pure quantum chromodynamics. Including the photons which interact with the quarks according to the standard rules results in the photon anomaly in $j^3_{\mu A}$. Further details are given in Sec. 2.4.

[3)] The arguments given below simplify the actual situation leaving aside some subtleties. In particular, in discussing the "baryon" scenario we consider only the baryon octet (2.29). Thus, it is tacitly assumed that baryons with other quantum numbers, e.g., $J^P=1/2^-$, do not contribute to the pole part of the amplitude $\langle 0|a^3_\mu|2\gamma\rangle$. Actually, it is rather difficult to prove that the combined contribution of all other baryons can not be equal to (2.28)—it turns out necessary to invoke additional assumptions (see in this connection Refs. 16, 17). Therefore although the main line of reasoning demonstrating the spontaneous breaking of the $SU(3)_A$ symmetry is reproduced correctly I would not like to give the impression that the derivation of the spontaneous chiral symmetry breaking is completely trivial and stems almost from nothing.

[4)] In QCD one can rigorously show that the spontaneous breaking of the vector $SU(3)_n$ symmetry is impossible.[17]

[5)] Relations stemming from "pure" PCAC are not discussed here. The corresponding low-energy results however, are widely used as intermediate building blocks. In particular, Subsec. 3.2 contains, in essence, the derivation of the old Glashow-Weinberg result that relates the correlation function of two pseudoscalar quark densities at zero momentum to the vacuum expectation value $\langle\bar qq\rangle$.

[6)] The canonical energy-momentum tensor stemming from Eq. (3.12) has the trace which does not coincide with (3.9). The reason is that the expression for the energy-momentum tensor of the scalar field is not unique. In order to reproduce the condition $\theta_{\mu\mu}=\sigma$ an extra term should be added to the canonical energy-momentum tensor, namely, $\{-(8\varepsilon_{\rm vac}/3m^2)(\Box g_{\mu\nu}-\partial_\mu\partial_\nu)\exp\chi\}$. This extra term is symmetric and is conserved by itself.

[7)] Eq. (3.30) is found in Ref. 24 without explicit derivation. The derivation proving the absence of ambiguities in the picnic energy-momentum tensor is worked out in Ref. 25.

[8)] The distance between $\bar Q$ and Q in the J/ψ and γ systems is noticeably less than 1 Fermi, the characteristic distance for conventional old hadrons.

[1] S. B. Treiman, R. Jackiw, D. Gross, *Lectures on Current Algebra and its Applications*, Princeton University Press, Princeton, 1972.
[2] L. D. Faddeev, Phys. Lett. **B145**, 81 (1984); L. D. Faddeev and S. L. Shatashvili, Phys. Lett. **B167**, 225 (1986); R. Jackiw and R. Rajaraman, Phys. Rev. Lett. **54**, 1219 (1985); R. Rajaraman, Phys. Lett. **B154**, 305 (1985); **B162**, 148 (1985); R. Rajaraman and J. Lott, Phys. Lett. **B165**, 321 (1985); S. L. Shatashvili, Teor. Mat. Fiz. **71**, 40 (1987) [Theor. Math. Phys. (USSR) **71**, 366 (1987)].
[3] M. Green and J. Schwarz, Phys. Lett. **B149**, 117 (1984).
[4] A. Yu. Morozov, Usp. Fiz. Nauk **150**, 337 (1986) [Sov. Phys. Usp. **29**, 993 (1986)].
[5] A. Casher, J. Kogut, and L. Susskind, Phys. Rev. Lett. **31**, 792 (1973).
[6] G. Kogut and L. Susskind, Phys. Rev. **D11**, 3594 (1975).
[7] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Fortschr. Phys. **32**, 585 (1984).
[8] S. L. Adler and W. A. Bardeen, Phys. Rev. **182**, 1517 (1969).
[9] J. S. Collins, L. Duncan, and S. D. Joglekar, Phys. Rev. **D16**, 438 (1977).
[10] M. A. Shifman and A. I. Vainshtein, Nucl. Phys. **B277**, 456 (1986).
[11] S. Ferrara and B. Zumino, Nucl. Phys. **B87**, 207 (1975).
[12] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Phys. Lett. **B57**, 169 (1985).
[13] K. Wilson and J. Kogut, Phys. Rep. **12**, 75 (1974).
[14] A. D. Dolgov and V. I. Zakharov, Nucl. Phys. **B12**, 68 (1971).
[15] G. 't Hooft, in: *Recent Developments in Gauge Theories*, Ed. by G. 't Hooft *et al.*, Plenum Press, N. Y., 1980, p. 241.
[16] S. Coleman and E. Witten, Phys. Rev. Lett. **45**, 100 (1980).
[17] S. Vafa and E. Witten, Nucl. Phys. **B234**, 173 (1984).
[18] A. I. Vaĭnshtein and V. I. Zakharov, Usp. Fiz. Nauk **100**, 225 (1970) [Sov. Phys. Usp. **13**, 73 (1971)].
[19] Y. Nambu, in: *Preludes in Theoretical Physics*, Ed. A.de Shalit, North Holland, Amsterdam, 1966; H. Fritzsch and M. Gell-Mann, *Proc. 16 Int. Conf. on High Energy Physics*, Batavia, 1972, V.2, p. 135; H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. **B47**, 365 (1973); S. Weinberg, Phys. Rev. Lett. **31**, 494 (1973).
[20] H. Leutwyler, Nucl. Phys. **B76**, 413 (1974); S. Weinberg, in: *Festschrift for I. I. Rabi*, Ed. by L. Motz, New York Academy of Sciences, N. Y., 1978.
[21] E. M. Lifshitz and L. P. Pitaevskii, *Relativistic Quantum Theory*, Part 2, Pergamon Press, Oxford, 1974 [Russ. original, Nauka, M., 1971, §141,142]; A. I. Akhiezer, V. B. Berestetskiĭ, *Quantum Electrodynamics*, Interscience, N. Y., 1965 [Russ. original, Nauka, M., 1964 and 1969, §40.2.
[22] R. Crewther, Phys. Lett. **B70**, 349 (1977).
[23] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B191**, 301 (1981).
[24] M. B. Voloshin and V. I. Zakharov, Phys. Rev. Lett. **45**, 688 (1980).
[25] V. A. Novikov and M. A. Shifman, Z. Phys. **C8**, 43 (1981).
[26] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B166**, 493 (1980).
[27] A. A. Migdal and M. A. Shifman, Phys. Lett. **B114**, 445 (1982).
[28] J. Ellis and J. Lanik, Phys. Lett. **B150**, 289 (1987).
[29] G. 't Hooft, Nucl. Phys. **B72**, 461 (1974).
[30] E. Witten, Nucl. Phys. **B160**, 57 (1979).
[31] S. Weinberg, Phys. Rev. **D11**, 3583 (1975).
[32] J. F. Kim, Phys. Rev. Lett. **43**, 103 (1979).
[33] M. B. Voloshin and A. D. Dolgov, Yad. Fiz. **35**, 213 (1982) [Sov. J. Nucl. Phys. **35**, 120 (1982)].
[34] D. Gross, S. B. Treiman, and F. Wilczek, Phys. Rev. **D19**, 2188 (1979).
[35] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B165**, 55 (1980).
[36] B. L. Ioffe and M. A. Shifman, Phys. Lett. **B95**, 99 (1980).

[37]M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Phys. Lett. **B78**, 443 (1978).
[38]L. B. Okun, Leptons and Quarks, North Holland, Amsterdam, 1983, chapter 24 [Russ. original, Nauka, M., 1981].
[39]T. Appelqist and J. Carrazone, Phys. Rev. **D11**, 2856 (1975).
[40]M. B. Voloshin, Yad. Fiz. **44**, 738 (1986) [Sov. J. Nucl. Phys. **44**, 478 (1986)].
[41]K. Gottfried, Phys. Rev. Lett. **40**, 598 (1978).
[42]M. B. Voloshin, Nucl. Phys. **B154**, 365 (1979); M. Peskin, Nucl. Phys. **B156**, 365 (1979).
[43]T. M. Yan, Phys. Rev. **D22**, 1652 (1980).
[44]G. 't Hooft, Phys. Rep. **142**, 357 (1986).
[45]V. N. Gribov, Preprint KFKI-1981-66, Budapest, 1981.
[46]Ya. Ya. Balitskii, D. I. D'yakonov, and A. V. Yung, Yad. Fiz. **35**, 1300 (1982) [Sov. J. Nucl. Phys. **35**, 761 (1982)]; Z. Phys. **C33**, 265 (1986).
[47]D. D'yakonov and M. Eides, Proc. XVI Winter School Phys. LINP, Leningrad, 1981, p. 123; D. D'yakonov, Proc. XXI Winter School Phys. LINP, Leningrad, 1986, p. 3.
[48]H. Leutwyler and M. Shifman, Bern University, Preprint BUTP-89/2.

Translated (and partially revised) by the author.