Measure of strong CP violation

Zheng Huang
Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6
(Received 17 September 1992)

We investigate a controversial issue on the measure of CP violation in strong interactions. In the presence of nontrivial topological gauge configurations, the θ term in QCD has a profound effect: it breaks the CP symmetry. The CP-violating amplitude is shown to be determined by the vacuum tunneling process, where the semiclassical method makes most sense. We discuss a long-standing dispute on whether or not the instanton dynamics satisfies the anomalous Ward identity (AWI). The strong CP-violation measure, when complying with the vacuum alignment, is proportional to the topological susceptibility. To solve the IR divergence problem of the instanton computation, we present a "classically gauged" Georgi-Manohar model and derive an effective potential which uniquely determines an explicit $U(1)_A$ symmetry breaking sector. The CP-violation effects are analyzed in this model. It is shown that the strong CP problem and the U(1) problem are closely related. Some possible solutions to both problems are also discussed with new insights.

PACS number(s): 11.30.Er, 11.15.Kc, 12.38.Aw, 12.40.Aa

I. INTRODUCTION

The discovery of instantons [1] has been associated with some of the most interesting developments in strong interaction theory. It has led to a resolution [2] of the long-standing U(1) problem [3], and also pointed to the existence in QCD [4] of vacuum tunneling and of a vacuum angle θ, which combined with the phase of the determinant of the quark mass matrix, signals the CP violation in strong interactions. The difficulty in understanding the very different hierarchies of the strong CP violation and weak CP violation in the standard model has been targeted as the so-called strong CP problem (for a review, see Ref. [5]).

The theoretical understanding of weak CP violation is well established in the framework of the Kobayashi-Maskawa mechanism [6] in spite of the challenge in higher-precision experimental measurements. It has been shown [7] that the determinant of the commutator of the up-type and down-type quark mass matrices $[M^u, M^d] \equiv iC$ given by

$$\det C = -2\mathcal{J}_{\rm weak}(m_t - m_c)(m_c - m_u)(m_u - m_t) \times(m_b - m_s)(m_s - m_d)(m_d - m_b), \quad (1.1)$$

where

$$\mathcal{J}_{\rm weak} \equiv \sin^2\theta_1 \sin\theta_2 \sin\theta_3 \cos\theta_1 \cos\theta_2 \cos\theta_3 \sin\delta \quad (1.2)$$

is the unique measure of the weak CP violation. All CP-violating effects in weak interactions must be proportional to $\det C$. Even though the CP-violating phase $\sin\delta$ can be of order 1, the physical amplitude is naturally suppressed by the product of Cabibbo mixing angles.

However, the measure of CP violation in QCD, which we shall denote as $\mathcal{J}_{\rm strong}$, is not so clear. It has long been realized that $\theta_{\rm QCD}$ and phases of quark masses are not independent parameters in QCD. In the presence of the chiral anomaly [8], they are related through the chiral transformations of quark fields. Thus $\mathcal{J}_{\rm strong}$ must be proportional to a combination

$$\bar\theta = \theta_{\rm QCD} + \arg\det M, \quad (1.3)$$

which is invariant under chiral rotations. It is well known that if one of the quarks is massless, $\bar\theta$ can take an arbitrary value since one can make arbitrary rotations on the chiral field. This suggests that the $\bar\theta$ dependence of $\mathcal{J}_{\rm strong}$ disappears in the chiral limit. Thus in the case of $L=2$ where L is the number of light quarks, $\mathcal{J}_{\rm strong}$ has the form

$$\mathcal{J}_{\rm strong} = m_u m_d K \sin\bar\theta, \quad (1.4)$$

where we have written $\sin\bar\theta$ instead of $\bar\theta$ to take care of the periodicity in $\bar\theta$. Is there any other common factor that we can extract from strong CP effects or is K in (1.4) only a kinematical factor which varies with different physical processes?

To answer the question, we need to know whether there is any other condition under which the strong CP violation vanishes. Recently, the reanalysis of strong CP effects has shed some light on this issue. Several authors [9] have pointed out, in studying an effective Lagrangian, that the strong CP violation should vanish if the chiral anomaly is absent. We regard their work as constructive and enlightening. However, how to appreciate such a feature in QCD with quarks is not apparent in their approaches. In QCD theory, indeed, if the chiral anomaly is absent, the phases of quark masses can be rotated away without changing the θ term. But it is not clear why $\theta_{\rm QCD}$ itself does not lead to CP violation in strong interactions. In addition, the presence of the chiral anomaly in a gauge theory may not be directly related to CP violation. One example is QED. It is well known that QED is a CP-conserving theory even if it is chirally anomalous, and, in principle, could have a θ term and a complex electron mass term.

In this paper, however, we show that the measure of strong CP violation does acquire a factor referred to as the measure of the nontriviality of the non-Abelian gauge vacuum. It is simply due to the fact that the θ term is a total divergence whose integration over space-time yields a surface term. It can be dropped unless there are nontrivial gauge configurations at the boundary. K in (1.4) will be shown to be the vacuum tunneling amplitude between different vacua characterized by the winding numbers

$$\nu = \int d^4x F\tilde{F} \equiv \frac{g^2}{32\pi^2}\int d^4x F_{\mu\nu}\tilde{F_{\mu\nu}}, \qquad (1.5)$$

where a semiclassical method makes the most sense to deal with it. To probe the property of the K factor, we proceed to consider a classically gauged linear σ model. A derivation of a $U(1)_A$ sector of the model can be made by taking into account the fermion zero modes in the instanton fields. Contrary to the conventional result [10, 11] where K has a singularity in quark masses such that $\mathcal{J}_{\rm strong}$ is a linear function of the quark mass, our model clearly shows that K is to be explained as the mass difference between the U(1) particle and pions. Thus, $\mathcal{J}_{\rm strong}$ has the form

$$\mathcal{J}_{\rm strong} = m_u m_d(m_\eta^2 - m_\pi^2)\sin\bar{\theta}. \qquad (1.6)$$

In the context of the effective model, the strong CP effects can be explicitly calculated and various solutions to the strong CP problem will be discussed with new insights.

The paper is organized as follows. In Sec. II, we discuss a long-standing problem raised by Crewther [10, 12] on whether or not the instanton is consistent with the anomalous Ward identity (AWI). We find that the AWI does not put any constraint on the topological susceptibility $\langle\langle\nu^2\rangle\rangle$ in QCD. The AWI is automatically satisfied by instanton dynamics if the singularity in the chiral limit of some fermionic operator is taken care of. Section III deals with an instanton computation of $\langle\langle\nu^2\rangle\rangle$ in the dilute gas approximation. The vacuum alignment equations of the quark condensates are derived based on the path integral formalism. Upon making alignment among strong CP phases, we rederive an effective CP-violating Lagrangian. In Sec. IV we present a classically gauged linear σ model. In the semiclassical approximation, the instanton fields are integrated out. An effective one-loop potential is obtained by integrating over fermions in the instanton background where the fermion zero modes are essential to yield an explicit $U(1)_A$ symmetry breaking. The strong CP effects and the U(1) particle mass are calculated in the model. Section V is devoted to discussions on various possible solutions to the strong CP problem. Section VI is reserved for conclusions.

II. DOES THE INSTANTON SATISFY THE AWI? THE TOPOLOGICAL SUSCEPTIBILITY $\langle\langle\nu^2\rangle\rangle$

Let us leave our discussion on $\mathcal{J}_{\rm strong}$ aside for the moment and turn to a problem which turns out to be the key to understanding both strong CP violation and the U(1) problem. It has long been pointed out that instanton physics, in some ways, suffers from difficulties. It is well known that the integration over the instanton size is infrared divergent. It is further argued by Witten [13] that the semiclassical method based on the instanton solution of the Yang-Mills equation is in conflict with the most successful idea of $\frac{1}{N_c}$ expansion in QCD. The reason is that instanton effects are of order $e^{-\frac{1}{g^2}}$ or e^{-N_c} (for g^2 of order $\frac{1}{N_c}$ in the large N_c limit), which is smaller than any finite power of $\frac{1}{N_c}$ obtained by summing Feynman diagrams. These problems, as they stand now, indeed reflect various defects in the instanton calculation (we will come back to these points in later sections).

However, there was another type of objection initiated by Crewther [10] followed by others [12], which would be even more serious if they were correct. For many years Crewther has emphasized that the breakdown of $U(1)_A$ symmetry by the chiral anomaly and the instanton is related to the breakdown of the $SU(L)\times SU(L)$ symmetry. This relation is represented by the so-called anomalous Ward identity. He claimed that instanton dynamics failed to satisfy the AWI and one would still expect the unwanted $U(1)_A$ Goldstone boson. He further showed that the topological susceptibility defined as

$$\langle\langle\nu^2\rangle\rangle = \int d^4x \langle T\, iF\tilde{F}(x)\, iF\tilde{F}(0)\rangle \qquad (2.1)$$

must be equal to $m\langle\bar{\psi}\psi\rangle$ in order to satisfy the AWI; m is the quark mass. We have assumed that all quarks have equal masses. As we shall see in Sec. III, $\langle\langle\nu^2\rangle\rangle$ is to be identified as the measure of strong CP violation. If Crewther is right, it would seem that strong CP has no direct relation to the topological vacuum structure.

To see where the problem lies, we carefully follow a path integral derivation of the AWI. Consider a fermion bilinear operator $\bar{\psi}_L\psi_R$ with chirality $+2$ (a sum over flavor indices is understood). Its vacuum expectation value (VEV) is formally given as

$$\begin{aligned}\langle\bar{\psi}_L\psi_R\rangle &= \frac{1}{V}\left\langle\int d^4x\,\bar{\psi}_L\psi_R(x)\right\rangle \\ &= \frac{1}{V}\frac{1}{Z}\int\mathcal{D}(A,\bar{\psi},\psi)\int d^4x\,\bar{\psi}_L\psi_R(x)e^{-S[A,\bar{\psi},\psi]}\end{aligned} \qquad (2.2)$$

where the QCD action in Euclidean space is

$$S[A,\bar{\psi},\psi] = \int d^4x\,(\bar{\psi}\,\not{D}\psi + m\bar{\psi}\psi + \tfrac{1}{4}F^2 - i\theta F\tilde{F}) \qquad (2.3)$$

and Z is the normalization factor, V is the volume of space-time. Under an infinitesimal $U(1)_A$ transformation,

$$\psi_R \to e^{i}\alpha(x)\psi_R, \quad \psi_L \to e^{-i}\alpha(x)\psi_L, \qquad (2.4)$$

the measure $\mathcal{D}(A,\bar{\psi},\psi)$ will change because of the chiral anomaly. However, the integral (2.2) will not change since (2.4) is only a matter of changing integration variables. Equation (2.2) then becomes

$$\langle\bar{\psi}_L\psi_R\rangle = \frac{1}{VZ}\int\mathcal{D}(A,\bar{\psi},\psi)\int d^4x\, e^{2i\alpha(x)}\bar{\psi}_L\psi_R(x)\exp\left(-S[A,\bar{\psi},\psi]+i\alpha(x)\int d^4x[\partial_\mu J^5_\mu - 2m\bar{\psi}\gamma_5\psi - 2L\,F\tilde{F}]\right), \tag{2.5}$$

where the $\mathrm{U}(1)_A$ current is $J^5_\mu = \bar{\psi}\gamma_\mu\gamma_5\psi$ and L is the number of light quarks. The independence of $\langle\bar{\psi}_L\psi_R\rangle$ on $\alpha(x)$ implies the vanishing of its first derivative which yields the AWI:

$$\int d^4x\,\partial_\mu\langle T\,J^5_\mu(x)\,\bar{\psi}_L\psi_R(0)\rangle = 2m\int d^4x\langle T\,\bar{\psi}i\gamma_5\psi(x)\,\bar{\psi}_L\psi_R(0)\rangle + 2L\int d^4x\langle T\,iF\tilde{F}(x)\,\bar{\psi}_L\psi_R(0)\rangle - 2i\langle\bar{\psi}_L\psi_R\rangle. \tag{2.6}$$

Crewther's arguments go as follows. If there is no $\mathrm{U}(1)_A$ Goldstone boson coupling to J^5_μ, the LHS of Eq. (2.6) vanishes. In the chiral limit, the first term on the RHS would vanish too. Thus one has, when $m \to 0$,

$$L\int d^4x\langle T\,F\tilde{F}(x)\,\bar{\psi}_L\psi_R(0)\rangle = \langle\bar{\psi}_L\psi_R\rangle. \tag{2.7}$$

The instanton dynamics assumes that the integration over the gauge field is separated into a sum over gauge configurations characterized by the integer winding number ν in (1.5), i.e., $\int[dA] = \sum_\nu\int[dA]_\nu$ and $\langle\bar{\psi}_L\psi_R\rangle = \sum_\nu\int\langle\bar{\psi}_L\psi_R\rangle_\nu$. Equation (2.7) would then imply

$$(L\nu - 1)\langle\bar{\psi}_L\psi_R\rangle_\nu = 0. \tag{2.8}$$

By assuming the spontaneous chiral symmetry breaking caused by $\langle\bar{\psi}_L\psi_R\rangle \neq 0$, (2.8) cannot be satisfied if ν is an integer. Moreover, by noting that

$$\frac{d\langle\bar{\psi}_L\psi_R\rangle}{d\theta} = i\int d^4x\langle T\,F\tilde{F}(x)\,\bar{\psi}_L\psi_R(0)\rangle \tag{2.9}$$

one obtains

$$\left(-i\frac{d}{d\theta} - \frac{1}{L}\right)\langle\bar{\psi}_L\psi_R\rangle = 0 \Rightarrow \langle\bar{\psi}_L\psi_R\rangle_\theta = \langle\bar{\psi}_L\psi_R\rangle_{\theta=0}\,e^{i\frac{\theta}{L}}, \tag{2.10}$$

which is unacceptable because the θ dependence of $\langle\bar{\psi}_L\psi_R\rangle$ would have a wrong periodicity $2\pi L$. Along the same line, one could derive the AWI for operator $\bar{\psi}_R\psi_L$ and $F\tilde{F}$ and combine them with (2.6) to obtain

$$\langle\langle\nu^2\rangle\rangle = \frac{m^2}{L^2}\int d^4x\langle T\,\bar{\psi}i\gamma_5\psi(x)\;\bar{\psi}i\gamma_5\psi(0)\rangle + \frac{m}{L^2}\langle\bar{\psi}\psi\rangle. \tag{2.11}$$

By noting that the first term on the RHS of (2.11) is of order $O(m^2)$, one would conclude that $\langle\langle\nu^2\rangle\rangle$ was a linear function of m, which, again, contradicts the instanton computation.

We argue, however, that all these inconsistencies arise from dropping the first term on the RHS of (2.6) in the chiral limit or treating it as a higher-order term. The $\mathrm{U}(1)_A$ fermion operator $\bar{\psi}i\gamma_5\psi$, when the fermion fields are integrated out *first* as they should be, may reveal a $\frac{1}{m}$ singularity in certain gauge configurations. To see this, we first calculate the VEV of $\bar{\psi}i\gamma_5\psi$ in a fixed background field A_μ. Upon the fermion integration, one has

$$\langle\bar{\psi}i\gamma_5\psi\rangle^A = \mathrm{Tr}\frac{i\gamma_5}{\not{D}(A)+m} = \frac{1}{m}T(m^2), \tag{2.12}$$

where

$$T(m^2) = \mathrm{Tr}\frac{i\gamma_5 m^2}{-\not{D}^2+m^2} = \mathrm{Tr}\frac{i\gamma_5 m^2}{-D^2+\frac{1}{2}g\sigma_{\mu\nu}F_{\mu\nu}+m^2}. \tag{2.13}$$

It is easy to check that $\frac{d}{dm^2}T(m^2) \equiv 0$, i.e., $T(m^2)$ is independent of m^2. Thus it can be calculated in the limit $m^2 \to \infty$ [14],

$$\lim_{m^2\to\infty} T(m^2) = -iL\int d^4x\;\mathrm{Tr}\gamma_5(\tfrac{1}{2}\sigma_{\mu\nu}F_{\mu\nu})^2\int\frac{d^4p}{(2\pi)^4}\frac{m^2}{(p^2+m^2)^3} = iL\,F\tilde{F} \tag{2.14}$$

and therefore

$$\langle\bar{\psi}i\gamma_5\psi\rangle^A = -iL\frac{F\tilde{F}}{m}. \tag{2.15}$$

It has a pole at $m = 0$. It is clear that $m\langle\bar{\psi}i\gamma_5\psi\rangle^A$ may be finite in the limit $m \to 0$ if $F\tilde{F}$ is nontrivial. Performing the fermion integration for the first term on the RHS of (2.6), we obtain

$$m\int d^4x\langle T\,\bar{\psi}i\gamma_5\psi(x)\,\bar{\psi}_L\psi_R(0)\rangle = \int d^4x\left\langle T\,\mathrm{Tr}\left(\frac{im\gamma_5}{\not{D}+m}\right)(x)\,\mathrm{Tr}\left(\frac{1+\gamma_5}{2(\not{D}+m)}\right)(0)\right\rangle - \left\langle\mathrm{Tr}\left(\frac{im\gamma_5}{\not{D}+m}\frac{1+\gamma_5}{2}\frac{1}{\not{D}+m}\right)\right\rangle \tag{2.16}$$

$$= -L\int d^4x\left\langle T\,iF\tilde{F}(x)\;\mathrm{Tr}\left(\frac{1+\gamma_5}{2}\frac{1}{\not{D}+m}\right)(0)\right\rangle - i\left\langle\mathrm{Tr}\left(\frac{1}{2}(1+\gamma_5)\frac{m}{-\not{D}^2+m^2}\right)\right\rangle. \tag{2.17}$$

Identifying the second term in (2.17) with $\langle\bar{\psi}_L\psi_R\rangle$, we find that the RHS of (2.6) vanishes identically for any m. This is not surprising since if we had considered a *global* $U(1)_A$ transformation instead of a local one in (2.4) at the beginning, we would have come up with the same conclusion immediately. Similarly, (2.11) is an identity to be satisfied (trivially) by any dynamics which respects the basic rule of the fermion quantization. [Of course it should also respect the anomaly relation. If there were no anomaly, the second term on the RHS of (2.6) would be absent. The cancellation would be incomplete indicating the existence of a massless excitation coupling with J^5_μ. Thus the chiral anomaly is essential to solve the U(1) problem.]

There is a delicate problem about taking the chiral limit. One may ask what if the quark mass term is simply absent in the Lagrangian in the first place. Crewther's problem seems to come back if the first term on the RHS of (2.6) is not present. Actually this is where the puzzle comes about. In this case, however, a nonvanishing value of the quark condensate is not well defined. It relates to a general feature of the spontaneous symmetry breaking mechanism. For example, in the ϕ^4 theory with spontaneous breaking of the reflection symmetry ($\phi \to -\phi$), the VEV of ϕ is calculated

$$\langle\phi\rangle = \frac{1}{Z}\int d\phi\,\phi\, e^{-\int d^4x(\partial_\mu\phi)^2+\frac{\lambda}{4}(\phi^2-v^2)^2}. \qquad (2.18)$$

Since the action is perfectly reflection symmetric and ϕ is an odd operator under reflection, we have $\langle\phi\rangle \equiv 0$. Mathematically this is true because of the equal weight of the degenerate vacua. But what is of physical interest is a situation where one of the degenerate vacua is *chosen* as the ground state. The way to do it is to introduce a source term $\int d^4x\, J\phi$ into the action which breaks the symmetry explicitly. The degeneracy of the vacua in the absence of the source implies that $\langle\phi\rangle_J$ is a multivalued function of J at $J=0$. The VEV's of ϕ crucially depend on the way that J tends to zero. In particular, $\langle\phi\rangle_{J\to 0^+} = -\langle\phi\rangle_{J\to 0^-} \neq 0$.

The same procedure should follow for the spontaneous chiral symmetry breaking in QCD. In order to define the quark condensate $\langle\bar{\psi}_L\psi_R\rangle$, one ought to add the source term $\int d^4x\, J\bar{\psi}_l\psi_R(x)$ to the action. Then a $U(1)_A$ transformation changes the source term as well

$$\int d^4x\, J\bar{\psi}_L\psi_R \to \int d^4x\, J e^{2i\alpha}\bar{\psi}_L\psi_R. \qquad (2.19)$$

We also need to take this change into account because $\langle\bar{\psi}_L\psi_R\rangle$ defined by the way that $J\to 0$ would be different from the one defined by $Je^{2i\alpha}\to 0$. By differentiating $\langle\bar{\psi}_L\psi_R\rangle$ with respect to α we obtain an equation exactly the same as (2.6) except that m is replaced by J. For the same reason as we have discussed, the RHS of the equation is identically zero for any value J (even in the limit $J\to 0$). There is no $U(1)_A$ Goldstone boson, and, in general, (2.7), (2.8), and (2.10) do not hold.

We have shown that the AWI for the isosinglet current J^5_μ is trivially satisfied by QCD dynamics including the axial anomaly. Equation (2.11) is an identity satisfied by any dynamics if the singularity of the singlet operator $\bar{\psi}i\gamma_5\psi$ in the zero-mass limit is appropriately handled. It does not put any constraint on how the topological susceptibility $\langle\langle\nu^2\rangle\rangle$ should behave as a function of the quark mass. Thus, it does not, from the context of the field theory, rule out the instanton computation. However, this should not be confused with the case of the AWI's for nonsinglet currents where the assumption on the lowest-lying resonances have to be made. For a nonsinglet axial-vector current $J^a_\mu = \bar{\psi}\gamma_\mu\gamma_5\frac{\lambda^a}{2}\psi$ [λ^a's are generators of SU(L), $a=1,\ldots,L^2-1$], the corresponding AWI reads

$$m^2\int d^4x\left\langle T\,\bar{\psi}i\gamma_5\frac{\lambda_a}{2}\psi(x)\;\bar{\psi}i\gamma_5\frac{\lambda_b}{2}\psi(0)\right\rangle - \delta^{ab}\frac{m}{L}\langle\bar{\psi}\psi\rangle = 0. \qquad (2.20)$$

It can be readily checked by integrating the fermion fields that (2.20) is satisfied in QCD. Unlike the singlet current in (2.12)

$$\left\langle\bar{\psi}i\gamma_5\frac{\lambda_a}{2}\psi\right\rangle^A = \mathrm{Tr}\frac{\lambda_a}{2}\frac{i\gamma_5}{\not{D}+m} = 0 \qquad (2.21)$$

because λ^a's are traceless. *Assuming* that pions are the lowest-lying resonances which dominate the correlation function, one obtains

$$m^2\int d^4x\left\langle T\,\bar{\psi}i\gamma_5\frac{\lambda_a}{2}\psi(x)\;\bar{\psi}i\gamma_5\frac{\lambda_a}{2}\psi(0)\right\rangle_{\rm res} = F_\pi^2 m_\pi^2\delta^{ab} \qquad (2.22)$$

leading to $F_\pi^2 m_\pi^2 = -\frac{1}{L}m\langle\bar{\psi}\psi\rangle$. Can we do the same analysis for the correlation function of the singlet operator

$$m^2\int d^4x\langle T\,\bar{\psi}i\gamma_5\psi(x)\;\bar{\psi}i\gamma_5\psi(0)\rangle_{\rm res} \qquad (2.23)$$

such that we may get a phenomenological value for $\langle\langle\nu^2\rangle\rangle$ from (2.11) without resorting to instanton computations? This turns out to be of some difficulty. For the axial singlet operator, we cannot generally assume pion dominance. In fact, $m\bar{\psi}i\gamma_5\psi$ does not couple to pions because λ^a's commute with the identity [11]. In addition, $\bar{\psi}i\gamma_5\psi$ has pole behavior at $m=0$ whose residue is $F\tilde{F}$. It may couple to a gauge ghost [15] as well as glueballs and the $U(1)_A$ particle. It may also exhibit a nonzero subtraction constant in the spectral dispersion representation [16], which by itself is not surprising in the presence of the anomaly. All these factors may further fall into overlap, causing double counting. They have made an estimation on (2.23) extremely difficult if not impossible.

In summary, the AWI and the low-energy phenomenology may not put a constraint on the topological susceptibility. Therefore, it leaves us the task of calculating $\langle\langle\nu^2\rangle\rangle$ and the measure of strong *CP* violation from instanton dynamics. To avoid the infrared divergence, we further relate $\langle\langle\nu^2\rangle\rangle$ to the $U(1)_A$ particle mass in an effective theory.

III. THE EFFECTIVE *CP*-VIOLATING LAGRANGIAN IN QCD

In Sec. II we have shown that the axial singlet operator $\bar{\psi} i\gamma_5\psi$ is related to $F\tilde{F}$ in a fixed gauge background. When the gauge fields are integrated out, (2.15) becomes a relation on VEV's. It can be easily proven that such a relation is true for each flavor. In general, when the quark mass is complex, one derives

$$-i(m_i e^{i\varphi_i}\langle\bar{\psi}_L^i\psi_R^i\rangle - m_i e^{-i\varphi_i}\langle\bar{\psi}_R^i\psi_L^i\rangle) = -i\left(m_i e^{i\varphi_i}\left\langle \mathrm{Tr}\frac{1}{2}\frac{1+\gamma_5}{\not{D}+m_i e^{i\varphi_i\gamma_5}}\right\rangle - m_i e^{-i\varphi_i}\left\langle \mathrm{Tr}\frac{1}{2}\frac{1-\gamma_5}{\not{D}+m_i e^{i\varphi_i\gamma_5}}\right\rangle\right)$$
$$= \langle iF\tilde{F}\rangle, \tag{3.1}$$

where φ_i is the phase of the ith quark mass ($i = 1, \ldots, L$), no sum over i is understood in (3.1). Now define

$$\langle\bar{\psi}_L^i\psi_R^i\rangle \equiv -\frac{C_i}{2}e^{i\beta_i}, \quad \langle\bar{\psi}_R^i\psi_L^i\rangle \equiv -\frac{C_i}{2}e^{-i\beta_i} \tag{3.2}$$

or

$$\langle\bar{\psi}^i\psi^i\rangle \equiv -C_i\cos\beta_i, \quad \langle\bar{\psi}^i i\gamma_5\psi^i\rangle \equiv C_i\sin\beta_i, \tag{3.3}$$

where $C_i > 0$ and β_i is the phase of the ith quark condensate. Equation (3.1) yields

$$\langle iF\tilde{F}\rangle = -m_i C_i \sin(\varphi_i + \beta_i) \quad (i = 1, 2, \ldots, L), \tag{3.4}$$

which is to be referred to as the vacuum alignment equation (VAE) [17]. It can also be derived directly by taking vacuum expectation values on both sides of the anomaly relation [18]. Equation (3.4) means that if the first moment of the topological charge is nonzero in the presence of instanton, the quark condensate develops a phase β_i different from $-\varphi_i$. If the phase of the fermion mass φ_i is zero as it can always be made so by making a chiral rotation, the fermion condensate has a nontrivial phase $\beta_i \neq 0$, i.e., develops an imaginary part which is determined by the topological structure of the theory. This of course would not happen in a theory like QED where only the trivial topological configuration exists. We shall see that it is the combination $\varphi_i + \beta_i$ that determines the *CP*-violating amplitude in strong interactions.

$\langle F\tilde{F}\rangle$ can be calculated from instanton dynamics in the dilute gas approximation (DGA) [19]. The vacuum-to-vacuum amplitude in the presence of the θ term is given by

$$Z(\bar{\theta}) = \sum_{\nu=0,\pm1,\ldots}^{\infty} \int \mathcal{D}(A,\bar{\psi},\psi)e^{i\bar{\theta}\nu} \times e^{-\int d^4x \sum_i \bar{\psi}^i(\not{D}+m_i)\psi^i + \frac{1}{4}F^2}, \tag{3.5}$$

where we have not explicitly included the gauge fixing and the ghost terms. Inclusion of them must be understood when a practical computation is performed. The phase of the quark masses have been rotated away and $\bar{\theta} = \theta_{\rm QCD} + \sum_i \varphi_i$. In the DGA,

$$Z(\bar{\theta}) = \sum_{n_+=0}^{\infty}\sum_{n_-=0}^{\infty}\frac{1}{n_+}\frac{1}{n_-}(Z_+)^{n_+}(Z_-)^{n_-} = e^{Z_+ + Z_-}, \tag{3.6}$$

where Z_+ (Z_-) is the single instanton (anti-instanton) amplitude

$$Z_+ = e^{i\bar{\theta}}\int d^4z\frac{d\rho}{\rho^5}C_{N_c}\left(\frac{8\pi^2}{g^2(\rho)}\right)^{2N_c} e^{-\frac{8\pi^2}{g^2(\rho)}}d(M\rho),$$
$$Z_- = Z_+^*, \tag{3.7}$$

with

$$C_{N_c} = \frac{4.6\,\exp(-1.68N_c)}{\pi^2(N_c-1)!(N_c-2)!}.$$

The factor $d(M\rho)$ in (3.8) is connected with the so-called fermion determinant, which introduces important physics. It was first discovered by 't Hooft [20] that there exists a zero mode of the operator $\not{D}$ in the instanton field. Thus we expect $d(M\rho) \propto \det M$ (M is the quark mass matrix). For small quark masses, $d(M\rho)$ is equal to [20, 21]

$$d(M\rho) = \prod_{i=1}^{L} f(m_i\rho),$$
$$f(x) = 1.34x(1 + x^2\ln x + \cdots), \quad x \ll 1. \tag{3.8}$$

Combining (3.8) and (3.8) with (3.6) one obtains

$$Z(\bar{\theta}) = \exp[2V\cos\bar{\theta}\, m_1 m_2\cdots m_L K(L)], \tag{3.9}$$

where $K(L)$ is of dimension $4 - L$:

$$K(L) \cong (1.34)^L \int \frac{d\rho}{\rho^{5-L}}C_{N_c}\left(\frac{8\pi^2}{g^2(\rho)}\right)^{2N_c} e^{-\frac{8\pi^2}{g^2(\rho)}}. \tag{3.10}$$

The first moment $\langle iF\tilde{F}\rangle$ is calculated by taking an average of the topological charge over four-space,

$$\langle iF\tilde{F}\rangle = \frac{1}{V}\left\langle\int d^4x\, iF\tilde{F}\right\rangle = \frac{1}{V}\frac{d}{d\bar{\theta}}\ln Z(\bar{\theta})$$
$$= -2m_u m_d\cdots m_L K(L)\sin\bar{\theta}, \tag{3.11}$$

and the topological susceptibility is equal to

$$\langle\langle\nu^2\rangle\rangle = \frac{1}{V}\frac{d^2}{d\bar{\theta}^2}\ln Z(\bar{\theta})$$
$$= -2m_u m_d\cdots m_L K(L)\cos\bar{\theta}. \tag{3.12}$$

Clearly, when $\bar{\theta}$ is small we have

$$\langle iF\tilde{F}\rangle = \langle\langle \nu^2\rangle\rangle\bar{\theta}. \quad (3.13)$$

The vacuum alignment in QCD can be readily made through the VAE in (3.4). By defining the quark field, one can change the phase of the quark mass φ_i and phase of the quark condensate β_i. However, the $\varphi_i+\beta_i$'s will not change under the redefinition. They are only functions of $\bar{\theta}$ as shown in (3.4). One may choose $\beta_i = 0$ $(i = 1,\ldots,L)$ such that the vacuum is *CP*-conserving:

$$\langle \bar{\psi}^i i\gamma_5\psi^i\rangle = 0 \quad (i = 1,2,\ldots,L). \quad (3.14)$$

Then the phases of the quark masses are no longer arbitrary. They are uniquely determined by the vacuum alignment equation (3.4),

$$\varphi_i = -\frac{\langle\langle\nu^2\rangle\rangle}{m_i C}\bar{\theta} \qquad (i = 1,2,\ldots,L),$$
$$\theta_{\rm QCD} = \bar{\theta} - \sum_i \varphi_i = \left(1 - \sum_i \frac{\langle\langle\nu^2\rangle\rangle}{m_i C}\right)\bar{\theta}, \quad (3.15)$$

where we have assumed φ_i's are small and C_i's are all equal to C. To be aligned with the vacuum, the strong *CP* phase $\bar{\theta}$ must be distributed among the θ term and the quark mass terms according to their determined weights. The effective *CP*-violating part of the QCD Lagrangian reads

$$\mathcal{L}_{CP}^{\beta_i=0} = i\theta_{\rm QCD}F\tilde{F} - \frac{2}{C}m_u m_d\cdots m_L K(L)\bar{\theta}\bar{\psi}i\gamma_5\psi \quad (3.16)$$

with $\theta_{\rm QCD}$ given in (3.15).

It is worth emphasizing that the effective *CP*-violating interactions in (3.16) are only valid in the *CP*-conserving vacuum where β_i's are zero. One can alternatively choose a certain pattern of the phase distribution and ask in what direction the vacuum is to align with it. In general, the vacuum angles are not zero and should be determined by the VAE (3.4). For example, we can choose $\varphi_i = 0$ $(i = 1,\ldots,L)$ such that $\mathcal{L}_{CP}^{\beta_i=0} = i\bar{\theta}F\tilde{F}$. In this case, the vacuum condensates are complex where $\beta_i = -\frac{\langle\langle\nu^2\rangle\rangle}{m_i C}\bar{\theta}$. A physical *CP*-violating amplitude is from both the *CP*-violating part of the Lagrangian and the *CP*-violating part of the quark condensate. A proof of the equivalence of different chiral frames on strong *CP* effects is given in Ref. [22] where it is shown that the vacuum alignment equation (3.4) plays an essential role.

Does the leftover θ term in the effective Lagrangians play any role in computing the strong *CP* effects? So far there have been only two *CP*-violating processes available: $\eta \to 2\pi$ and the electric dipole moment (EDM) of the neutron. The latter process depends on a computation on the effective *CP*-odd π-N coupling [23,24]. Both of them would involve in an evaluation of the commutator $[Q_5^a, F\tilde{F}]$ if the θ term were to contribute

$$\langle \pi^a\pi^b|\theta_{\rm QCD}F\tilde{F}|\eta\rangle = -\frac{i\theta_{\rm QCD}}{F_\pi}\langle\pi^b|[Q_5^a,\, F\tilde{F}]|\eta\rangle,$$
$$\langle \pi^a N|\theta_{\rm QCD}F\tilde{F}|N'\rangle = -\frac{i\theta_{\rm QCD}}{F_\pi}\langle N|[Q_5^a,\, F\tilde{F}]|N'\rangle, \quad (3.17)$$

where we have used the soft-pion theorem. It is obvious that $[Q_5^a, F\tilde{F}] = 0$ since Q_5^a is a nonsinglet charge and thus the canonical commutation relation applies. The θ term in our particular choice of the effective Lagrangian and the vacuum direction can be ignored. However, it is emphasized that this should not be considered as a general statement. The whole point has to do with the vacuum alignment. What really matters is the correlation relation between ϕ_i's and β_i's given by (3.4).

The above statement can be exemplified in the following. For simplicity, let us assume $m_u = m_d = \cdots = m_L = m$ and $L = 3$ where pions and η are all light pseudoscalars and the soft-pion theorem applies. The amplitude of $\eta \to 2\pi$ is readily calculated when the β_i's are zero:

$$\begin{aligned}A(\eta\to 2\pi) &= \langle\pi^0\pi^0|\mathcal{L}_{CP}^{\beta_i=0}|\eta\rangle\\ &= \bar{\theta}\left(\frac{-i}{F_\pi}\right)^3 \langle[Q_5^3,[Q_5^3,[Q_5^8,\,\bar{\psi}i\gamma_5\psi]]]\rangle\\ &= \frac{4}{\sqrt{3}}\frac{1}{F_\pi^3}m_u m_d m_s K\bar{\theta}.\end{aligned} \quad (3.18)$$

In deriving (3.18), we have dropped the $F\tilde{F}$ term. In a chiral frame where the ϕ_i's are zero, we can still drop the θ term. But the *CP*-conserving part of the Lagrangian will contribute because the vacuum condensates are *CP* violating:

$$\begin{aligned}A(\eta\to 2\pi) &= -\langle\pi^0\pi^0|m\bar{\psi}\psi|\eta\rangle\\ &= -m\left(\frac{-i}{F_\pi}\right)^3 \langle[Q_5^3,[Q_5^3,[Q_5^8,\,\bar{\psi}\psi]]]\rangle\\ &= -\frac{2}{\sqrt{3}}\frac{1}{F_\pi^3}mC\sin\beta\\ &= \frac{4}{\sqrt{3}}\frac{1}{F_\pi^3}m_u m_d m_s K\bar{\theta},\end{aligned} \quad (3.19)$$

where $\beta_i = -\frac{\langle\langle\nu^2\rangle\rangle}{m_i C}\bar{\theta}$. Both (3.12) and (3.19) yield the same result.

We conclude that the measure of strong *CP* violation is given by the topological susceptibility

$$\mathcal{J}_{\rm strong} = -\frac{1}{2}\langle\langle\nu^2\rangle\rangle\bar{\theta} = m_1 m_2\cdots m_L K(L)\bar{\theta}. \quad (3.20)$$

However, $K(L)$ is still an unknown factor, in addition, the integral in (3.10) is simply divergent for large instanton density. This is the shortcoming of all instanton computations if one uses the dilute gas approximation (DGA). More seriously, as we shall see below, $K(L)$ is to be related to the mass of the $U(1)_A$ particle. If $K(L)$ is of order e^{-N_c} as argued by Witten [13], it would be in conflict with Witten and Veneziano's solution to the $U(1)_A$ problem [15] in which the mass of the $U(1)_A$ particle is of $O(\frac{1}{N_c})$. This suggests that we should not take the expression for $K(L)$ in (3.10) too seriously since it is divergent after all. Furthermore, the DGA may not be valid in the IR region of the QCD theory. It has been suggested in Ref. [25] that the instanton liquid can in principle avoid the IR problem, and gives rise to a description on the $U(1)_A$ particle mass consistent with Witten and Veneziano's scenario. Nevertheless, some main features of the instanton computation do not depend on the detail of the topological configurations. For instance, those mass factors appearing in (3.20) will not change since they are a direct result of the Atiyah-Singer index theorem [26] on the fermion zero modes.

IV. THE EFFECTIVE CHIRAL MODEL

A. The model and the instanton induced quantum corrections

We consider an effective chiral theory where meson degrees of freedom are explicitly introduced. The virtue of the model is that it reflects all flavor symmetries in strong interactions as described by QCD and the mesons as independent field excitations coupled to fermions through Yukawa couplings. Unlike a conventional effective theory [27] in which the nucleons are involved, the model that we will be discussing contains quarks, gluons, and mesons. It is a linear version of the gauged σ model suggested by Georgi and Manohar [28], which describes strong interactions in the intermediate energy region between the scale of chiral symmetry breaking and the scale of quark confinement.

The model reads

$$\mathcal{L} = -\bar{\psi}\,\not{D}\psi - \frac{1}{4}F^2 + i\theta F\tilde{F} - f\bar{\psi}_L\phi\psi_R - f\bar{\psi}_R\phi^\dagger\psi_L$$
$$-\mathrm{Tr}\partial_\mu\phi\partial_\mu\phi^\dagger - V_0(\phi\phi^\dagger) - V_m(\phi,\phi^\dagger), \qquad (4.1)$$

where ϕ is a complex $L\times L$ matrix, $V_0(\phi\phi^\dagger)$ is the most general form of a potential invariant under $U(L)\times U(L)$ (renormalizable):

$$V_0(\phi\phi^\dagger) = -\mu^2\mathrm{Tr}\phi\phi^\dagger + \frac{1}{2}(\lambda_1-\lambda_2)(\mathrm{Tr}\phi\phi^\dagger)^2$$
$$+\lambda_2\mathrm{Tr}(\phi\phi^\dagger)^2 \qquad (4.2)$$

and

$$V_m(\phi,\phi^\dagger) = -\frac{1}{4}me^{i\chi}\mathrm{Tr}\phi - \frac{1}{4}me^{-i\chi}\mathrm{Tr}\phi^\dagger. \qquad (4.3)$$

Equation (4.1) needs some explanations. Under $U(L)_L\times U(L)_R$, the quark field as well as the complex meson field transforms as

$$\psi_L \to U_L\psi_L, \quad \psi_R \to U_R\psi_R,$$
$$\phi \to U_L\phi U_R^\dagger, \quad \phi^\dagger \to U_R\phi^\dagger U_L^\dagger. \qquad (4.4)$$

In the absence of V_m, $\mathcal{L}$ is invariant *classically* under (4.4) but broken down to $SU(L)_L\times SU(L)_R\times U(1)_V$ by the chiral anomaly. V_m, replacing the quark mass (m now is of dimension 3), serves as an explicit symmetry breaking and must be treated as a perturbation. f is the Yukawa coupling, chosen to be real by redefining ϕ. Under $U(1)_A$ transformation

$$\psi_L \to e^{i\omega}\psi_L, \quad \psi_R \to e^{-i\omega}\psi_R,$$
$$\phi \to e^{2i\omega}\phi, \quad \phi^\dagger \to e^{-2i\omega}\phi^\dagger, \qquad (4.5)$$

the θ term and V_m change as $\theta\to\theta-2L\omega$, $\chi\to\chi+2\omega$. But $\bar{\theta}=\theta+L\chi$ remains unchanged. Except for the meson sector, the gauge interaction in (4.1) looks identical to QCD. One may wonder if we are double counting the degrees of freedom. This is explained in [28] by pointing out that these quarks and gluons are not the same as in QCD. In particular, quarks are supposed to acquire *constituent* masses about 360 MeV, which is huge compared to the current mass in QCD. The gauge coupling g_s between quarks and gluons in the effective theory is found to be

$$\alpha_s \cong 0.28, \qquad (4.6)$$

much less than its QCD counterpart. This may explain why the nonrelativistic quark model works since the quarks inside a proton could be treated as weakly interacting objects.

However, the drawback of the model is that it has a very serious U(1) problem. Indeed, if one calculates the physical spectrum from V_0+V_m, one finds L^2 would-be Goldstone modes. In addition, the nontrivial topological structure of the theory has been totally overlooked. The classical excitations such as instantons have not been accounted for in the model, which, according to the original idea of 't Hooft [2], are crucial to solving the U(1) problem.

We therefore consider the quantum correction to the Lagrangian (4.1) in the presence of nontrivial classical gauge fields known as instantons. We argue that the effective gauge coupling α_s in (4.6) is obtained *only if* those classical extrema to the action have been effectively summed over by semiclassical methods. We find that the one-loop quantum fluctuations around instantons lead to a dramatic change on the $U(1)_A$ sector of the model. The U(1) particle acquires an extra mass from the vacuum tunneling effects, which, in turn, result in the so-called strong CP problem.

The effective action of the meson field is calculated as

$$Z = \int\mathcal{D}(\phi,\phi^\dagger)e^{-S_0[\phi,\phi^\dagger]}\int\mathcal{D}(A,\bar{\psi},\psi)e^{-S[\bar{\psi},\psi;A;\phi,\phi^\dagger]},$$
$$= \int\mathcal{D}(\phi,\phi^\dagger)e^{-S_{\rm eff}[\phi,\phi^\dagger]}, \qquad (4.7)$$

where

$$S_{\rm eff}[\phi,\phi^\dagger] = S_0[\phi,\phi^\dagger] + \Delta S[\phi,\phi^\dagger] \qquad (4.8)$$

and the quantum correction is given

$$\Delta S[\phi,\phi^\dagger] = -\ln\int\mathcal{D}(A,\bar{\psi},\psi)e^{-S[\bar{\psi},\psi;A;\phi,\phi^\dagger]}$$
$$\equiv -\ln\tilde{Z}[\phi,\phi^\dagger]. \qquad (4.9)$$

The calculation of $\tilde{Z}[\phi,\phi^\dagger]$ in the instanton background follows the standard derivation of the vacuum-to-vacuum amplitude as illustrated in [20]

$$\tilde{Z}[\phi,\phi^\dagger]=\sum_\nu\int \mathcal{D}A_{\rm cl}\,e^{i\theta\nu-S[A_{\rm cl}]}\times(\mathrm{Det}\,\mathcal{M}_A)^{-1/2}\mathrm{Det}\,\mathcal{M}_\psi\mathrm{Det}\,\mathcal{M}_{gh},\qquad(4.10)$$

where

$$\mathcal{M}_A=-D^2-2F,\quad \mathcal{M}_{gh}=-D^2,\quad \mathcal{M}_\psi=\not{D}+\frac{f}{2}(\phi+\phi^\dagger)+\frac{f}{2}(\phi-\phi^\dagger)\gamma_5.\qquad(4.11)$$

If only the effective potential is of concern, ϕ and $\phi^\dagger$ in $\mathcal{M}_\psi$ are to be taken as constant fields. The fermion determinant, as usual, needs special treatment:

$$\mathrm{Det}\,\mathcal{M}_\psi=\mathrm{Det}^{(0)}\mathcal{M}_\psi\mathrm{Det}'\mathcal{M}_\psi.\qquad(4.12)$$

$\mathrm{Det}^{(0)}$ denotes contributions from the subspace of zero modes of $\not{D}$. In a single instanton field, $\not{D}$ has a zero mode with chirality -1 ($\gamma_5=-1$) [14]. Thus we have

$$\mathrm{Det}^{(0)}\mathcal{M}_\psi=\det\left[\frac{f}{2}(\phi+\phi^\dagger)+\frac{f}{2}(\phi-\phi^\dagger)(-1)\right]=\det(f\phi^\dagger),$$

where det only acts upon flavor indices. The prime in $\mathrm{Det}'\mathcal{M}_\psi$ reminds us of excluding zero modes from the eigenvalue product. Since $[\not{D},\gamma_5]\neq 0$, $\mathcal{M}_\psi$ cannot be diagonalized in the basis of eigenvectors of $\not{D}$. The nonvanishing eigenvalues always appear in pairs, i.e., if $\not{D}\varphi_n=\lambda_n\varphi_n$, where $\lambda_n\neq 0$, then $\not{D}\gamma_5\varphi_n=-\gamma_5\not{D}\varphi_n=-\lambda_n\gamma_5\varphi_n$, namely both λ_n and $-\lambda_n$ are eigenvalues of $\not{D}$. In addition, γ_5 takes φ_n to φ_{-n}. Therefore

$$\begin{aligned}\mathrm{Det}'\mathcal{M}_\psi&=\det\prod_{\lambda_n>0}\begin{pmatrix} i\lambda_n+\frac{f}{2}(\phi+\phi^\dagger) & \frac{f}{2}(\phi-\phi^\dagger)\\ \frac{f}{2}(\phi-\phi^\dagger) & -i\lambda_n+\frac{f}{2}(\phi+\phi^\dagger)\end{pmatrix}\\&=\det\prod_{\lambda_n>0}(\lambda_n^2+f^2\phi\phi^\dagger)\\&=\mathrm{Det}'^{1/2}(-\not{D}^2+f^2\phi\phi^\dagger).\end{aligned}\qquad(4.13)$$

Now we are ready to make the DGA. We need to further assume a weak-field approximation of ϕ and $\phi^\dagger$. This can be justified since ϕ and $\phi^\dagger$ fluctuate about their VEV's, which are about 300 MeV. The large fluctuations are exponentially suppressed by $\exp(-\lambda_1|\phi|^4)$. In the DGA

$$\tilde{Z}[\phi,\phi^\dagger]=\mathrm{Det}^{1/2}(-\partial^2+f^2\phi\phi^\dagger)\exp(\tilde{Z}_++\tilde{Z}_-),\qquad(4.14)$$

where

$$\begin{aligned}\tilde{Z}_+[\phi,\phi^\dagger]&=e^{i\theta}\det(f\phi^\dagger)\int dz\frac{d\rho}{\rho^5}C_{N_c}\left(\frac{8\pi^2}{g^2(\rho)}\right)^{2N_c}e^{-\frac{8\pi^2}{g^2(\rho)}}\det\left[1.34\rho\left(1+f^2\phi\phi^\dagger\ln f^2\phi\phi^\dagger+\cdots\right)\right]\\&\cong VK(L)e^{i\theta}\det(f\phi^\dagger),\\ \tilde{Z}_-[\phi,\phi^\dagger]&=\tilde{Z}_+^\dagger[\phi,\phi^\dagger],\end{aligned}\qquad(4.15)$$

and $K(L)$ is given in (3.10).

Combining (4.14) with (4.9) and noticing that $\ln\mathrm{Det}(-\partial^2+f^2\phi\phi^\dagger)$ contains terms which can be absorbed into the tree-level Lagrangian by redefinition of its bare parameters, we obtain the following effective Lagrangian:

$$\mathcal{L}_{\rm eff}=-\bar\psi\not{D}_s\psi-\frac{1}{4}F_s^2-(f\bar\psi_L\phi\psi_R+\mathrm{H.c.})-\mathrm{Tr}(\partial_\mu\phi\partial_\mu\phi^\dagger)-V_0(\phi\phi^\dagger)-V_m(\phi,\phi^\dagger)-V_k(\phi,\phi^\dagger),\qquad(4.16)$$

where

$$V_k(\phi,\phi^\dagger)=-K(L)f^Le^{i\theta}\det\phi^\dagger-K(L)f^Le^{-i\theta}\det\phi.\qquad(4.17)$$

Several remarks on (4.16) are in order. The presence of V_k in (4.16) is the direct result of fermion zero modes in the instanton field. It is invariant under $\mathrm{SU}(L)_L\times\mathrm{SU}(L)_R\times\mathrm{U}(1)_V$ but not invariant under $\mathrm{U}(1)_A$. Under $\mathrm{U}(1)_A$ rotation (4.5), $e^{i\theta}\det\phi\to e^{i(\theta-2\omega L)}\det\phi$. Thus V_k takes over the role of the θ term and respects the anomaly relation. Again, $\bar\theta=\theta+\chi L$ remains invariant. The prototype of V_k was suggested a long time ago by several authors [29] and rediscussed by 't Hooft [30] in the context of instanton. It is different from the model originally proposed by Di Vecchia [31] and recently analyzed in Ref. [9], although the physical contents of both models may be similar. The gauge interactions between quarks and gluons are still present in (4.16) as required in the nonrelativistic quark model. However, they differ from QCD in that the gauge coupling g_s has a smaller value, and most importantly, the gauge field A_s now possesses a *trivial* topology at infinity. The gauge interaction sector in (4.16) is very analogous to QED: the fermion chiral anomaly still exists, but any θ term $\int d^4x\,\theta F_s\tilde{F}_s$ in the action would be simply a vanishing surface term and can be dropped.

B. U(1) particle mass and strong *CP* violation

We would like to discuss the physical spectrum of the model (4.16) (this part has been worked out in Ref. [30])

and show how the strong CP effects can be calculated effectively. To simplify the problem, we take $L = 2$ and u and d quarks have equal masses. In this case, η is identified as the U(1) particle and there will not be a mixing between π^0 and η.

The complex meson field ϕ contains eight particle excitations σ, η, π_a, and α_a $(a = 1, 2, 3)$:

$$\phi = \frac{1}{2}(\sigma + i\eta) + \frac{1}{2}(\boldsymbol{\alpha} + i\boldsymbol{\pi}) \cdot \boldsymbol{\tau}, \tag{4.18}$$

where $\tau^{1,2,3}$ are the Pauli matrices. In terms of physical fields, V_0, V_m, and V_k can be rewritten as

$$\begin{aligned} V_0(\phi\phi^\dagger) = & -\frac{\mu^2}{2}(\sigma^2 + \eta^2 + \boldsymbol{\alpha}^2 + \boldsymbol{\pi}^2) \\ & + \frac{\lambda_1}{8}(\sigma^2 + \eta^2 + \boldsymbol{\alpha}^2 + \boldsymbol{\pi}^2)^2 \\ & + \frac{\lambda_2}{2}[(\sigma\boldsymbol{\alpha} + \eta\boldsymbol{\pi})^2 + (\boldsymbol{\alpha} \times \boldsymbol{\pi})^2], \end{aligned} \tag{4.19}$$

$$V_m(\phi, \phi^\dagger) = -\frac{1}{4} m e^{i\chi}(\sigma + i\eta) - \frac{1}{4} m e^{-i\chi}(\sigma - i\eta), \tag{4.20}$$

$$\begin{aligned} V_k(\phi, \phi^\dagger) = & -\frac{1}{2} K f^2(\sigma^2 - \eta^2 - \boldsymbol{\alpha}^2 + \boldsymbol{\pi}^2)\cos\theta \\ & - K(\sigma\eta - \boldsymbol{\alpha} \cdot \boldsymbol{\pi})\sin\theta. \end{aligned} \tag{4.21}$$

Assuming, for convenience,

$$\langle \phi \rangle = \frac{1}{2}\langle \sigma + i\eta \rangle = \frac{1}{2} v e^{i\varphi} \qquad (v > 0), \tag{4.22}$$

we get, by taking the extremum of $V_0 + V_m + V_k$ with respect to v and φ,

$$v^2 = \frac{2\mu^2}{\lambda_1} + \frac{2m}{\lambda_1 v}\cos(\chi + \varphi) - \frac{2Kf^2}{\lambda_1}\cos(\theta - 2\varphi) \tag{4.23}$$

and

$$m\sin(\chi + \varphi) - Kf^2 v \sin(\theta - 2\varphi) = 0. \tag{4.24}$$

Equation (4.24) plays a role of the vacuum alignment in the effective theory. If we take $\varphi = 0$ as we wish, (4.24) implies a consistency constraint on χ and θ: They are not separately independent parameters. They can be expressed in terms of the physical parameter $\bar{\theta} = \theta + 2\chi$ as

$$\sin\chi \cong -\frac{Kf^2 v}{m + 2Kf^2 v}\sin\bar{\theta}, \tag{4.25}$$

$$\sin\theta \cong -\frac{m}{m + 2Kf^2 v}\sin\bar{\theta}, \tag{4.26}$$

where we have assumed that $\sin\chi$ is very small ($\ll 1$).

Rewriting $\mathcal{L}_{\text{eff}}$ in terms of the shifted field $\phi \to \langle\phi\rangle + \phi$, we get

$$\begin{aligned} \mathcal{L}_{\text{eff}} = & -\bar{\psi}(\not{D}_s + \tfrac{1}{2} f v)\psi - \frac{1}{4}F_s^2 - (f\bar{\psi}_L \phi \psi_R + \text{H.c.}) - \text{Tr}(\partial_\mu\phi\partial_\mu\phi^\dagger) - \frac{1}{2}(\sigma, \eta) M^2_{\sigma\eta} \begin{pmatrix} \sigma \\ \eta \end{pmatrix} - \frac{1}{2}(\boldsymbol{\alpha}, \boldsymbol{\pi}) M^2_{\alpha\pi} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\pi} \end{pmatrix} \\ & - \frac{\lambda_1 v}{2}\sigma(\sigma^2 + \eta^2 + \boldsymbol{\alpha}^2 + \boldsymbol{\pi}^2) - \lambda_2 v \boldsymbol{\alpha} \cdot (\sigma\boldsymbol{\alpha} + \eta\boldsymbol{\pi}) - \frac{\lambda_1}{8}(\sigma^2 + \eta^2 + \boldsymbol{\alpha}^2 + \boldsymbol{\pi}^2)^2 - \frac{\lambda_2}{2}(\sigma\boldsymbol{\alpha} + \eta\boldsymbol{\pi})^2 - \frac{\lambda_2}{2}(\boldsymbol{\alpha} \times \boldsymbol{\pi})^2, \end{aligned} \tag{4.27}$$

where the meson mass matrices are given

$$M^2_{\sigma\eta} = \begin{pmatrix} \lambda_1 v^2 + \frac{m}{v}\cos\chi & -\frac{1}{2}Kf^2\sin\theta \\ -\frac{1}{2}Kf^2\sin\theta & \frac{m}{v}\cos\chi + 2Kf^2\cos\theta \end{pmatrix}, \tag{4.28}$$

$$M^2_{\alpha\pi} = \begin{pmatrix} \lambda_1 v^2 + \frac{m}{v}\cos\chi + 2Kf^2\cos\theta & \frac{1}{2}Kf^2\sin\theta \\ \frac{1}{2}Kf^2\sin\theta & \frac{m}{v}\cos\chi \end{pmatrix}.$$

The quark acquires a large constituent mass

$$m_Q = \frac{1}{2} f v \cong \frac{f\mu^2}{\lambda_1} + \frac{fm}{\lambda_1 v} + \frac{Kf^3}{\lambda_1}. \tag{4.29}$$

It is interesting to note that m_Q arises from three parts: the spontaneous chiral symmetry breaking (from V_0), the explicit chiral symmetry breaking (from V_m), and the instanton induced symmetry breaking (from V_k). The instanton does *spontaneously* break chiral symmetry $\text{SU}(L)_L \times \text{SU}(L)_R$ [32]. The mass spectrum of mesonic states can be read off from diagonalizing (4.28). The mixing probability is proportional to $(Kf^2\sin\theta)^2 = m^2\sin^2\chi$, which is of high order, thus it hardly affects the physical masses:

$$\begin{aligned} m_\eta^2 &= \frac{m}{v}\cos\chi + 2Kf^2\cos\theta\,, \\ m_\pi^2 &= \frac{m}{v}\cos\chi, \\ m_\sigma^2 &= \lambda_1 v^2 + \frac{m}{v}\cos\chi\,, \\ m_\alpha^2 &= \lambda_2 v^2 + \frac{m}{v}\cos\chi + 2Kf^2\cos\theta\,. \end{aligned} \tag{4.30}$$

Equation (4.30) clearly shows how the instanton induced V_k leads to a mass splitting between pions and the U(1) particle η. When $\bar{\theta}$ thus θ is small,

$$m_\eta^2 - m_\pi^2 = 2Kf^2, \tag{4.31}$$

and in the chiral limit $m \to 0$, $m_\pi^2 \to 0$, but $m_\eta^2 \to 2Kf^2$. We conclude that the U(1) problem is solved in the framework of the effective theory if $2Kf^2$ is big enough.

The CP-violating effects originate from the mixing between the scalar and pseudoscalars even though the mixing is negligible in computing the meson masses. To diagonalize the quadratic terms in (4.27), we define the

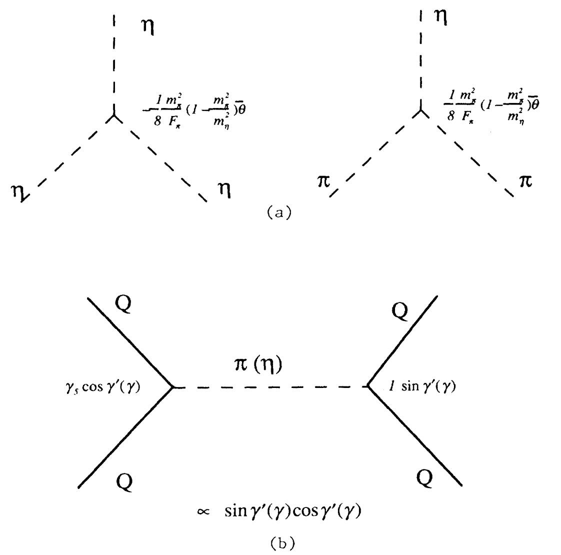

FIG. 1. (a) Feynman rules for η^3 and $\eta\pi^2$ couplings. (b) The CP-violating $qq \to qq$ scattering. We have assumed that $m_\sigma^2 \gg m_\eta^2$, $m_\alpha^2 \gg m_\pi^2$ and $v = 2F_\pi$.

physical meson fields (the primed fields)

$$\sigma = \sigma' \cos\gamma + \eta \sin\gamma, \quad \eta = -\sigma' \sin\gamma + \eta\cos\gamma, \tag{4.32}$$

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}' \cos\gamma' + \boldsymbol{\pi}\sin\gamma', \quad \boldsymbol{\pi} = -\boldsymbol{\alpha}\sin\gamma' + \boldsymbol{\pi}\cos\gamma', \tag{4.33}$$

such that the off-diagonal elements in (4.28) vanish. The mixing angles γ and γ' are determined

$$\gamma = \frac{Kf^2 \sin\theta}{m_\sigma^2 - m_\eta^2} = \frac{1}{2}\frac{m_\pi^2}{m_\sigma^2 - m_\eta^2}\left(1 - \frac{m_\pi^2}{m_\eta^2}\right)\bar{\theta}, \tag{4.34}$$

$$\begin{aligned}\gamma' &= -\frac{Kf^2\sin\theta}{m_\alpha^2 - m_\pi^2} \\ &= -\frac{1}{2}\frac{m_\pi^2}{m_\alpha^2 - m_\pi^2}\left(1 - \frac{m_\pi^2}{m_\eta^2}\right)\bar{\theta},\end{aligned} \tag{4.35}$$

which meet the criteria that the mixing and thus the strong CP violation must disappear as $m_\pi^2 \to 0$ or $m_\eta^2 = m_\pi^2$ or $\bar{\theta} = 0$. In terms of the physical fields, the CP-violating part of the effective potential is identified (for simplicity we drop the prime notations),

$$\begin{aligned}V_{CP} = {} & \frac{\lambda_1 v}{2}\sin\gamma\,\eta(\sigma^2 + \eta^2 + \boldsymbol{\alpha}^2 + \boldsymbol{\pi}^2) \\ & + \lambda_2 v \cos\gamma' \sin(\gamma - \gamma')\boldsymbol{\alpha}\cdot(\eta\boldsymbol{\alpha} - \sigma\boldsymbol{\pi}) \\ & + \lambda_2 v \sin\gamma\cos(\gamma - \gamma')\boldsymbol{\pi}\cdot(\sigma\boldsymbol{\alpha} + \eta\boldsymbol{\pi}),\end{aligned} \tag{4.36}$$

and the Yukawa coupling between quarks and mesons contains a CP-violating part too:

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} = {} & -\frac{1}{2}\bar{\psi}(\sin\gamma + i\gamma_5\cos\gamma)\psi\eta \\ & -\frac{1}{2}\bar{\psi}(\sin\gamma' + i\gamma_5\cos\gamma')\boldsymbol{\tau}\psi\cdot\boldsymbol{\pi}.\end{aligned} \tag{4.37}$$

The Feynman rules for CP-violating vertices and the typical CP-violating $qq \to qq$ amplitude are shown in Fig. 1.

The amplitude of $\eta \to 2\pi$ decays reads from (4.36)

$$A(\eta \to 2\pi) = \frac{1}{4}\frac{m_\pi^2}{F_\pi}\left(1 - \frac{m_\pi^2}{m_\eta^2}\right)\bar{\theta}, \tag{4.38}$$

where $F_\pi = \frac{v}{2}$. It is worth noting that (4.38) does not have a direct comparison with the QCD calculation (3.18) and (3.19) where we worked in the $L = 3$ case and η is one of the would-be Goldstone bosons. In (4.38), however, η has been referred to as the U(1) particle.

C. The EDM for the constituent quark

The CP-violating Yukawa coupling in (4.37) results in an important strong CP effect: the electric dipole moment (EDM) of the constituent quark. It can be examined by computing the effective interaction of the type when an external electromagnetic field A_μ^{em} is introduced:

$$\mu_{\text{EDM}}\bar{\psi}\gamma_5\sigma_{\mu\nu}\psi F_{\mu\nu}^{\text{em}}. \tag{4.39}$$

The coefficient μ_{EDM} is defined as the EDM of the quark. Since (4.39) is not invariant under the chiral rotation, the EDM can be converted into the magnetic moment if the fermion field is chiral (the chirality flip). When $m_Q \neq 0$, we have to check the phase of the constituent quark mass m_Q since only the phase difference between the quark mass and the effective interaction makes sense. In our convention, m_Q is real at the tree level. At a higher level, the mass acquires infinite renormalization. The renormalizability of our model guarantees that the renormalized mass will not develop a γ_5-dependent part. However, m_Q may acquire a finite renormalization which may contain a γ_5 part at high order. But that phase would be too small to cancel (4.39).

In the background of the em field, the charged quarks and pions couple to A_μ^{em} through the covariant derivative D_μ^{em}

$$-\bar{\psi}_Q\, \not{D}_Q^{\text{em}}\psi_Q - \left|D_\mu^{\text{em}}\pi^+\right|^2, \tag{4.40}$$

where

$$D_{\mu,Q}^{\text{em}} = \partial_\mu + eQA_\mu^{\text{em}} \tag{4.41}$$

and Q is the electric charge of the particle. Following Schwinger's formalism [33] on the derivation of the anomalous magnet moment of electron, we obtain the effective interactions

$$\begin{aligned}\int d^4x\, \mathcal{L}_{\text{eff}}^{\text{em}} = {} & -\int d^4x \sum_{Q=u,d}\bar{\psi}_Q(\not{D}_Q^{\text{em}} + m_Q)\psi_Q - \frac{f^2}{2!}\int d^4x\, d^4y \sum_{Q=u,d}\bar{\psi}_Q(x)e^{i\gamma'\gamma_5}S_{\pi^0\pi^0}S_F^Q(x,y)e^{i\gamma'\gamma_5}\psi_Q(y) \\ & -\frac{f^2}{2!}\int d^4x\, d^4y\bar{u}(x)e^{i\gamma'\gamma_5}S_{\pi^+\pi^-}S_F^d(x,y)e^{i\gamma'\gamma_5}u(y) - \frac{f^2}{2!}\int d^4xd^4y\bar{d}(x)e^{i\gamma'\gamma_5}S_{\pi^+\pi^-}S_F^Q(x,y)e^{i\gamma'\gamma_5}d(y),\end{aligned} \tag{4.42}$$

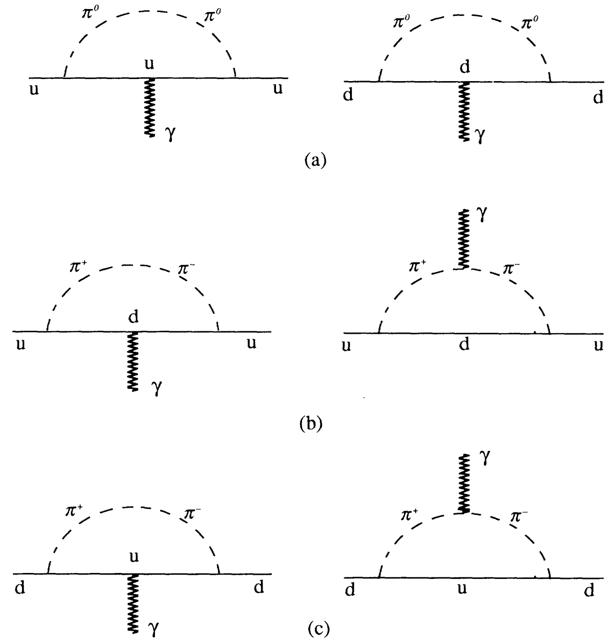

FIG. 2. Diagrammatic representations of Schwinger's formulation on the EDM's for constituent quarks.

where the $S_{\pi\pi}$'s and S_F^Q's are pion and quark propagators in the background of $A_\mu^{\rm em}$,

$$S_{\pi^0\pi^0}=\frac{1}{\partial^2-m_\pi^2},\quad S_{\pi^+\pi^-}=\frac{1}{(D_\mu^{\rm em})^2-m_\pi^2},$$
$$S_F^Q=\frac{1}{\not{D}_Q^{\rm em}+m_Q}. \tag{4.43}$$

Because $\frac{e^2}{4\pi}\ll 1$, we can expand these propagators perturbatively in e:

$$S_F^Q=\frac{\not{D}_Q^{\rm em}-m_Q}{(D_Q^{\rm em})^2-m_Q^2}\left(1+\frac{\frac{1}{2}eQ\sigma_{\mu\nu}F_{\mu\nu}^{\rm em}}{(D_Q^{\rm em})^2-m_Q^2}+\cdots\right), \tag{4.44}$$

$$S_{\pi^+\pi^-}=\frac{1}{\partial^2-m_\pi^2}\left(1+\frac{eA_\mu^{\rm em}\partial_\mu+e\partial_\mu A_\mu^{\rm em}}{\partial^2-m_\pi^2}+\cdots\right), \tag{4.45}$$

where the ellipses denote $O(e^2)$. The extraction of the effective interaction of (4.39) is done with the aid of Feynman diagrams in Fig. 2. The contributions from the second term in (4.42) correspond to Fig. 2(a), the third to Fig. 2(b), and the fourth to Fig. 2(c). Summing them up, we get

$$\mu_{\rm EDM}^u=\mu_{\rm EDM}^d$$
$$=\frac{ef^2}{32\pi^2}\sin 2\gamma' m_Q$$
$$\times\left[-\frac{2}{3}\frac{1}{m_Q^2-m_\pi^2}+\frac{m_Q^2}{(m_Q^2-m_\pi^2)^2}\ln\frac{m_Q^2}{m_\pi^2}\right]. \tag{4.46}$$

The EDM of the neutron is obtained by applying the SU(6) quark model,

$$\mu_{\rm EDM}^{\rm neutron}=\frac{4}{3}\mu_{\rm EDM}^d-\frac{1}{3}\mu_{\rm EDM}^u$$
$$\cong\frac{e}{2m_Q}\frac{f^2}{16\pi^2}\sin 2\gamma'\ln\frac{m_Q^2}{m_\pi^2}, \tag{4.47}$$

where we have used $m_Q^2\ll m_\pi^2$ and γ' is given in (4.35).

V. POSSIBLE SOLUTIONS TO THE STRONG *CP* PROBLEM

In above, we have studied extensively the measure of strong *CP* violation and its physical effects from viewpoints of QCD and an effective chiral theory. $\mathcal{J}_{\rm strong}$ is a product of quark masses, $\bar\theta$, and the instanton amplitude $K(L)$. It should vanish when any one of them vanishes. The most stringent experimental constraint on $\mathcal{J}_{\rm strong}$ comes from the EDM of the neutron, which has been measured at a very high precision [34]:

$$\mu_{\rm EDM}^{\rm neutron}<1.2\times 10^{-25}e\,{\rm cm}. \tag{5.1}$$

This implies

$$\mathcal{J}_{\rm strong}<10^{-16}\,{\rm GeV}^4. \tag{5.2}$$

At a typical hadron energy scale, one would suspect $\mathcal{J}_{\rm strong}\simeq\Lambda_{\rm QCD}^4\simeq 10^{-4}\sim 10^{-6}{\rm GeV}^4$, enormously larger than the upper limit. This is the so-called strong *CP* problem. It has puzzled us for more than a decade, ever since the instanton was discovered.

A. $m_u=0$ scenario

When $m_u=0$ and $\mathcal{J}_{\rm strong}=0$, the strong *CP* problem is most neatly and elegantly solved. In the meantime, the U(1) problem can be solved by instantons without resorting to other assumptions. There is an additional $\mathrm{U}(1)_A$ symmetry associated with the u quark. $m_u=0$, unlike setting $\bar\theta=0$, does increase the symmetry of the system and thus does not violate 't Hooft's naturalness principle. However, $m_u=0$ seems to contradict with the phenomenology where $m_u^{\rm expt}\simeq 5\sim 10$ MeV [35].

However, there is a loophole in this argument [36]. The instanton *explicitly* breaks $\mathrm{U}(1)_A$, as well as $\mathrm{U}(1)_A^u$ associated with the massless u quark if all other light quarks are *massive*. The instanton is acting as a flavor-changing force, as a result, the u quark acquires a radiative mass from other flavors. This is again due to the existence of the zero modes of $\not{D}$ in the nontrivial instanton field. In the presence of a massless fermion, the vacuum tunneling effect is suppressed unless we insert an operator that contains enough Grassmann fields to eliminate all the zero modes. In the $\nu=\pm1$ sector, the only operator which survives is $\bar uu$. To see how it works, let us recall the partition function $Z(\theta)$ in (3.9). $\langle\bar uu\rangle$ is calculated by taking the average over space-time:

$$\langle \bar{u}u \rangle_{\text{instanton}} = \frac{1}{V}\left\langle \int d^4x \bar{u}u(x) \right\rangle = -\frac{1}{V}\frac{d}{dm_u}\ln Z(\bar{\theta}) = -2m_d \cdots m_L K(L), \tag{5.3}$$

where we have rotated $\bar{\theta}$ to zero as we can when $m_u = 0$. Equation (5.3) implies that the $\mathrm{U}(1)_A^u$ symmetry is broken by the instanton. Of course we would not have the Goldstone boson associated with it since it is referred to as an explicit breaking. We should not confuse the condensate $\langle \bar{u}u \rangle$ caused by the spontaneous symmetry breaking with $\langle \bar{u}u \rangle_{\text{instanton}}$. The former can be nonzero even if all quarks are massless while the latter vanishes if the d quark mass is zero. The instanton induced u quark mass can be roughly estimated [37] in the $L = 2$ case where $K(2)$ is related to m_η^2,

$$m_u^{\text{instanton}} \cong -\pi\alpha_s(\bar{\rho})C_F\bar{\rho}^2\langle \bar{u}u \rangle_{\text{instanton}} = \frac{4}{3}\pi\alpha_s(\bar{\rho})\bar{\rho}^2F_\pi^2\frac{m_\eta^2 - m_\pi^2}{m_Q^2}m_d \cong 4\ \text{MeV}, \tag{5.4}$$

where we take $\bar{\rho} \simeq (\frac{1}{3}\Lambda_{\text{QCD}})^{-1}$, $K = -\frac{1}{2f^2}(m_\eta^2 - m_\pi^2)$, and $f = \frac{2m_Q}{F_\pi}$. $m_u^{\text{instanton}}$ must be viewed as an explicit mass because of its proportionality to m_d. What seems remarkable is that the order of magnitude of $m_u^{\text{instanton}}$ is consistent with the phenomenological value. The massless u quark is still the most favorable solution to the strong *CP* problem.

B. Peccei-Quinn symmetry

Another possibility of rendering $\mathcal{J}_{\text{strong}} = 0$ is that $\bar{\theta} = 0$ for some dynamical reason. This is realized if the phase of the quark masses $\theta_{QFD} = \sum_i \varphi_i$ is equal to $-\theta_{\text{QCD}}$. A decade ago, Peccei and Quinn [38] suggested that the strong *CP* problem may be naturally solved if one or more quarks acquire the current masses entirely through the Higgs mechanism where the Lagrangian of quarks and scalars exhibits an adjoint chiral symmetry: the Peccei–Quinn (PQ) symmetry.

For simplicity, let us examine a toy model of a single quark

$$\mathcal{L}_{\text{toy}} = -\bar{\psi}\,\not{D}\psi - \frac{1}{4}F^2 + i\theta F\tilde{F} - (f\bar{\psi}_L\psi_R\phi + \text{H.c.}) - \partial_\mu\phi\partial_\mu\phi^* - V_0(\phi,\phi^*), \tag{5.5}$$

where

$$V_0(\phi,\phi^*) = -\mu^2\phi\phi^* + \frac{1}{4}\lambda(\phi\phi^*)^2. \tag{5.6}$$

Equation (5.5) is invariant under the PQ symmetry:

$$\psi_R \to e^{i\alpha}\psi_R, \quad \psi_L \to e^{-i\alpha}\psi_L, \qquad \phi \to e^{-2i\alpha}\phi, \quad \phi^* \to e^{2i\alpha}\phi^*. \tag{5.7}$$

The PQ symmetry is broken at the quantum level by the chiral anomaly, and effectively

$$\mathcal{L}_{\text{toy}} \to \mathcal{L}_{\text{toy}} - 2i\alpha F\tilde{F}. \tag{5.8}$$

Choosing $\alpha = \frac{\theta}{2}$ yields $\bar{\theta} = 0$.

The effective potential of the scalar fields can be calculated in a similar way to (4.16),

$$V_{\text{eff}}(\phi,\phi^*) = -\mu^2\phi\phi^* + \frac{1}{4}\lambda(\phi\phi^*)^2 - Kf^*e^{-i\theta}\det\phi^* - Kfe^{i\theta}\det\phi, \tag{5.9}$$

where K is the instanton amplitude. The last two terms in the effective potential breaks the PQ symmetry. The VEV's of ϕ and ϕ^* are found to be

$$\langle f\phi \rangle = ve^{-i\theta}, \quad \langle f^*\phi^* \rangle = ve^{i\theta}, \tag{5.10}$$

and

$$v^2 = \frac{2\mu^2|f|^2}{\lambda} + \frac{2K|f|^4}{\lambda v}. \tag{5.11}$$

Thus the fermion mass is read off from the Yukawa interaction $m = fve^{-i\theta}$ and

$$\bar{\theta} = \theta + \arg\langle f\phi \rangle = 0. \tag{5.12}$$

The axion [39] mass is readily derived from (5.9) by diagonalizing the quadratic terms

$$m_{\text{axion}}^2 = \frac{2K|f|^2}{v}. \tag{5.13}$$

Unfortunately, we have not yet been able to discover this particle so far.

VI. CONCLUSIONS

We have studied the measure of *CP* violation in strong interactions. It arises from the nontrivial topological structure of Yang-Mills fields, a nonzero vacuum angle $\bar{\theta}$, as well as nonvanishing quark current masses. The instanton dynamics makes the most sense in dealing with the topological gauge configurations where the semiclassical method applies. It has been shown that the instanton dynamics, as a consistent field theory, automatically satisfies the so-called anomalous Ward identity. Crewther's original complaints on the topological susceptibility and θ periodicity of the fermion operator are a result of inconsistently handling the singularities in some fermion operators. We conclude that QCD theory itself does not put any constraint on the instanton computation.

In the presence of the chiral anomaly, there is no would-be $\mathrm{U}(1)_A$ Goldstone particle. By studying an effective chiral theory, we find that the instanton leads to an explicit $\mathrm{U}(1)_A$ symmetry breaking. If the instanton is to solve the U(1) problem, the measure of the strong *CP* violation is connected to the mass of the U(1) particle. It may be natural to think that the strong *CP* problem is the side effect of the U(1) problem and both problems cannot be solved simultaneously in the context of QCD.

However, we point out that the massless u quark scenario to solve the strong CP problem may not be such a silly idea. The u quark may acquire a mass from the d quark through the instanton interaction in which the fermion zero modes play an essential role. In any case, with the failure so far to observe axions experimentally, the strong CP problem is wide open to new mechanisms.

ACKNOWLEDGMENTS

I wish to thank Dr. D.D. Wu and Dr. K.S. Viswanathan for useful discussions.

---

[1] A. Belavin, A. Polyakov, A. Schwartz, and Y. Tyupkin, Phys. Lett. **59B**, 85 (1975).

[2] G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. D **14**, 3432 (1976).

[3] S. Weinberg, Phys. Rev. D **11**, 3594 (1975).

[4] C. Callan, Jr., R. Dashen, and D. Gross, Phys. Lett. **63B**, 334 (1976); R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 72 (1976).

[5] R. Peccei, in *CP Violation*, edited by C. Jarlskog (World Scientific, Singapore, 1989).

[6] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 625 (1973).

[7] C. Jarlskog, in *CP Violation* [5].

[8] J. S. Bell and R. Jackiw, Nuovo Cimento A **60**, 47 (1969); S. L. Adler and W. A. Bardeen, Phys. Rev. **182**, 1517 (1969).

[9] S. Aoki and T. Hatsuda, CERN Report No. CERN-TH-5808/90, 1990 (unpublished); H. Y. Cheng, Phys. Rev. D **44**, 166 (1991); A. Pich and E. de Rafael, Nucl. Phys. **B367**, 313 (1991).

[10] R. J. Crewther, Phys. Lett. **70B**, 349 (1977); Riv. Nuovo Cimento **2**, 63 (1979); Phys. Lett. **93B**, 75 (1980); Nucl. Phys. **B209**, 413 (1982).

[11] M. Shifman, A. Vainshtein, and V. Zakharov, Nucl. Phys. **B166**, 493 (1980).

[12] G. A. Christos, Phys. Rep. **116**, 251 (1984).

[13] E. Witten, Nucl. Phys. **B149**, 285 (1979).

[14] L. Brown, R. Carlitz, and C. Lee, Phys. Rev. D **16**, 417 (1977).

[15] G. Veneziano, Nucl. Phys. **B159**, 213 (1979).

[16] D. I. Dyakonov and M. I. Eides, Zh. Eksp. Teor. Fiz. **81**, 434 (1981) [Sov. Phys. JETP **54**, 232 (1981)].

[17] R. Dashen, Phys. Rev. D **3**, 1879 (1971).

[18] Z. Huang, K. S. Viswanathan, and D. D. Wu, Mod. Phys. Lett. A **6**, 711 (1991); Z. Huang and D. D. Wu, Commun. Theor. Phys. **16**, 363 (1991).

[19] C. G. Callan, Jr., R. F. Dashen, and D. J. Gross, Phys. Rev. D **17**, 2717 (1978); S. Coleman, *Aspects of Symmetry* (Cambridge University Press, Cambridge, England, 1985); N. A. McDougall, Nucl. Phys. **B211**, 139 (1983).

[20] G. 't Hooft, Phys. Rev. D **14**, 3432 (1976).

[21] R. D. Carlitz and D. B. Creamer, Ann. Phys. (N.Y.) **118**, 429 (1979); N. Andrei and D. J. Gross, Phys. Rev. D **18**, 468 (1978).

[22] Z. Huang, K. S. Viswanathan, and D. D. Wu, Mod. Phys. Lett. A **7**, 3147 (1992).

[23] V. Baluni, Phys. Rev. D **19**, 2227 (1979).

[24] R. Crewther, P. Di Vecchia, G. Veneziano, and E. Witten, Phys. Lett. **88B**, 123 (1979).

[25] D. I. Dyakonov and V. Yu. Petrov, Nucl. Phys. **B245**, 259 (1984); **B272**, 475 (1986); E. V. Shuryak, *ibid.* **B302**, 559 (1988).

[26] M. Atiyah and I. Singer, Ann. Math. **87**, 484 (1968); M. Atiyah, R. Bott, and V. Patodi, Invent. Math. **19**, 279 (1973).

[27] M. Gell-Mann and M. Levy, Nuovo Cimento **16**, 705 (1960).

[28] A. Manohar and H. Georgi, Nucl. Phys. **B234**, 189 (1984).

[29] P. Carruthers and R. Haymaker, Phys. Rev. D **4**, 406 (1971); S. Raby, *ibid.* **13**, 2594 (1976).

[30] G. 't Hooft, Phys. Rep. **142**, 357 (1986); E. Mottola, Phys. Rev. D **21**, 3401 (1980); E. P. Shabalin, Pis'ma Zh. Eksp. Teor. Fiz. **36**, 981 (1982) [Sov. J. Nucl. Phys. **36**, 575 (1982)].

[31] P. Di Vecchia and G. Veneziano, Nucl. Phys. **B171**, 253 (1980).

[32] D. G. Caldi, Phys. Rev. Lett. **39**, 121 (1977); R. D. Carlitz, Phys. Rev. D **17**, 3225 (1978).

[33] J. Schwinger, *Particles, Sources, and Fields* (Addison-Wesley, Reading, 1989).

[34] K. F. Smith *et al.*, Phys. Lett. B **234**, 191 (1990).

[35] J. Gasser and H. Leutwyler, Phys. Rep. **87**, 771 (1982).

[36] K. Choi, C. W. Kim, and W. K. Sze, Phys. Rev. Lett. **61**, 794 (1988).

[37] H. D. Politzer, Nucl. Phys. **B117**, 397 (1976); M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *ibid.* **B163**, 43 (1980); **B165**, 45 (1981).

[38] R. Peccei and H. Quinn, Phys. Rev. Lett. **38**, 1440 (1977); Phys. Rev. D **16**, 1791 (1977).

[39] S. Weinberg, Phys. Rev. Lett. **40**, 223 (1978); F. Wilczek, *ibid.* **40**, 279 (1978).