Ref.TH.2522-CERN

EFFECTS OF TOPOLOGICAL CHARGE IN GAUGE THEORIES

R.J. Crewther

CERN -- Geneva

Lectures presented at the
XVII Internationale Universitätswochen für Kernphysik,
Schladming, Austria
21 February -- 3 March, 1978

Ref.TH.2522-CERN
18 August 1978

CONTENTS

1. INTRODUCTION

These lectures concern the properties of topological charge in gauge theories and the physical effects which have been attributed to its existence.

No introduction to this subject would be adequate without a discussion of the original work of Belavin, Polyakov, Schwarz and Tyupkin[1], of the beautiful calculation by 't Hooft[2,3], and of the occurrence of θ-vacua[4-6]. Other important topics include recent progress on solutions of the Yang-Mills equation of motion[7,8], and the problem of parity and time-reversal invariance in strong interactions[9] [axions[10,11], etc.].

In a few places, I have strayed from the conventional line and in one important case, disagreed with it. The important remark concerns the connection between chirality and topological charge first pointed out by 't Hooft[2]: in the literature, the rule is repeatedly quoted *with the wrong sign!* If Q_5 is the generator for Abelian chiral transformations of massless quarks with N flavours, the correct form of the rule is

$$\Delta Q_5 = -2N\{\text{topological charge}\} \tag{1.1}$$

where ΔQ_5 means the out eigenvalue of Q_5 minus the in eigenvalue. The sign can be checked by consulting the standard WKB calculation[2,3], rotating to Minkowski space, and observing that the sum of *right*-handed chiralities of operators in a Green's function equals $-\Delta Q_5$. The wrong sign is an automatic consequence of a standard but incorrect derivation in which the axial charge is misidentified. These points are examined in Sections 7 and 8, and influence the discussion in Section 9. Elsewhere, I discuss loopholes in the argument that finite Euclidean action implies integer topological charge (Section 2), offer some opinions about the problem of integration over instanton sizes (Section 4), and include some brief remarks (Section 6) about generalizing θ-vacua to non-WKB situations (e.g. to the real world).

Inevitably, some important topics have had to be omitted:

a) There is no explicit treatment of the U(1) problem[12], although related issues, such as Eq. (1.1) and the derivation of anomalous Ward identies, are discussed here. A full review would require my going far afield into current algebra.

b) Nor have I considered merons[13], which are best explained by the proponents of the program[14].

c) I have restricted the discussion to Yang-Mills theories in four dimensions. For topological considerations in gravity and the Schwinger model, and for the underlying mathematics of fibre bundles, see the lectures of Schroer[15] and Thirring[16].

2. TOPOLOGICAL CHARGE

Let $A_\mu^a(x)$ be a Yang-Mills field. The gauge (colour) index a and Lorentz index μ are gauge-transformed as follows[17]

$$A_\mu^a \tau^a \longrightarrow \left(A_\mu^a\right)' \tau^a = G^{-1} A_\mu^a \tau^a G + i g^{-1} G^{-1} \partial_\mu G \tag{2.1}$$

where g is a coupling constant,

$$G[\xi] = \exp\, i \xi^a(x) \tau^a \tag{2.2}$$

is an element of the gauge group parametrized by functions $\xi^a(x)$, τ^a generate the group in a suitable representation,

$$\left[\tau^a, \tau^b\right] = i c^{abc} \tau^c \tag{2.3}$$

and c^{abc} are the structure constants of the gauge group. The field-strength tensor $F_{\alpha\beta}$ and its dual $F^*_{\alpha\beta}$ are given by

$$F_{\alpha\beta}{}^a = \partial_\alpha A_\beta^a - \partial_\beta A_\alpha^a + g c^{abc} A_\alpha^b A_\beta^c \tag{2.4}$$

$$F_{\alpha\beta}{}^* = \frac{1}{2} \varepsilon_{\alpha\beta\mu\lambda} F^{\mu\lambda} \tag{2.5}$$

Topological charge (or topological number, Pontryagin index, second Chern class number, etc.) for Yang-Mills fields is given by the formula[1]

$$\begin{aligned} \nu &= \left(g^2/64\pi^2\right) \varepsilon^{\alpha\beta\gamma\delta} \int d^4x \;\; F_{\alpha\beta}{}^a(x)\, F_{\gamma\delta}{}^a(x) \\ &= \left(g^2/32\pi^2\right) \int d^4x \;\; F.F^*(x) \end{aligned} \tag{2.6}$$

This definition is valid in four-dimensional Euclidean space (metric $\delta_{\mu\lambda}$; $\mu,\lambda = 1, 2, 3, 4$) if the antisymmetric tensor $\varepsilon_{\alpha\beta\gamma\delta}$ is given by $\varepsilon_{1234} = +1$. The same formula works in Minkowski space (metric $g_{\mu\lambda} = (1, -1, -1, -1)$; $\mu,\lambda = 0, 1, 2, 3$) with

$$\varepsilon_{0123} = - \varepsilon^{0123} = +1 \tag{2.7}$$

The significance of ν is:

i) It classifies configurations A_μ^a in functional integrals

$$\int \left[dA_\mu^a\right] \; \exp\, i \left\{ \text{Action} \right\}.$$

Contributions with $\nu \neq 0$ are necessarily non-perturbative.

ii) It can appear as a new gauge-invariant term in the (Minkowskian) action of quantum chromodynamics (QCD)[18]

$$S_{QCD}(\phi) \quad = \quad S_{QCD}(0) \quad - \quad \phi\nu \tag{2.8}$$

where ϕ is a coupling constant and

$$S_{QCD}(0) \quad = \quad \int d^4x \; \mathcal{L}_{QCD}(x) \tag{2.9}$$

is constructed from the QCD Lagrangian

$$\mathcal{L}_{QCD} \quad = \quad - \tfrac{1}{4} F^2 \quad + \quad \sum_{i=1}^{N} \bar{q}_i (i\not{D} \; - \; M_i)\, q_i \qquad (q_i \; = \; \text{quark field}, \; i^{th} \text{ flavour}). \tag{2.10}$$

The new term conserves charge conjugation C, but is odd under parity P and time-reversal T.

Why call ν a "topological charge"? At first sight, ν appears to have a complicated dependence on the field A_μ. However, consider the effect of a local variation $\delta/\delta A_\mu$ of the field

$$\begin{aligned} \delta A_\alpha^a(x)/\delta A_\beta^b(y) \quad &= \quad \delta_\alpha^{\;\beta}\, \delta^{ab} \delta^4(x-y) \quad , \\ \delta F_{\alpha\beta}^{\;a}(x)/\delta A_\gamma^b(y) \quad &= \quad \left(D_\alpha^{ab}\, \delta_\beta^{\;\gamma} \; - \; D_\beta^{\;ab} \delta_\alpha^{\;\gamma}\right)_x \delta^4(x-y) \end{aligned} \tag{2.11}$$

where D_α^{ab} is the covariant derivative

$$D_\alpha \quad = \quad \partial_\alpha \quad - \quad i g\, A_\alpha^c \tau^c \tag{2.12}$$

evaluated in the adjoint representation. Equations (2.6) and (2.11) imply the result

$$\begin{aligned} \delta\nu[A]/\delta A^\alpha \quad &= \quad (g^2/8\pi^2)\, D^\beta F_{\alpha\beta}^{\;*} \\ &= \quad 0 \end{aligned} \tag{2.13}$$

where the last line follows from the Bianchi identity. Therefore $\nu[A]$ is a topological invariant: its value does not change when A_μ is locally deformed. Fields can be grouped in classes labelled by the value of the topological charge[1].

Since $\nu[A]$ does not change under local variations of A_μ, it must be determined by the large-scale properties of A_μ. Therefore we look for a surface-integral representation. This can be obtained from the divergence relation[1,19]

$$\begin{aligned} \partial_\mu K^\mu \quad &= \quad (g^2/32\pi^2)\, F.F^* \\ K^\mu \quad &= \quad (g^2/32\pi^2)\, \varepsilon^{\mu\alpha\beta\gamma} A_\alpha^{\;a}\left(F_{\beta\gamma}^{\;a} \; - \; \tfrac{1}{3} g\, c^{abc} A_\beta^{\,b} A_\gamma^{\;c}\right) \end{aligned} \tag{2.14}$$

where K^μ is fixed up to the effect of gauge transformations (2.1):

$$\begin{aligned} K^\mu \rightarrow (K^\mu)' \quad = \quad K^\mu \quad &+ \quad (24\pi^2)^{-1} \mathrm{Tr}\, \varepsilon^{\mu\alpha\beta\gamma} G^{-1}\partial_\alpha G\, G^{-1}\partial_\beta G\, G^{-1}\partial_\gamma G \\ &+ \quad (ig/8\pi^2)\, \partial_\beta\, \mathrm{Tr}\, \varepsilon^{\mu\alpha\beta\gamma} \partial_\alpha G\, G^{-1} A_\gamma^{\;a} \tau^a \, . \end{aligned} \tag{2.15}$$

In Euclidean space (to which the remainder of this section refers), the result is

$$\nu \quad = \quad \lim_{R\to\infty} \oint_R d\sigma_\mu \, K_\mu(x) \tag{2.16}$$

where the surface integration is performed over the hypersphere in four dimensions

$$S_3(R) \quad = \quad \left\{ x_1, \ldots, x_4 \; ; \quad x_\mu x_\mu \;=\; R^2 \right\} \tag{2.17}$$

[The subscript 3 in $S_3(R)$ refers to the dimensionality of the surface of the hypersphere.] Of course, the limit $R \to \infty$ exists only if $\int d^4x \, F.F^*(x)$ converges.

The inequality

$$- F^2 \;\leqslant\; F.F^* \;\leqslant\; F^2 \tag{2.18}$$

implies that $\int d^4x \, F.F^*(x)$ is absolutely convergent if the Euclidean action

$$S[A] \quad = \quad \frac{1}{4} \int d^4x \;\; F^2(x) \tag{2.19}$$

is finite. [The finiteness of the classical action is relevant for the semi-classical calculations of 't Hooft[3]; see Section 4.] Integration of (2.18) over all x_μ yields the result[1)]

$$S[A] \quad \geqslant \quad 8\pi^2 \left|\nu[A]\right| / g^2 \tag{2.20}$$

with equality if and only if $F_{\mu\nu}$ is self-dual or anti-self-dual:

$$F_{\mu\nu} \quad = \quad \pm \; F_{\mu\nu}{}^* \tag{2.21}$$

Note that $S[A]$ is not a topological invariant, because

$$\delta S[A]/\delta A_\beta \quad = \quad - \; D_\alpha F_{\alpha\beta} \tag{2.22}$$

need not vanish. Fields corresponding to stationary points of the action are solutions of the equation of motion

$$D_\alpha F_{\alpha\beta} \quad = \quad 0 \tag{2.23}$$

In particular, fields satisfying (2.21) are solutions because of the Bianchi identity.

There is a non-rigorous (?) argument due to Belavin et al.[1)] that $\nu[A]$ should be an integer if $S[A]$ is finite. The finiteness of $S[A]$ implies

$$x^2 \, F_{\mu\nu}(x) \quad \longrightarrow \quad 0 \tag{2.24}$$

as x_μ tends to ∞ in almost every direction, so the field A^a_μ should tend to a pure-gauge configuration:

$$g A_\mu^\alpha \tau^\alpha \quad \rightarrow \quad i\, G^{-1}\partial_\mu G \tag{2.25}$$

$$K_\mu \quad \longrightarrow \quad (24\pi^2)^{-1}\, \varepsilon_{\mu\alpha\beta\gamma}\, \mathrm{Tr}\; G^{-1}\partial_\alpha G\, G^{-1}\partial_\beta G\, G^{-1}\partial_\gamma G \tag{2.26}$$

This asymptotic form for K_μ is conserved in regions where G(x) is non-singular, so the surface integral (2.16) can be written

$$\begin{aligned} \nu \;&=\; (24\pi^2)^{-1} \oint_R d\sigma_\mu\; \varepsilon_{\mu\alpha\beta\gamma}\, \mathrm{Tr}\; G^{-1}\partial_\alpha G\, G^{-1}\partial_\beta G\, G^{-1}\partial_\gamma G \\ &=\; \text{independent of } R \end{aligned} \tag{2.27}$$

if R is large enough to avoid singularities of the integrand in Eq. (2.27).

This integrand can be regarded as the Jacobian of a mapping. For example, in the simplest case of the SU(2) gauge group, the group elements can be parametrized as follows,

$$G \;=\; V_4(x) \;-\; i\vec{\sigma}.\vec{V}(x) \quad , \quad (\vec{\sigma} \;=\; \text{Pauli matrices}) \tag{2.28}$$

where the conditions

$$G G^\dagger \;=\; I \quad , \qquad \det G \;=\; 1 \tag{2.29}$$

restrict V_μ to be real and to have unit length:

$$\vec{V}^2 \;+\; V_4^2 \;=\; V^2 \;=\; 1 \tag{2.30}$$

Equation (2.27) becomes

$$\nu \;=\; (12\pi^2)^{-1} \oint d\sigma(x)\; \varepsilon_{\mu\alpha\beta\gamma}\, \varepsilon_{\zeta\lambda\delta\eta}\; \hat{x}_\mu\, V_\zeta\, \partial_\alpha V_\lambda\, \partial_\beta V_\delta\, \partial_\gamma V_\eta \tag{2.31}$$

Now consider the mapping of the hypersphere S_3 in x_μ-space onto the hypersphere $V^2 = 1$ in V_μ-space (Fig. 1). For each sphere, there is a normal direction ($\hat{x}_\mu$ or V_μ) and three mutually orthogonal tangential directions. Let indices

$$i, j, \;\ldots\;, n \quad = \quad 1, 2, 3$$

label the tangential directions, and let ∇_i denote the derivative with respect to x in the i^{th} tangential direction. Then the Jacobian of the mapping is given by

$$\det_{3\times 3} \left(\nabla_i V_j \right) \quad = \quad \tfrac{1}{6}\; \varepsilon_{ijk}\, \varepsilon_{\ell mn}\, \nabla_i V_\ell\, \nabla_j V_m\, \nabla_k V_n \tag{2.32}$$

Since $\varepsilon_{\mu\alpha\beta\gamma}\hat{x}_\mu$ is the same as ε_{ijk} when (α,β,γ) are identified with (i,j,k), and similarly for $\varepsilon_{\zeta\lambda\delta\eta}V_\zeta$, Eq. (2.32) can be written in the equivalent form

$$\det_{3\times 3} \left(\nabla_i V_j \right) \;=\; \tfrac{1}{6}\; \varepsilon_{\mu\alpha\beta\gamma}\, \varepsilon_{\zeta\lambda\delta\eta}\; \hat{x}_\mu\, V_\zeta\, \partial_\alpha V_\lambda\, \partial_\beta V_\delta\, \partial_\gamma V_\eta \tag{2.33}$$

If the mapping is assumed to be differentiable, the Jacobian relates surface elements on the spheres

$$d\sigma(V) \quad = \quad \det_{3\times 3}\left(\nabla_i V_j\right) \; d\sigma(x) \tag{2.34}$$

so Eq. (2.31) simplifies to the result

$$\nu \quad = \quad \left(2\pi^2\right)^{-1} \oint_{\text{mult.}} d\sigma(V) \tag{2.35}$$

where the sign of integration carries a notation indicating that S_3(x-space) may be "wound" several times onto S_3(V-space) -- in general, the mapping is not one-to-one. Each winding corresponds to four-dimensional solid angle $2\pi^2$, so $|\nu|$ must equal the number of windings:

$$\nu \quad = \quad \text{integer} \tag{2.36}$$

The simplest example of a mapping is provided by the gauge-group element

$$G_1(x) \quad = \quad \left(x_4 - i\vec{\sigma}.\vec{x}\right)/\sqrt{x^2} \tag{2.37}$$

which is well-behaved everywhere except at $x_\alpha = 0$. Here $V_\alpha(x)$ equals $\hat{x}_\alpha$, so the mapping $S_3 \to S_3$ is one-to-one. Therefore, if G_1 is substituted for G in (2.27), we expect $\nu = 1$, and indeed, that is the result obtained by direct calculation. If we had substituted $G_1^{+} = G_1^{-1}$, the answer would have been $\nu = -1$. The minus sign corresponds to a right-handed system of tangential coordinates on S_3(x-space) being mapped into a left-handed set on S_3(V-space).

Equation (2.36) remains valid for any simple Lie group, given the same assumptions about smoothness of the asymptotic $x_\mu \to \infty$ configuration. However, the analysis is less direct. It depends on the fact (discussed in the Appendix) that any continuous mapping

$$S_3 \longrightarrow \text{simple gauge group}$$

can be continuously deformed into a mapping

$$S_3 \longrightarrow SU(2) \text{ subgroup}$$

Then the result follows, because the formula (2.27) for ν is invariant under small deformations δG of G,

$$\begin{aligned} \delta\nu \quad &= \quad -\left(8\pi^2\right)^{-1} \oint d\sigma_\mu\, \varepsilon_{\mu\alpha\beta\gamma}\, \partial_\alpha \left\{ \operatorname{Tr} G^{-1}\partial_\beta G\, \partial_\gamma G^{-1}\, \delta G \right\} \\ &= \quad 0 \end{aligned} \tag{2.38}$$

and hence under the full continuous deformation.

There is an unresolved question[12,20] about whether $S[A] < \infty$ is really sufficient to ensure ν = integer for smooth fields A_μ. The analysis depends crucially on the mapping from S_3 to S_3 being smooth, i.e. the mapping is supposed to belong to a homotopy class. However, the assumed smoothness of $A_\mu^a(x)\tau^a$ is not sufficient to ensure smoothness of the corresponding pure gauge configuration $ig^{-1}G^{-1}\partial_\mu G$. It may be possible to start with a completely smooth function $A_\mu^a(x)$ which asymptotically approximates (in almost every direction) a configuration which is not smooth. For example, suppose the asymptotic configuration is a pure gauge term obtained from the gauge group element

$$G_1^p = \exp\left\{-ip\vec{\sigma}.\vec{x}\,|\vec{x}|^{-1}\arccos\left(x_4/\sqrt{x^2}\right)\right\} \tag{2.39}$$

If p is not an integer, $ig^{-1}G_1^{-p}\partial_\mu G_1^p$ is singular for $\vec{x} = 0$, $x_4 \leq 0$. However, it is still possible to perform the surface integration in Eq. (2.27) with the result

$$\begin{aligned}(24\pi^2)^{-1}\oint d\sigma_\mu\ \varepsilon_{\mu\alpha\beta\gamma}\ \mathrm{Tr}\ G_1^{-p}\partial_\alpha G_1^p\, G_1^{-p}\partial_\beta G_1^p\, G_1^{-p}\partial_\gamma G_1^p \\ = p - (2\pi)^{-1}\sin(2p\pi)\end{aligned} \tag{2.40}$$

which is not an integer for $p \neq$ integer. I have not been able to close the loophole or find a counter-example. For example, I looked at configurations A_μ obtained by substituting smooth functions for the $|\vec{x}|^{-k}$ singularities of $ig^{-1}G_1^{-p}\partial_\mu G_1^p$; a typical substitution is

$$|\vec{x}|^{-k} \longrightarrow \left[\vec{x}^2 + \mu^2\exp - \Lambda\,\vec{x}^2x^2\right]^{-k/2} \tag{2.41}$$

The result is certainly a smooth configuration $A_\mu^a(x)$ which asymptotically approximates $ig^{-1}G_1^{-p}\partial_\mu G_1^p$ in almost every direction. Unfortunately, the approximation is not good enough -- the Euclidean action diverges. All attempts to construct a counter-example gave the same result*).

It is certainly true that ν must be an integer if the configuration $A_\mu^a(x)$ can be smoothly mapped onto the unit hypersphere in five dimensions, or onto any smooth compact four-dimensional surface without boundary. Consider the hypersphere (Fig. 2)

$$S_4 = \left\{\xi_\alpha\ ,\ \alpha = 1, \ldots, 5\ ;\ \xi^2 \equiv \sum_{\alpha=1}^{5}\xi_\alpha^2 = 1\right\} \tag{2.42}$$

---

*) Even if a counter-example should be found in the future, this would not necessarily imply the existence of a counter-example satisfying (2.21) or (2.23).

Note added: Marino and Swieca [Cath. Univ. (Rio de Janeiro) preprint PUC-RIO-NC 78-7 (1978)] have shown that the action is necessarily infinite for a large class of fields with asymptotic behaviour given by $ig^{-1}G_1^{-p}\partial_\mu G_1^p$.

The origin of Euclidean space R_4 is mapped*) onto the N pole of S_4, $\xi = (0, 0, 0, 0, 1)$, and infinity is mapped onto a *single* point, the S pole of S_4, irrespective of the direction in which x_μ tends to ∞:

$$\begin{aligned} \xi_\mu &= 2x_\mu/(1 + x^2) \quad , \quad (\mu = 1, \ldots, 4), \\ \xi_5 &= (1 - x^2)/(1 + x^2) \end{aligned} \tag{2.43}$$

The field on the hypersphere[21)]

$$\begin{aligned} \hat{A}_\mu(\xi) &= \tfrac{1}{2}(1 + x^2) A_\mu - x_\mu x_\nu A_\nu \quad , \\ \hat{A}_5(\xi) &= - x_\mu A_\mu \end{aligned} \tag{2.44}$$

points in a tangential direction

$$\xi_\alpha \hat{A}_\alpha = 0 \tag{2.45}$$

The configuration $\hat{A}_a(\xi)$ is assumed to be regular in "gauge patches" (Fig. 3) which cover S_4, i.e. within the i^{th} patch, the field is described by a regular function $\hat{A}_a^{(i)}(\xi)$, and in the overlap region between patches 1 and 2, $\hat{A}^{(1)}$ and $\hat{A}^{(2)}$ are related by a smooth gauge transformation. This means that gauge-invariant quantities are smooth everywhere on S_4.

The fact that ν is an integer is a consequence of the Atiyah-Singer index theorem[15,22-27)], which is formulated for compact manifolds with gauge-patching. It relates the number of zero eigenvalues of covariant operators D_μ to the topological charge of the field contained in D_μ. In particular, consider spin-½ S_4-mappable solutions ψ of the Euclidean equation

$$\not{D}\psi \equiv (\not{\partial} - ig\not{A}.\tau)\psi = 0 \tag{2.46}$$

where ψ transforms under the fundamental representation of the gauge group. If A_μ is S_4-mappable, there are n_R right-handed and n_L left-handed solutions, with

$$\nu = n_R - n_L \tag{2.47}$$

Since n_R and n_L are integers, so is ν.

---

*) It is important to distinguish the mappings considered in Figs. 1 and 2. Figure 1 refers to the three-parameter gauge group element associated with the asymptotic $x \to \infty$ limit of the configuration, so there are three orthogonal tangential directions on the surfaces of the hyperspheres. In Fig. 2, the configuration itself is being mapped; it depends on $(x_1, \ldots, x_4)$, so a hypersphere S_4 with four orthogonal tangential directions is needed.

Note: The Bjorken-Drell conventions in Minkowski space give

$$\psi_R^{\ M} = \gamma_5^{\ M} \psi_R^{\ M} \qquad (2.48)$$
$$[\gamma_\mu , \gamma_\nu]_+^{\ M} = 2 g_{\mu\nu} \quad , \quad \gamma_5^{\ M} = i(\gamma^0\gamma^1\gamma^2\gamma^3)^M$$

for a right-handed spinor ψ_R^M. The rotation to Euclidean space introduces a minus sign

$$\psi_R^{\ M} \longleftrightarrow \psi_R^{\ E} \quad , \quad \gamma_5^{\ M} = -\gamma_5^{\ E} \qquad (2.49)$$
$$\psi_R^{\ E} = -\gamma_5^{\ E} \psi_R^{\ E}$$

if γ_5^E is defined in the usual way:

$$[\gamma_\mu , \gamma_\nu]_+^{\ E} = 2\delta_{\mu\nu} \quad , \quad \gamma_5^{\ E} = (\gamma_1\gamma_2\gamma_3\gamma_4)^E \qquad (2.50)$$
$$\gamma_0^{\ M} = \gamma_4^{\ E} \quad , \quad \vec{\gamma}^M = i\vec{\gamma}^E$$

Mappability onto S_4 is a stronger condition than $S[A] < \infty$. Consider the projected field-strength tensor $\hat{F}_{abc}$ given by[21,27]

$$\hat{F}_{\mu\nu\lambda} = \tfrac{1}{2}(1 + x^2)\left[x_\mu F_{\nu\lambda} + x_\nu F_{\lambda\mu} + x_\lambda F_{\mu\nu} \right] \qquad (2.51)$$
$$\hat{F}_{\mu\nu 5} = \tfrac{1}{2}(1 + x^2)\left[\tfrac{1}{2}(1 - x^2) F_{\mu\nu} - x_\mu x_\lambda F_{\nu\lambda} + x_\nu x_\lambda F_{\mu\lambda} \right]$$

and assume that the gauge-invariant quantity

$$\hat{F}(\xi)^2 = \hat{F}_{abc}\hat{F}_{abc} = \tfrac{3}{16}(1 + x^2)^4 F_{\mu\nu}F_{\mu\nu} \qquad (2.52)$$

is smooth everywhere on S_4, including the S pole and its neighbourhood. In Euclidean space, this corresponds to the asymptotic behaviour

$$x^8 F(x)^2 \longrightarrow C \quad , \quad (x_\mu \longrightarrow \infty) \qquad (2.53)$$

where C is zero or is a finite constant which does not depend on the direction in which x_μ is taken to ∞.

If Eq. (2.53) is not satisfied for finite $S[A]$, $\hat{F}^2(\xi)$ has an integrable singularity at the S pole. The problem is easily analysed in the Abelian case by slightly smearing $\hat{A}$ to obtain an approximating smooth potential. However, this method fails for non-Abelian fields because of the non-linear relation between $\hat{F}$ and $\hat{A}$.

3. SOLUTIONS OF THE EQUATION OF MOTION

All known finite-action solutions of the Euclidean equation of motion (2.23) are self-dual or anti-self-dual, or have a trivial direct-product structure such as

$$A.\tau \quad = \quad \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \tag{3.1}$$

for the gauge group $\mathcal{G}_1 \otimes \mathcal{G}_2$, where A_1 and A_2 are (anti-)self-dual solutions for $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively.

The simplest SU(2) solution, the instanton[1])

$$A^a_\mu(x) \quad = \quad \frac{2}{g}\ \eta_{a\mu\sigma}\,(x - x_0)_\sigma \Big/ \left[(x - x_0)^2 \ + \ \lambda^2\right] \tag{3.2}$$

involves a tensor $\eta_{a\mu\sigma}$ given by[2,3])

$$\begin{aligned} \eta_{a\mu\sigma} &= -\ \eta_{a\sigma\mu} \quad , \quad (a = 1,2,3\ ;\ \mu,\sigma = 1,\ldots,4) \\ \eta_{aij} &= \varepsilon_{aij} \quad , \quad (i,j = 1,2,3) \\ \eta_{ai4} &= \delta_{ai} \end{aligned} \tag{3.3}$$

and five explicit arbitrary parameters, the instanton position $(x_0)_\mu$ and its size λ. The solution (3.2) is self-dual, so ν is positive. Direct calculation of the volume integral (2.6), or the observation

$$g\, A_\mu(x)\,.\,\tau \quad \longrightarrow \quad i\, G_1^{-1}\partial_\mu\, G_1 \quad , \quad (x_\mu \rightarrow \infty) \tag{3.4}$$

[where G_1 is defined in Eq. (2.37)] leads to the result $\nu = +1$.

Of course, any solution which is gauge-equivalent to (3.2) is equally acceptable. The solution (3.2) corresponds to the Landau gauge $\partial_\mu A_\mu = 0$. When quantum amplitudes are being computed in this gauge, it must be remembered that the choice (3.3) for the tensor $\eta_{a\mu\sigma}$ picks out a particular direction in colour SU(2) space. Any x-independent rotation in colour space yields another Landau-gauge solution. Thus the general Landau-gauge solution contains three additional parameters $\vec{\Omega}$

$$\left(A_\mu . \tau\right)_{8\ \text{para.}} \quad = \quad U(\vec{\Omega})\left(A_\mu . \tau\right)_{5\ \text{para.}} U(\vec{\Omega})^{-1} \tag{3.5}$$

where $U(\vec{\Omega})$ is an x-independent rotation through an angle $\vec{\Omega}$ in SU(2) group space.

The anti-instanton is obtained by replacing $\eta_{a\mu\sigma}$ by $\bar{\eta}_{a\mu\sigma}$ in Eq. (3.2):

$$\bar{\eta}_{a\mu\sigma} = -\,\bar{\eta}_{a\sigma\mu} \quad , \quad \bar{\eta}_{aij} = \varepsilon_{aij} \quad , \quad \bar{\eta}_{ai4} = -\,\delta_{ai} \tag{3.6}$$

It is anti-self-dual, with $\nu = -1$. Equations (2.20) and (2.21) imply

$$S[A] \quad = \quad 8\pi^2/g^2 \tag{3.7}$$

for the instanton or anti-instanton.

Solutions for which ν is an arbitrary integer can be generated from the prescription[28])

$$A_\mu^a = -g^{-1}\,\bar{\eta}_{a\mu\sigma}\,\partial_\sigma \ln\rho(x) \tag{3.8}$$

Self-duality (or anti-self-duality if $\bar{\eta}_{a\mu\sigma}$ is replaced by $\eta_{a\mu\sigma}$) implies

$$\rho^{-1}\partial^2\rho = 0 \tag{3.9}$$

i.e. $\partial^2\rho$ vanishes everywhere except at the singularities of $\rho(x)$. The general solution of Eq. (3.9),

$$\rho(x) = \sum_{i=0}^{|\nu|} \lambda_i^2/(x-x_i)^2 \quad , \quad (\lambda_0 = 1) \tag{3.10}$$

due to Jackiw, Nohl and Rebbi[29]), includes special cases found earlier by Witten[30]) and 't Hooft[31]). The result has singularities at $x = x_i$, but these are pure-gauge terms which can be gauge-transformed to infinity to obtain a completely smooth representation in Euclidean space[32]). For $|\nu| = 1, 2$, some of the parameters (x_i, λ_i) correspond to gauge transformations. The number of independent gauge-invariant parameters in the solution (3.10) is[29]) 5 for $|\nu| = 1$, 13 for $|\nu| = 2$, and $(5|\nu| + 4)$ for $|\nu| \geq 3$.

Equation (3.8) also provides an interesting connection between gauge theory and scalar field theory with quartic coupling. The equation of motion (2.23) implies the restriction

$$\rho^2\,\partial_\lambda(\rho^{-3}\partial^2\rho) = 0 \tag{3.11}$$

so non-singular solutions $\rho(x)$ of (3.11) obey the equation of motion of an interacting scalar field[28]):

$$\partial^2\rho + C\rho^3 = 0 \quad , \quad (C = \text{constant}) \tag{3.12}$$

The only known non-singular solution of Eq. (3.12)

$$\rho(x) = (8\lambda^2/C)^{1/2}\,[(x-x_0)^2 + \lambda^2]^{-1} \tag{3.13}$$

yields the anti-instanton in Eq. (3.8).

An application of the index theorem[24,33]) shows that the most general (anti-) self-dual SU(2) solution mappable onto S_4 must involve $(8|\nu| - 3)$ gauge-invariant parameters. The intuitive reason for this[34,35]) is that for an appropriate choice of its parameters, the general solution should approximate a superposition of $|\nu|$ widely separated (anti-) instantons, each with a position x_i, a size λ_i, and an SU(2) gauge orientation $\vec{\Omega}_i$; of these $8|\nu|$ parameters, three correspond to the over-all SU(2) gauge orientation, so there remain $8|\nu| - 3$ gauge-invariant parameters. This means that the prescription (3.8) is insufficiently general for $|\nu| \geq 3$.

A prescription for the general case which includes all compact classical Lie groups

$$\mathcal{G} = SO(n),\ SU(n),\ Sp(n) \tag{3.14}$$

has been derived from fibre bundle theory by Drinfeld, Manin, Atiyah and Hitchin[7,8]*). Let τ^a generate $\mathcal{G}$ in the fundamental representation. Then all self-dual fields mappable onto S_4 are given by the prescription

$$\begin{aligned} -ig A_\mu^a \tau^a &= M^\dagger(x)\, \partial_\mu M(x) \quad , \\ M^\dagger(x)\, M(x) &= I \quad , \\ M^\dagger(x)\, \Delta(x) &= 0 \end{aligned} \tag{3.15}$$

where $M(x)$ and $\Delta(x)$ are rectangular matrices with real, complex, or quaternionic elements (depending on $\mathcal{G}$) subject to the constraints

$$\begin{aligned} \Delta(x) &= B + Cx \quad , \quad (B, C = x_\mu\text{-independent matrices}), \\ x &= x_4 + i\vec{\sigma}.\vec{x} \quad , \\ [\Delta^\dagger\Delta, \vec{\sigma}] &= 0 \quad , \\ (\Delta^\dagger\Delta)^{-1} &\text{ exists} \end{aligned} \tag{3.16}$$

(For anti-self-dual fields, replace x by x^+.) The matrices $\frac{1}{2}\vec{\sigma}$ generate an SU(2) subgroup acting on the domain of Δ. As predicted previously[20], the result for $A_\mu^a(x)$ is a rational function of x_μ.

Remarks:

i) The dependence of the result on the instanton number $|\nu|$ arises when one specifies the numbers of rows and columns of M and Δ. For the special case $\mathcal{G} = SU(2)$, $M(x)$ is a column vector of $(|\nu| + 1)$ quaternions [i.e. a $(2|\nu| + 2) \times 2$ complex matrix] and $\Delta(x)$ is a $(|\nu| + 1) \times |\nu|$ array of quaternions.

ii) Gauge transformations $G(x)$ in Eq. (2.1) correspond to the transformation

$$M(x) \longrightarrow M(x)\, G(x) \tag{3.17}$$

iii) The number of gauge-invariant parameters of A_μ^a can be counted by parametrizing the constraint equations for Δ and taking the gauge freedom (3.17) into account. Of course, the answer agrees with the prediction of the index theorem. However, it is not clear that the dependence of A_μ^a on its parameters can be

---

*) Details of the prescription and the construction of fermionic solutions (2.46) and propagators are explained in recent papers by Corrigan, Fairlie, Templeton and Goddard[36], Osborn[37], and Christ, Weinberg and Stanton[38].

expressed in terms of simple functions as in Eq. (3.10). The general case seems to involve the solution of coupled quadratic equations.

The proof[8] of the complete generality of the construction (3.15) and (3.16) is non-trivial, but it is not difficult to check its self-duality. Consider the operators

$$\begin{aligned} P &= MM^\dagger \\ \bar{P} &= I - \Delta(\Delta^\dagger\Delta)^{-1}\Delta^\dagger \end{aligned}$$

which project onto spaces orthogonal to Δ:

$$P^2 = P \quad , \quad \bar{P}^2 = \bar{P}$$
$$P\Delta = 0 = \bar{P}\Delta$$

In fact, the identities

$$\bar{P}P = \bar{P} \quad , \quad P\bar{P} = P$$

imply that these spaces are identical, so P and $\bar{P}$ must be identical operators[36,38]:

$$I - MM^\dagger = \Delta(\Delta^\dagger\Delta)^{-1}\Delta^\dagger \tag{3.18}$$

As a result, the field-strength tensor can be written in the form

$$\begin{aligned} -igF_{\mu\nu}{}^a\tau^a &= \partial_\mu M^\dagger(I - MM^\dagger)\partial_\nu M - \{\mu \leftrightarrow \nu\} \\ &= M^\dagger\partial_\mu\Delta(\Delta^\dagger\Delta)^{-1}\partial_\nu\Delta^\dagger M - \{\mu \leftrightarrow \nu\} \end{aligned} \tag{3.19}$$

Equation (3.16) implies $\partial_\mu\Delta = C\partial_\mu x$, where $\partial_\mu x$ commutes with $(\Delta^\dagger\Delta)^{-1}$:

$$-igF_{\mu\nu}{}^a\tau^a = M^\dagger C(\partial_\mu x\,\partial_\nu x^\dagger - \partial_\nu x\,\partial_\mu x^\dagger)(\Delta^\dagger\Delta)^{-1}C^\dagger M \tag{3.20}$$

The tensor

$$\partial_\mu x\,\partial_\nu x^\dagger - \partial_\nu x\,\partial_\mu x^\dagger = 2i\sigma^a\eta_{a\mu\nu}$$

is self-dual and therefore so is $F^a_{\mu\nu}$.

A particularly elegant result has been found for the propagator $\Delta(x,y)$ of a spin-0 field belonging to the fundamental representation of the gauge group. If the covariant derivative D_μ is constructed from the field given by Eqs. (3.15) and (3.16), the solution of the equation

$$-D^2\Delta(x,y) = \delta(x-y) \tag{3.21}$$

is[36,38]

$$\Delta(x,y) = M^\dagger(x)M(y)/4\pi^2(x-y)^2 \tag{3.22}$$

The construction of spin-½ and spin-1 propagators from $\Delta(x,y)$, together with special cases of (3.22), were given previously by Brown et al.[39].

The problem of finding Euclidean finite-action SU(2) solutions which are neither self-dual nor anti-self-dual seems to be very difficult. In fact, it has not even been established whether such configurations exist or not. All that has been proved[40] so far is that they do not exist in the function space neighbourhood of (anti-)self-dual fields. If solutions exist outside this neighbourhood, they probably correspond to saddle points in the action, because the action of an instanton-anti-instanton pair can be reduced below 16 π^2/g^2 by deformation[41]. The latest suggestion[42] is to search for solutions by regarding the second-order equation (2.23) as a projection of first-order equations in an eight-dimensional space.

4. WKB APPROXIMATION

Classical configurations become important in quantum theory when the path-integral approach[17,43] to quantum amplitudes is considered. A quantum amplitude is a sum over classical paths weighted by the factor

$$\exp i\left[\text{Action}\right]$$

When dealing with Green's functions, it is often convenient to rotate to Euclidean space. Field components and γ-matrices [Eq. (2.50)] are slightly changed, because the Lorentz group O(3,1) becomes the orthogonal group O(4):

$$\begin{aligned} &\text{quarks}: \quad && (q, \bar{q}) \quad \longrightarrow \quad (q, q^{\dagger}) \\ &\text{gluons}: \quad && (A_0, \vec{A}) \quad \longrightarrow \quad (iA_4, \vec{A}) \end{aligned} \tag{4.1}$$

This tends to improve the convergence of expressions for amplitudes, because the weighting factor becomes

$$\exp -\left[\text{Euclidean} \quad \text{Action}\right]$$

and typically the Euclidean action is bounded below. [Gravity is an exception[44].]

Thus the problem considered by 't Hooft in his definitive paper[3]*) was the computation of Euclidean QCD Green's functions

$$\left\langle \prod_k O_k(x_k) \right\rangle = \frac{\int [dA\, dq^{\dagger} dq]\, \prod_k O_k\left[A(x_k),\, q(x_k)\,,\, q^{\dagger}(x_k)\right]\, e^{-S}\, \delta[f(A)]\,\Delta_f[A]}{\int [dA\, dq^{\dagger} dq]\;\; e^{-S}\, \delta[f(A)]\,\Delta_f[A]} \tag{4.2}$$

for which paths $A^a_\mu(x)$ with non-zero topological charge ν are important. In Eq. (4.2),

$$S \quad = \quad S[A, q, q^{\dagger}] \tag{4.3}$$

*) See also Refs. 45-47.

is the gauge-invariant Euclidean action of QCD. The denominator is included so that the vacuum state for the corresponding Minkowskian amplitudes

$$T \langle vac | \prod_k O_k(x_k) | vac \rangle$$

has unit normalization:

$$\langle vac | vac \rangle \quad = \quad 1 \tag{4.4}$$

The operators $O_k(x_k)$ can be simple or composite operators constructed from the fields (A,q,q^+). Usually one is interested in gauge-invariant observables such as currents and their divergences. The factors $\delta[f]\Delta_f$ arise from the need to avoid summing over unphysical directions in A_μ-space corresponding to gauge transformations. The introduction of a δ-function $\delta[f]$ to fix the gauge $f[A] = 0$ must be compensated by a Jacobian factor $\Delta_f[A]$, the Faddeev-Popov determinant. As in perturbation theory, calculations are simpler if $\delta[f]\Delta_f$ is written in the form

$$\int [dc^+ dc] \; exp \; - [S_{gauge-fix} \; + \; S_{ghost}]$$

where $[c,c^+]$ are the ghost fields of Feynman, Faddeev and Popov et al.[48]. [The standard prescription was derived for perturbative use, but works for WKB calculations about smooth configurations. In a study of the Coulomb gauge, Gribov[49] observed that modifications may arise in non-perturbative calculations and become important for the non-WKB problem of quark confinement. In WKB calculations, this phenomenon is merely a gauge artifact -- it arises only if the choice of gauge forces the relevant classical field to be discontinuous. See Ref. 50.]

The sum over paths of a real Bose field ϕ is obtained by considering a parametrization

$$\phi(x) \quad = \quad \sum_E z_E \, \phi_E(x) \tag{4.5}$$

where $\{\phi_E(x)\}$ is a complete orthonormal set:

$$\begin{aligned} \int d^4x \; \phi_E(x) \, \phi_{E'}(x) \quad &= \quad \delta_{EE'} \\ \sum_E \phi_E(x) \, \phi_E(x') \quad &= \quad \delta^4(x - x') \end{aligned} \tag{4.6}$$

This set can be regarded as defining orthogonal axes with unit length in function space. As the parameters z_E are varied over all real values, all paths ϕ are produced. Thus $\{z_E\}$ is a set of coordinates for an infinite-dimensional Cartesian space, where each point labels a path. The metric for path summation is given by a volume element in $\{z_E\}$ space,

$$[d\phi] \quad = \quad \prod_E \left(dz_E / \sqrt{2\pi} \right) \tag{4.7}$$

and does not depend on the choice of axes $\{\phi_E\}$. The factors $(2\pi)^{-1/2}$ which fix the over-all normalization are chosen arbitrarily, but this does not matter because the normalizing factor cancels in the expression (4.2).

The standard example is the "Gaussian" integral

$$I = \int [d\phi] \exp - \tfrac{1}{2} \int d^4x \; \phi(x) D_x \phi(x) \tag{4.8}$$

where D_x is a differential operator. Let $\{\phi_E\}$ be the set of eigenfunctions of D_x:

$$D_x \phi_E(x) = E \, \phi_E(x) \tag{4.9}$$

Then Eqs. (4.5)-(4.7) imply

$$\begin{aligned} I &= \prod_E \left\{ \int_{-\infty}^{\infty} dz_E \, (2\pi)^{-1/2} \exp - \tfrac{1}{2} E z_E^2 \right\} \\ &= (\det D)^{-1/2} \end{aligned} \tag{4.10}$$

where the determinant of the operator D is defined to be the product of its eigenvalues

$$\det D = \prod_E E \tag{4.11}$$

This example can be easily generalized to include a source $j(x)$ for ϕ:

$$I[j] = \int [d\phi] \exp - \int d^4x \left[\tfrac{1}{2} \phi(x) D_x \phi(x) + j(x) \phi(x) \right] \tag{4.12}$$

A shift of integration variable

$$\phi \longrightarrow \phi + D_x^{-1} j$$

yields the result

$$\begin{aligned} I[j] &= (\det D)^{-1/2} \exp \tfrac{1}{2} \int d^4x \; j(x) D_x^{-1} j(x) \\ &= (\det D)^{-1/2} \exp \tfrac{1}{2} \iint d^4x \, d^4y \; j(x) \, G(x,y) \, j(y) \end{aligned} \tag{4.13}$$

where

$$G(x,y) = \sum_E E^{-1} \phi_E(x) \phi_E(y) = D_x^{-1} \delta(x-y) \tag{4.14}$$

is the propagator of the operator D_x. The result can be further generalized to include a potential $V(\phi)$:

$$\begin{aligned} &\int [d\phi] \exp - \int d^4x \left[\tfrac{1}{2} \phi D \phi + j\phi + V(\phi) \right] \\ &\qquad = \exp - \int d^4x \; V[-\delta/\delta j(x)] \; I[j] \end{aligned} \tag{4.15}$$

There are two circumstances which result in modifications of these formulas. Sometimes D_x has zero eigenvalues: then integrations $\int dz_0$ associated with the corresponding eigenfunctions $\phi_0(x)$ must be handled separately (see below). The other problem is the presence of infinities associated with infinite products and sums Π_E, $\sum_E$. Some of these infinities do not cancel between the numerator and denominator of Eq. (4.2). Thus regulation and renormalization are necessary: the infinities have to be absorbed in the definition of the mass parameters, coupling constant, and normalizations of the operators $O_k(x_k)$. 't Hooft[3] adds Pauli-Villars regulator fields. An alternative is Hawking's zeta function regularization[47,51]

$$\begin{aligned} \prod_E E &= \exp \sum_E \ln E \\ &\longrightarrow \exp - \frac{\partial}{\partial s} \sum_E E^{-s} \end{aligned} \tag{4.16}$$

For Re s sufficiently large, the series $\sum_E E^{-s}$ converges to a generalized zeta function

$$\zeta(s) = \mathrm{Tr}\ D^{-s} \tag{4.17}$$

which can be analytically continued to s = 0 to yield the renormalized determinant

$$(\det D)_{ren} = \exp - \zeta'(0) \tag{4.18}$$

For fermions, an expansion similar to (4.5) is used, but the ordinary integration variables $\{z_E\}$ are replaced by independent Grassmann variables (ξ_E, η_E):

$$\begin{aligned} \Psi(x) &= \sum_E \xi_E u_E(x) \\ \Psi^\dagger(x) &= \sum_E \eta_E u_E^\dagger(x) \\ [d\Psi^\dagger d\Psi] &= \prod_E (d\eta_E d\xi_E) \end{aligned} \tag{4.19}$$

The functions $u_E(x)$ are c-numbers with spinor indices. Grassmann variables anticommute,

$$[\xi_E, \xi_{E'}]_+ = [\xi_E, \eta_{E'}]_+ = [\eta_E, \eta_{E'}]_+ = 0 \tag{4.20}$$

and in particular, they are nilpotent:

$$\xi_E^2 = 0 = \eta_E^2 \tag{4.21}$$

Integration over these variables is fixed by the rules*)

$$\int d\xi\ 1 = 0 \quad , \quad \int d\xi\ \xi = 1 \tag{4.22}$$

where infinitesimals $d\xi$, $d\eta$ are also Grassmann variables, e.g.,

$$[d\xi, \eta]_+ = 0 \tag{4.23}$$

---

*) The standard reference is the textbook by Berezin[52].

The fermionic "Gaussian" integral is given by

$$I = \int [d\Psi^\dagger d\Psi] \ \exp - \int d^4x \ \Psi^\dagger(x) \, D_x \, \Psi(x)$$
$$= \prod_E \left\{ \int d\eta_E \, d\xi_E \ \exp - E \, \eta_E \xi_E \right\} \tag{4.24}$$

where $\{E\}$ is the set of eigenvalues of D_x. The nilpotent property and integration rules imply that the only non-zero contribution from the power-series expansion of $\exp -E\eta_E\xi_E$ is the linear term $-E\eta_E\xi_E$, so the result is simply

$$I = \prod_E E = \det D \tag{4.25}$$

The aim of the WKB method is to keep the Euclidean action S as small as possible so that e^{-S} is maximized. The gauge field A_μ is split into classical and "quantum" fields

$$A = A^{cl} + A^{qu} \tag{4.26}$$

where $\int [dAdq^+dq]$ is shifted to $\int [dA^{qu}dq^+dq]$; (in this context, "quantum" field means a dummy functional integration variable). The classical field is chosen to be a solution of the equation of motion (2.23), i.e. it corresponds to a stationary point of the action:

$$\delta S/\delta A \Big|_{A = A^{cl}} = 0 \tag{4.27}$$

Then, if $S[A^{cl} + A^{qu}, q, q^+]$ is expanded in powers of the quantum fields $Q = (A^{qu}, q, q^+)$, the term linear in Q is absent:

$$S = S[A^{cl}, 0, 0] + \tfrac{1}{2}\iint d^4x \, d^4y \ A^{qu}(x) \, A^{qu}(y) \left[\delta^2 S/\delta A(x)\delta A(y)\right]_{A = A^{cl}}$$
$$+ \int d^4x \ q^+(x) \left\{ \not{D}[A^{cl}] + M \right\} q(x) \tag{4.28}$$
$$+ O[Q^3] + O[Q^4]$$

The effective action associated with the choice of gauge is expanded in the same variables, with ghosts treated as quantum fields:

$$S_{gauge\text{-}fix} + S_{ghost} = S_g[A^{cl}, A^{qu}, c, c^+] \tag{4.29}$$

[S_g need not be a functional of the combination $(A^{cl} + A^{qu})$]. The choice of gauge-fixing term should correspond to the choice of gauge for A^{cl} so that S_g is quadratic in the quantum fields:

$$S_g = O[Q^2] + O[Q^3] + (rarely) \ O[Q^4] \tag{4.30}$$

Why not expand about classical fermion fields ψ^{cl}, ψ^{+cl}, as well as the classical gluon field A_μ^{cl}? The difference is that such fields would be Grassmann variables, not c-numbers:

$$\psi^{cl}(x) = u(x)\,\xi^{cl} \quad , \quad \psi^{+cl}(x) = u^{\dagger}(x)\,\eta^{cl}$$

$$u(x) = \text{c-number function}$$

The answer has to be a polynomial in the variables (ξ^{cl}, η^{cl}) because they obey Eq. (4.21) and are not integration variables:

$$\text{Integral} = I_0 + I_1\xi^{cl} + I_2\eta^{cl} + I_3\eta^{cl}\xi^{cl}$$

However, the answer is known to be a c-number, i.e. the rules for Grassmann integration cause I_1, I_2, and I_3 to vanish. Consequently, the correct answer is found by setting ξ^{cl}, η^{cl} to zero from the beginning. In all known cases, the classical field is a c-number*).

Thus the vacuum expectation value (4.2) is given by a sum over stationary points X (or "quasi-stationary" points, to be discussed below) of the action:

$$\left\langle \prod_k O_k \right\rangle = \frac{\sum_X \int [dQ] \left\{ \prod_k O_k \exp -(S + S_g) \right\}_X}{\sum_X \int [dQ] \left\{ \exp -(S + S_g) \right\}_X} \tag{4.31}$$

Here $\int [dQ]$ means $\int [dA^{qu}\, dq^{+} dq\, dc^{+} dc]$, and each label X refers to a particular stationary point with classical field A_X^{cl}:

$$\left\{ \prod_k O_k \right\}_X = \prod_k O_k [A_X^{cl} + A^{qu}, q, q^{\dagger}] \quad , \quad \text{etc.} \tag{4.32}$$

The weight factor is a product of exponentials

$$\exp -(S + S_g) = \exp - S[A^{cl}, 0, 0]\; e^{-O[Q^2]}\; e^{-O[Q^3]}\; e^{-O[Q^4]} \tag{4.33}$$

Gaussian integrals

$$\int [dQ]\; Q^{p} \exp - O[Q^2]$$

are obtained by expanding the exponential

$$\exp -\left\{ O[Q^3] + O[Q^4] \right\} = \text{Taylor series in } g \tag{4.34}$$

and considering each term of the expansion separately.

---

*) The situation may be different in Adler's theory[53] of "algebraic chromodynamics", where the classical theory contains non-commuting gluon fields. A procedure for integration over these fields has yet to be developed.

The absolute minimum of S is given by $A^{cl} = 0$, which means that the classical factor $\exp - S[A^{cl}]$ equals unity. The result is that the Taylor series (4.34) generates the usual perturbative expansion of $\langle \Pi_k \, O_k \rangle$.

The next minima occur at

$$S[A^{cl}, 0, 0] = 8\pi^2/g^2 \tag{4.35}$$

with A^{cl} equal to an instanton or anti-instanton configuration. The classical factor

$$\exp - S[A^{cl}, 0, 0] = \exp - 8\pi^2/g^2 \tag{4.36}$$

is now a non-trivial function of g which has no Taylor expansion about g = 0, so it cannot be generated from perturbation theory. The factor (4.36) multiplies a power series in g obtained from (4.34). The calculations of 't Hooft[2,3] concern the leading term of this series:

$$\exp - \left\{ O[Q^3] + O[Q^4] \right\} = 1 + O(g) \tag{4.37}$$

In principle, contributions from the remaining stationary points X should be included in the same way. However, there are some technical difficulties which have yet to be overcome. One of these is the lack of a proper analysis of stationary points for fields which are neither self-dual nor anti-self-dual (this was mentioned at the end of Section 3). Furthermore, a successful analysis of that problem would not necessarily suffice for the WKB problem. For example, let us suppose that (as many people suspect) it can be proved that all finite-action Euclidean solutions are self-dual or anti-self-dual. Then it is *not* correct to restrict contributions to (4.31) to minima associated with the general (anti-) self-dual solutions (3.15) -- there must be additional contributions from "quasi-stationary" points[5,14] of the action which correspond to instantons and anti-instantons being infinitely far apart. In particular, if the instanton contributes to $\langle \Pi'_k \, O_k \rangle$ and the anti-instanton to $\langle \Pi''_\ell \, O_\ell \rangle$, the amplitude $\langle (\Pi'_k \, O_k)(\Pi''_\ell \, O_\ell) \rangle$ must receive a contribution

$$\sim \left(\exp - 16\pi^2/g^2\right) \left\{ 1 + a_1 g^2 + a_2 g^4 + \ldots \right\}$$

from a quasi-stationary instanton-anti-instanton configuration in order to be able to satisfy the cluster property

$$\left\langle \prod_k{}' O_k(x_k) \, . \, \prod_\ell{}'' O_\ell(y_\ell) \right\rangle \sim \left\langle \prod_k{}' O_k(x_k) \right\rangle \left\langle \prod_\ell{}'' O_\ell(y_\ell) \right\rangle , \qquad \text{Euclidean } (x_k - y_\ell)^2 \longrightarrow \infty \tag{4.38}$$

The absence of a precise WKB analysis for multi-instantons has led to the suggestion[5,14] that the dilute-gas approximation[54] be used. The classical field A^{cl} in (4.26) is taken to be a sum of widely separated instantons $A^{(+)}$ and anti-instantons $A^{(-)}$:

$$A_{gas} = \sum_{i=1}^{n_+} A^{(+)}(x - x_i^+, \lambda_i^+) + \sum_{j=1}^{n_-} A^{(-)}(x - x_j^-, \lambda_j^-) \tag{4.39}$$

The main approximation is to ignore the $O[Q]$ term

$$\int d^4x \; A^{qu}(x) \, [\delta S/\delta A]_{A = A_{gas}}$$

in the action, i.e. to ignore the overlap of various $A^{(\pm)}$ where A_{gas} does not exactly satisfy the equation of motion. This makes sense if the sizes $\lambda_i^{\pm}$ are much smaller than the distances $|x_i - x_j|$. The topological charge ν is $(n_+ - n_-)$; the summations in Eq. (4.31) are given by

$$\sum_X \longrightarrow \sum_{n_+, n_-}$$

and the action $S[A_{gas}]$ tends to $8\pi^2(n_+ + n_-)/g^2$ in the infinitely dilute limit.

Remarks:

i) Use of the expansion (4.34) means that the result can be taken seriously only for small values of g. Thus the general result is that an amplitude possesses a perturbative expansion about $g \to 0$ which is improved by the inclusion of smaller non-perturbative WKB contributions. Exceptions will be considered in Section 5.

ii) The condition

$$S[A^{cl}, 0, 0] < \infty \tag{4.40}$$

is obviously essential for WKB calculations. However, one can imagine other configurations A_∞ being dominant for strong interactions where g is not asymptotically small. The functional metric about A_∞ could be infinitely larger than the Gaussian metric*) about A^{cl} and hence compensate the zero from $\exp - S[A_\infty]$:

$$[\infty \text{ metric factor}] \; \exp - \infty \overset{?}{=} \text{finite} \tag{4.41}$$

Thus theorems for which (4.40) is a necessary condition (e.g. the index theorem and the result ν = integer) are applicable only within the WKB approximation.

---

*) This remark is entirely separate from the fact (Ref. 55, Appendix C) that $S[A^{cl} + A_{Gauss}]$ diverges for non-zero contributions to the Gaussian metric.

Now there is a slight complication due to the fact that each classical solution depends on several continuous parameters Λ:

$$A^{c\ell} = A^{c\ell}(x, \Lambda) \tag{4.42}$$

For example, the instanton depends on five gauge-invariant parameters (x_0, λ) and three global SU(2) gauge parameters. This means that each stationary point of the action is part of a continuum of stationary points labelled by Λ:

$$S[A^{c\ell}(x, \Lambda)] = \text{independent of } \Lambda \tag{4.43}$$

In Fig. 4, the base of the trough corresponds to the set of functions $\{A^{cl}(x, \Lambda)\}$ for various values of Λ. Small oscillations of the quantum field in this trough are Gaussian if they are perpendicular to the base. However, it is obvious that the flat direction along the base is not Gaussian; it does not contribute to $\int d^4x\, A^{qu} D\, A^{qu}$, the $O[(A^{qu})^2]$ term in the action. Thus this direction corresponds to a zero eigenvalue of D which causes the result of the functional integration $(\det D)^{-1/2}$ to be ill-defined.

The systematic procedure for dealing with this problem is the method of collective coordinates[56,57]. If there is just one parameter Λ, the method can be summarized as follows. The direction in function space along the base of the trough is given by

$$A^{c\ell}(x, \Lambda + d\Lambda) - A^{c\ell}(x, \Lambda) = a_0(x, \Lambda)\, dz_0 \tag{4.44}$$

where

$$a_0(x, \Lambda) = \frac{\partial}{\partial\Lambda} A^{c\ell}(x, \Lambda) \Big/ \left[\int d^4x \left\{ \frac{\partial}{\partial\Lambda} A^{c\ell}(x, \Lambda) \right\}^2 \right]^{1/2} \tag{4.45}$$

is the zero-eigenvalue member of the orthonormal set $\{a_E(x)\}$ used in the expansion of the quantum field

$$A^{qu}(x) = z_0\, a_0(x) + \sum_{E \neq 0} z_E\, a_E(x) \tag{4.46}$$

The integral over A^{qu} is split into a Gaussian $E \neq 0$ part and an ordinary integral over z_0:

$$\int [dA^{qu}] = \int dz_0 (2\pi)^{-1/2} \int [dA^{qu}]_{E \neq 0} \tag{4.47}$$

Then the integration variable z_0 is changed to Λ:

$$\begin{aligned} dz_0 &= a_0^{-1} \frac{\partial}{\partial\Lambda} A^{c\ell}\, d\Lambda \\ &= \left[\int d^4x \left\{ \partial A^{c\ell} / \partial\Lambda \right\}^2 \right]^{1/2} d\Lambda \end{aligned} \tag{4.48}$$

This affects the coupling constant dependence of the answer, because the factor $[\ldots]^{1/2}$ is proportional to g^{-1}.

If there are p independent parameters Λ_i, there are p zero modes with p parameters z_{oi} in the expansion of A^{qu}. Let us choose the parameters Λ_i such that $\partial A^{cl}/\partial\Lambda_i$ and $\partial A^{cl}/\partial\Lambda_j$ are *orthogonal* for $i \neq j$. Then each zero mode can be handled as above with the result

$$\int[dA^{qu}] \quad = \quad (2\pi)^{-p/2}\int d^p\Lambda\left[\prod_{i=1}^{p}\int d^4x\left\{\frac{\partial}{\partial\Lambda_i}A^{cl}(x,\vec{\Lambda})\right\}^2\right]^{1/2}\int[dA^{qu}]_{E\neq 0} \tag{4.49}$$

so the zero modes contribute a factor $O(g^{-p})$ to the amplitude.

The integration over global gauge rotations involves a subtlety[3,47,57]. The conventional gauge-fixing procedure is a collective-coordinate method for zero modes associated with local gauge transformations, so it must be checked that these are orthogonal to the modes $\partial A^{cl}/\partial\Lambda$. This is correct for the Landau gauge

$$\partial.A \quad = \quad \partial.A^{qu} \quad = \quad 0 \tag{4.50}$$

due to its invariance under global gauge transformations and changes in the gauge-invariant parameters of A^{cl} (with $\partial.A^{cl} = 0$). If other gauges are used [because of their convenience in the computation of determinants (4.11)], this orthogonality property need not be valid and the Faddeev-Popov procedure has to be combined with the collective coordinates for the Λ_i parameters. In particular, the background-field gauge

$$\partial_\mu A_\mu^{qu} \quad + \quad g\,A_\mu^{cl} \times A_\mu^{qu} \quad = \quad 0 \tag{4.51}$$

must be handled with great care [duly exercised by 't Hooft in Section 11 of Ref. 3, where the problem was circumvented by comparing (4.50) and (4.51)].

Thus a single instanton contributes

$$g^{-C}\left(\exp - 8\pi^2/g^2\right)\left\{\text{Taylor series in } g\right\}$$

to the asymptotic $g \to 0$ expansion of an amplitude. The power C depends on the gauge group. In general, many parameters $\vec{\Omega}$ are needed to specify the orientation of an instanton in group space:

$$\vec{\tau}.\vec{A}_\mu(x-x_0,\lambda,\Omega) \quad = \quad \frac{2}{g}\,e^{i\vec{\tau}.\vec{\Omega}}\,\eta_{\alpha\mu\kappa}\,\sigma^\alpha(x-x_0)_\kappa\,e^{-i\vec{\tau}.\vec{\Omega}}\Big/\left[(x-x_0)^2 + \lambda^2\right] \tag{4.52}$$

Here $\sigma^a/2$ (a = 1, 2, 3) generate an SU(2) subgroup of the gauge group $\mathcal{G}$. For $\mathcal{G}$ = SU(2), there are eight parameters:

$$C \quad = \quad 8 \quad , \qquad \left(\mathcal{G} = SU(2)\right) \tag{4.53}$$

For SU(3), there is one direction, colour hypercharge τ_8, which commutes with

$$(\tau_1, \tau_2, \tau_3) \quad = \quad \tfrac{1}{2}(\sigma_1, \sigma_2, \sigma_3)$$

so there are only seven $\vec{\Omega}$ parameters in the general solution (4.52):

$$C = 12 \quad , \quad \left(\mathcal{G} = SU(3) \right) \tag{4.54}$$

The result of including contributions from multi-instantons and anti-instantons is a double asymptotic series for the $g \to 0$ behaviour of Green's functions:

$$\left\langle \prod_k 0_k \right\rangle_{g \to 0} \sim \sum_{m,n \geq 0} a_{mn} \left\{ g^{-C} \exp - 8\pi^2/g^2 \right\}^m g^n \tag{4.55}$$

The coefficients a_{on} are generated by the Feynman diagrams of ordinary perturbation theory, so they are precisely defined. The same cannot be said (yet) for the remaining WKB coefficients a_{mn} ($m \geq 1$): there is trouble with the collective-coordinate integration over the size λ of the (anti-) instanton[2,3].

The problem arises in the following way. After computing regulated determinants, renormalizing, and carrying out all collective-coordinate integrations except that over λ, we obtain an amplitude $a(\lambda)$ for fixed instanton size λ which depends on the renormalized coupling constant g, mass parameters M_i, and a dimensional parameter μ which specifies the renormalization prescription. Consider the special case of a momentum-independent amplitude with $M_i = 0$ and no wave-function renormalization. Assuming renormalization-group invariance for fixed size λ, we have

$$a(\lambda, g, \mu) = \lambda^{-d} F\left[\bar{g}(\lambda\mu)\right] \tag{4.56}$$

where d is the mass dimensionality of $a(\lambda)$ and F is a dimensionless function of the running coupling constant $\bar{g}$ associated with the Callan-Symanzik β-function

$$\int_g^{\bar{g}(\lambda\mu)} dx/\beta(x) = -\ln(\lambda\mu) \tag{4.57}$$

The final step should be to carry out the integration over λ indicated by Eq. (4.49) with

$$\left[\int d^4x \left(\partial A^{cl}/\partial\lambda \right)^2 \right]^{1/2} = 4\pi/|g| \tag{4.58}$$

i.e. to compute $\int_0^\infty d\lambda\, a(\lambda)$. However, this integral diverges at*) ∞ when the lowest-order WKB contribution [with $\beta(x) \to -bx^3$]

$$F[\bar{g}] \sim (\text{constant})\; \bar{g}^{-C} \exp - 8\pi^2/\bar{g}^2 \tag{4.59}$$

is substituted in Eq. (4.56)[2,3,5].

---

*) If the number of quark flavours is not too large.

The conventional and possibly correct conclusion is that this difficulty is unavoidable and can be overcome only by successfully computing strong-coupling amplitudes. Certainly, if the prescription described above is the correct one, it is obvious that for large sizes λ, Eq. (4.59) is not correct and must be replaced by an expression valid for large $\bar{g}$.

What is not obvious (in my opinion) is that the scale-invariant zero mode $\partial A^{cl}/\partial\lambda$ is being handled correctly. The existence of continuously degenerate minima of the action due to zero modes means that the functional integral is ambiguous. This ambiguity can be resolved only if a new principle or assumption is externally imposed on the quantum amplitude. The principle underlying the method of collective coordinates is that symmetries broken by the parameters Λ in $A^{cl}(x, \Lambda)$ should be restored. Thus collective coordinates are appropriate when it is desired that quantum amplitudes should satisfy a symmetry principle of the classical theory. This is the case for the zero modes $\partial A^{cl}/\partial x_0$, $\partial A^{cl}/\partial\vec{\Omega}$. We certainly want quantum amplitudes to respect translation and global gauge invariance, so there is good reason*) to believe that the use of the corresponding collective coordinates is correct. The situation is different for the mode $\partial A^{cl}/\partial\Lambda$: classical scale invariance is not a symmetry of quantum amplitudes. Scale invariance is broken when μ-dependent renormalization counterterms are introduced.

In the absence of a suitable quantization procedure to decide the issue, we must be content with a list of possibilities.

For the conventional prescription, let us assume that strong interaction effects cause $\int_0^\infty d\lambda\ a(\lambda)$ to converge. Then the small g-dependence of amplitudes is not given by Eq. (4.55). For example, the amplitude

$$\int_0^\infty d\lambda\ a(\lambda) \ = \ \mu^{d-1}\int_0^{g_\infty} dy\ F(y)\,\beta(y)^{-1}\exp\left[(d-1)\int_g^y dx/\beta(x)\right] , \qquad (4.60)$$
$$\left(\ \bar{g}(\lambda\mu) \longrightarrow g_\infty\ ,\quad \lambda \longrightarrow \infty\ \right)$$

satisfies the equation

$$\left[\mu\,\partial/\partial\mu \ + \ \beta(g)\,\partial/\partial g\ \right]\int_0^\infty d\lambda\ a(\lambda) \ = \ 0 \qquad (4.61)$$

so it can be written[58])

$$\left|\int_0^\infty d\lambda\ a(\lambda)\right| \ = \ \mu^{d-1}\ \exp - (d-1)\int_\kappa^g dx/\beta(x) \qquad (4.62)$$

where κ is a (g,μ) independent constant. Thus the semiclassical factor $\exp - 8\pi^2/g^2$ is changed by the size integration to

*) In the theory of solitons, collective coordinates for translations have been checked for consistency with quantization in the one-soliton sector[56]).

$$\int_0^\infty d\lambda\, a(\lambda) \sim (\text{constant})\, \mu^{d-1} \exp - (d-1)/2bg^2 \quad , \quad (g \to 0) \qquad (4.63)$$

where b is the one-loop coefficient of the β-function

$$\beta(x) \sim - bx^3 + O(x^5) \qquad (4.64)$$

A purely semiclassical calculation at small g does not appear to be possible. (For a one-dimensional analogue, see Ref. 59.)

Alternatively, quantization may permit or require the inclusion of a certain weight function $\omega(\lambda)$ which cuts off the size integral at large λ and hence permits a purely semiclassical approximation at small g:

$$\text{Amplitude} = \int_0^\infty d\lambda\, \omega(\lambda)\, a(\lambda) \qquad (4.65)$$

For example, it may not be correct to treat renormalization and the size integration separately. A weight function of this type would have to depend on a characteristic size λ_0 which is renormalization-group invariant. There appear to be two cases:

i) The cut-off size λ_0 is proportional to μ^{-1} but is otherwise arbitrary; i.e. the pure number c in the formula

$$\lambda_0 = \mu^{-1} \exp \int_c^g dx/\beta(x) \qquad (4.66)$$

is unspecified, apart from the requirement that $\beta(x)$ should not vanish on $0 < x \leq c$.

ii) The non-perturbative parameter λ_0 is completely arbitrary. Both λ_0 and μ set the scale of hadronic amplitudes.

In case (i), the small g-dependence is changed as in Eq. (4.63). Only in case (ii) is it possible to maintain the semiclassical form (4.55). Of course, the suggestion that arbitrary renormalization-group invariant constants (λ_0 or c) are permitted is pure speculation.

The problem is even more complicated for multi-instantons. The difficulties with size integrations are also present in integrations over distances between instantons and anti-instantons.

Despite this uncertainty, there are situations in which it can be argued that practical calculations with the dilute-gas approximation are feasible. To justify the approximation, one must argue that short-distance effects (and hence small instanton sizes) are important for the particular process being considered. The cleanest applications involve amplitudes in which the external momenta q are non-exceptional and large, because it has been found[60,61] that non-exceptional momenta effectively cut off size integrals at $\lambda = O(q^{-1})$. For such processes, it is reasonable to hope that the importance of the theoretical ambiguities discussed above is minimal. Specific applications are listed in Section 10.

5. FERMIONIC AMPLITUDES

It was observed by 't Hooft[2] that some fermionic amplitudes which vanish in perturbation theory are non-zero in the WKB approximation and contribute to the series (4.55):

$$a_{on} = 0 \quad ; \quad a_{mn} \neq 0 \quad , \quad (\text{certain } m \neq 0) \tag{5.1}$$

This is due to the existence of zero modes of the fermionic covariant derviative $\not{D}$.

The simplest example is the unique normalizable solution[2]

$$u_0(x) = \left[(x - x_0)^2 + \lambda^2 \right]^{-3/2} \begin{pmatrix} \chi \\ \chi \end{pmatrix} \tag{5.2}$$

of the quark equation of motion

$$\left(\not{\partial} - ig \not{A}.\tau \right) u_0 = 0 \tag{5.3}$$

where A_μ^a is the SU(2) instanton (3.2). The column vector in Eq. (5.2) displays the positive and negative frequency components of the spinor u_0. It corresponds to the γ-matrix convention [Eqs. (2.48)-(2.50)]

$$\vec{\gamma}^E = \begin{pmatrix} 0 & -i\vec{\sigma} \\ i\vec{\sigma} & 0 \end{pmatrix} , \qquad \gamma_4^E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \qquad \gamma_5^E = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \tag{5.4}$$

so u_0 is right-handed:

$$\gamma_5^E u_0 = - u_0 \tag{5.5}$$

The spinor χ has two Dirac and two gauge components subject to the constraint

$$\left(\frac{1}{2} \vec{\sigma} + \vec{\tau} \right) \chi = 0 \tag{5.6}$$

For the anti-instanton, the spinor u_0 is left-handed:

$$\begin{pmatrix} \chi \\ \chi \end{pmatrix} \longrightarrow \begin{pmatrix} \chi \\ -\chi \end{pmatrix} \tag{5.7}$$

If there are N massless quark flavours, there are N orthonormal solutions $u_{oi}(x)$ $(i = 1, \ldots, N)$, i.e. N zero eigenvalues of $\not{D}$. The corresponding equations for multi-instantons have also been solved[36,37,62].

't Hooft considered the fermionic determinant

$$\det(\not{D} + J) = \prod_{\varepsilon} \varepsilon[J] \quad , \quad (A_\mu = \text{instanton}) \tag{5.8}$$

where J(x) is a source of gauge-invariant composite operators

$$q^\dagger \{\text{flavour and } \gamma\text{-matrices}\} q$$

The eigenvalues $\varepsilon[J]$ are given by

$$\left(\not{D} + J(x)\right) v_\varepsilon(x) = \varepsilon\, v_\varepsilon(x) \tag{5.9}$$

so the perturbation series in J(x) begins as follows

$$\varepsilon[J] = iE + \int d^4x\, u_E^\dagger(x)\, J(x)\, u_E(x) + O[J^2] \tag{5.10}$$

with

$$\begin{aligned} \not{D}\, u_E(x) &= iE u_E(x) \quad , \quad (E \text{ real}) \\ \int d^4x\, u_E^\dagger(x)\, u_{E'}(x) &= \delta_{EE'} \end{aligned} \tag{5.11}$$

If there are N quark flavours, the series in J(x) for the determinant begins with the $O[J^N]$ term[2]):

$$\det(\not{D} + J) = \left\{\prod_{E>0} E^2\right\} \det_{ij} \int d^4x\, u_{0i}^\dagger(x)\, J(x)\, u_{0j}(x) + O[J^{N+1}] \tag{5.12}$$

Thus in QCD with N massless quarks, the amplitude[2,3])

$$\left\langle \prod_{i=1}^{N} q_L^\dagger \Lambda_i q_R(x_i) \right\rangle_{\text{inst.}}$$
$$= \int_{\substack{\text{coll.} \\ \text{coords}}} (\text{constant}) \left\{g^{-C} \exp -8\pi^2/g^2\right\} \left(\prod_{i=1}^{N} \delta/\delta J_i(x_i)\right)_{J=0} \det\left[\not{D} + \sum_{j=1}^{N} J_j \Lambda_j (1 - \gamma_5{}^E)\right] \tag{5.13}$$

is non-zero if the flavour-space matrices Λ_i are suitably chosen (e.g. Λ_i = unity). Each operator $q_L^\dagger \Lambda_i q_R$ corresponds to a Minkowski-space operator $\bar{q}_L \Lambda_i q_R$ with flavour-independent chirality equal to +2 in right-handed units. The sum of these chiralities is 2N, so the amplitude (5.13) receives no contributions from perturbation theory.

In general, other non-perturbative Green's functions are generated by applying $(\delta/\delta J)^{N+1}$, $(\delta/\delta J)^{N+2}$, ..., to Eq. (5.8). A typical example is the amplitude

$$\left\langle q_R^\dagger q_L(x_{N+2}) \prod_{i=1}^{N+1} q_L^\dagger q_R(x_i) \right\rangle_{\text{inst.}} \neq 0 \tag{5.14}$$

However, in certain cases, simple exact expressions for det $(\not{D} + J)$ can be given because of the existence of a selection rule[2,4,5,12])

$$\sum(\text{R.H. chiralities}) = 2N\nu \tag{5.15}$$

whose derivation will be reviewed in Sections 7 and 8. To illustrate this rule, the example (5.8) will be modified in the following way:

i) allow any configuration A_μ mappable onto the O(5) hypersphere S_4;

ii) restrict the source to be purely right-handed

$$J(x) \longrightarrow J_R(x) = (1 - \gamma_5^E)\,\Lambda(x) \tag{5.16}$$

where $\Lambda(x)$ is a matrix in flavour space but contains no γ-matrices. Each variation $\delta/\delta J_R$ inserts an operator with (Minkowskian) chirality +2.

Instead of considering the expansion (5.10), we evaluate

$$\det(\not{D} + J_R) = \int [dq^\dagger dq]\, \exp - \int d^4x\, q^\dagger(x)\{\not{D}_x + J_R(x)\}\, q(x) \tag{5.17}$$

by introducing Grassmann variables ξ_E, η_E associated with the eigenfunctions $u_E(x)$ of $-i\not{D}$:

$$\det(\not{D} + J_R) = \left(\prod_E \int d\eta_E\, d\xi_E\right) \exp - S[J_R] \tag{5.18}$$

$$S[J_R] = i\sum_E E\,\eta_E \xi_E + \sum_{E,E'} \eta_E \xi_{E'} \int d^4x\, u_E^\dagger(x)\, J_R(x)\, u_{E'}(x) \tag{5.19}$$

The spectrum of eigenvalues E of $-i\not{D}$ depends on the choice of Yang-Mills fields A_μ.

The γ_5-invariance of the fermionic action is reflected in the chiral properties of the eigenfunctions u_E. The E = 0 eigenfunctions $u_{o\sigma}(x)$ are also eigenvectors of γ_5 (Euclidean):

$$\begin{aligned} \not{D}\, u_{o\sigma} &= 0 \\ \gamma_5\, u_{o\sigma} &= -u_{o\sigma}\ , \quad (\sigma = 1, \ldots, n_R) \\ \gamma_5\, u_{o\sigma} &= u_{o\sigma}\ , \quad (\sigma = n_R + 1, \ldots, n_R + n_L) \end{aligned} \tag{5.20}$$

The degeneracy of $\{u_{o\sigma}\}$ is constrained by the generalization of the index theorem (2.46) to N flavours:

$$n_R - n_L = N\nu \tag{5.21}$$

For each eigenvalue E > 0, there is an eigenvalue -E with eigenfunction

$$u_{-E} = -\gamma_5\, u_E \tag{5.22}$$

The expansion of exp - $S[J_R]$ becomes relatively simple if E = 0 and E $\neq$ 0 contributions to Eq. (5.19) are separated and Eq. (5.22) is applied:

$$
\begin{aligned}
S[J_R] &= i\sum_{E>0} E\left(\eta_E\xi_E - \eta_{-E}\xi_{-E}\right) + \sum_{\sigma,\tau}\eta_{0\sigma}\xi_{0\tau}\int d^4x\, u_{0\sigma}^\dagger J_R u_{0\tau} \\
&+ \sum_{\substack{\sigma\\E>0}}\left[\eta_{0\sigma}(\xi_E+\xi_{-E})\int d^4x\, u_{0\sigma}^\dagger J_R u_E + (\eta_E+\eta_{-E})\xi_{0\sigma}\int d^4x\, u_E^\dagger J_R u_{0\sigma}\right] \\
&+ \sum_{E,E'>0}(\eta_E+\eta_{-E})(\xi_{E'}+\xi_{-E'})\int d^4x\, u_E^\dagger J_R u_{E'}
\end{aligned}
\tag{5.23}
$$

The identities

$$
\begin{aligned}
(\xi_E+\xi_{-E})^2 &= 0 = (\eta_E+\eta_{-E})^2 \\
(\eta_E\xi_E - \eta_{-E}\xi_{-E})^2 &= 0 \\
(\eta_E+\eta_{-E})(\eta_E\xi_E - \eta_{-E}\xi_{-E})(\xi_E+\xi_{-E}) &= 0
\end{aligned}
\tag{5.24}
$$

imply that the J_R-dependent $E > 0$ terms in $S[J_R]$ cannot contribute terms of the form

$$\prod_{E>0} \eta_E\xi_E\eta_{-E}\xi_{-E}$$

to $\exp - S[J_R]$, and hence do not contribute to the integral (5.17):

$$
\begin{aligned}
&\det(\not{D} + J_R) \\
&= \left(\prod_{E'}\int d\eta_{E'}\,d\xi_{E'}\right)\exp\left[-i\sum_{E>0}E\left(\eta_E\xi_E-\eta_{-E}\xi_{-E}\right) - \sum_{\sigma,\tau}\eta_{0\sigma}\xi_{0\tau}\int d^4x\, u_{0\sigma}^\dagger J_R u_{0\tau}\right]
\end{aligned}
\tag{5.25}
$$

The integration over $E \neq 0$ modes yields the factor $\Pi_{E>0}\, E^2$, while the $E = 0$ modes produce a $(n_L + n_R) \times (n_L + n_R)$ determinant. This determinant vanishes if there is a left-handed eigenfunction $u_{0\sigma}$, because J_R is right-handed. Thus the answer is

$$
\begin{aligned}
\det(\not{D} + J_R) &= 0 \quad , \qquad (n_L \neq 0)\ ; \\
\det(\not{D} + J_R) &= \left(\prod_{E>0}E^2\right)\det_{\sigma\tau}\int d^4x\, u_{0\sigma}^\dagger(x)\,J_R(x)\,u_{0\tau}(x) \quad , \\
&\qquad (n_L = 0\ ,\ n_R = N\nu)
\end{aligned}
\tag{5.26}
$$

where the indices σ,τ run from 1 to $N\nu$. This is an exact result, not just the first term in a series expansion in powers of J_R. In other words, $\det(\not{D} + J_R)$ is homogeneous in J_R:

$$\det(\not{D} + \lambda J_R) = \lambda^{N\nu}\det(\not{D} + J_R) \quad , \quad (\nu \geqslant 0) \tag{5.27}$$

This example shows that Eq. (5.15) should be treated as a selection rule -- WKB amplitudes vanish if the sum of the chiralities is not equal to $2N\nu$. If the sum equals $2N\nu$, some amplitudes do not vanish, but not all. For example, suppose that Eq. (5.21) is satisfied with $\nu > 0$, $n_L \neq 0$. Then Eq. (5.15) holds, but the amplitude obtained from $\det(\not{D} + J_R)$ vanishes. Note that, within the WKB approximation, the rule depends on A_μ^{cl} being S_4 mappable.

The same rule (5.15) governs Green's functions of operators of arbitrary twist and dimension. These amplitudes can be obtained in short-distance expansions of fermionic Green's functions associated with the integral

$$I[j, j^\dagger] = \int [dq^\dagger dq] \exp - \int d^4x \, [q^\dagger \not{D} q + j^\dagger q + q^\dagger j] \tag{5.28}$$

where the Grassmann variables $j(x)$, $j^+(x)$ are sources of fermion fields q^+, q. Let

$$S'(x, y) = \sum_{E \neq 0} E^{-1} u_E(x) u_E^\dagger(y) \tag{5.29}$$

be the propagator of $-i\not{D}$ with zero modes removed[39]:

$$- i \not{D} S'(x, y) = \delta(x - y) - \sum_\sigma u_{0\sigma}(x) u_{0\sigma}^\dagger(y) \tag{5.30}$$

Then $I[j, j^+]$ is easily evaluated by shifting the integration variables

$$\begin{aligned} q(x) &\rightarrow q(x) - i\int d^4y \, S'(x, y) j(y) \quad , \\ q^\dagger(x) &\rightarrow q^\dagger(x) + i\int d^4y \, j^\dagger(y) S'(y, x) \end{aligned} \tag{5.31}$$

with the result

$$I[j, j^\dagger] = \left(\prod_{E>0} E^2\right) \left\{ \exp - i \iint d^4x \, d^4y \, j^\dagger(x) S'(x, y) j(y) \right\} \prod_\sigma j_\sigma^* j_\sigma \tag{5.32}$$

$$\begin{aligned} j_\sigma &= \int d^4x \, u_{0\sigma}^\dagger(x) j(x) \quad , \\ j_\sigma^* &= \int d^4x \, j^\dagger(x) u_{0\sigma}(x) \end{aligned} \tag{5.33}$$

The chirality properties of the sources determine whether a particular Green's function generated by (5.32) is zero or not. Let j_L, j_R be the left- and right-handed components of j:

$$\begin{aligned} j_L &= \tfrac{1}{2}(1 + \gamma_5^E) j \quad , \quad & (j^\dagger)_L &= \tfrac{1}{2} j^\dagger (1 - \gamma_5^E) \\ j &= j_L + j_R \quad , \quad & j^\dagger &= j^\dagger_L + j^\dagger_R \end{aligned} \tag{5.34}$$

Equation (5.22) implies that $S'(x,y)$ connects j_L with j_R^+, but not $j_L \leftrightarrow j_L^+$ or $j_R \leftrightarrow j_R^+$:

$$\iint d^4x \, d^4y \, j^\dagger(x) S'(x,y) j(y) = \iint d^4x \, d^4y [j^\dagger_L S' j_R + j^\dagger_R S' j_L] \tag{5.35}$$

Thus the exponential factor in Eq. (5.32) carries zero chirality. The zero-mode factor $\Pi_\sigma j_\sigma^* j_\sigma$ carries right-handed chirality $2N\nu$ because of Eqs. (5.20) and (5.21), so the rule (5.15) is satisfied. [Note that (j_σ, j_σ^*) are independent Grassmann variables, so $\Pi_\sigma j_\sigma^* j_\sigma$ is actually a determinant. It corresponds to the zero-mode determinants in Eqs. (5.12) and (5.26).]

The presence of the exponential term in (5.32) suggests the existence of analogues of the perturbative Feynman rules; the latter are generated by Eqs. (4.13) and (4.15) with G(x,y) equal to the free propagator. However, non-trivial contributions from zero modes tend to upset the analogy. In particular, the fact that the fermionic zero-mode term $\Pi_\sigma\, j^*_\sigma j_\sigma$ is not exponentiated makes the formulation of systematic rules cumbersome. Zero modes also complicate the rules for pure Yang-Mills theories in higher orders of the expansion (4.34)[63]. The safest procedure is to work directly with expressions derived from the relevant functional integral.

6. θ-VACUA

The WKB calculations reviewed above contain a phase ambiguity

$$\int [dA] \;=\; \sum_\nu e^{i(\mathrm{Phase})_\nu} \int [dA]_\nu \tag{6.1}$$

where $[dA]_\nu$ denotes Gaussian integration about classical fields with topological charge ν. There is a Minkowski-space argument[4-6] which results in all of these phases being described by a single angle*) θ. The original argument was specifically designed to apply only to the WKB approximation, and involved the choice of gauge

$$A_0 \;=\; 0 \tag{6.2}$$

The choice (6.2) is convenient but not essential[50,55,64], i.e. the results are not gauge artifacts. The generalization of θ-vacua to non-WKB situations will be indicated at the end of the section.

Consider the Minkowski-space version of the surface integral (2.16) for topological charge, where the surface is taken to be a large cylinder $\sigma(r, T)$ with radius r and length 2T:

$$\sigma(r,T) \;=\; \left\{ x_\mu \,;\; |\vec{x}| = r\,, \quad \text{or} \quad t = \pm T \right\} \tag{6.3}$$

$$\nu \;=\; \lim_{r,T\to\infty} \oint_\sigma d\sigma^\mu\, K_\mu(t,\vec{x}) \tag{6.4}$$

The asymptotic $x_\mu \to \infty$ behaviour of the Minkowskian field configurations dominating the WKB integral is assumed to be a pure gauge as in the Euclidean case [Eq. (2.25)]. For the gauge (6.2), this condition is

*) This result was suggested by the presence of a similar θ-angle in two-dimensional QED[65].

$$-igA_i^a(t,\vec{x})\tau^a \longrightarrow \begin{cases} G_m(\vec{x})^{-1}\nabla_i G_m(\vec{x}) \quad , \quad (t \to -\infty) & (6.5a) \\ G_n(\vec{x})^{-1}\nabla_i G_n(\vec{x}) \quad , \quad (t \to \infty) & (6.5b) \\ g(\Omega)^{-1}\nabla_i g(\Omega) \quad , \quad (|\vec{x}| \to \infty) & (6.5c) \end{cases}$$

where G_m and G_n are time-independent and tend to the same function $g(\Omega)$ as $|\vec{x}| \to \infty$. The index i denotes spatial components 1, 2, 3, and Ω is three-dimensional solid angle. The subscripts m,n refer to the value of the integral

$$K[G] \;=\; -(24\pi^2)^{-1}\int d^3x \; \mathrm{Tr}\,\varepsilon_{ijk}\, G^{-1}\nabla_i G\, G^{-1}\nabla_j G\, G^{-1}\nabla_k G \tag{6.6}$$

$$n \;=\; K[G_n] \tag{6.7}$$

Because of Eq. (6.5c), the sides of the cylinder do not contribute to the surface integral (6.4):

$$\nu \;=\; \int d^3x \left\{ K_0(\infty, \vec{x}) \;-\; K_0(-\infty, \vec{x}) \right\} \tag{6.8}$$

The values of $\int d^3x\, K_0$ at $t = \pm\infty$ are given by Eq. (6.7):

$$\nu \;=\; n - m \tag{6.9}$$

It follows that n and m must differ by an integer. If so desired, a change of gauge can be performed:

$$\begin{aligned} A.\tau \;\longrightarrow\; A'.\tau \;&=\; g(\Omega)\,A.\tau\, g(\Omega)^{-1} \;+\; ig^{-1}g(\Omega)\nabla g(\Omega)^{-1} \\ G_m \;\longrightarrow\; G_{m'} \;&=\; G_m\, g(\Omega)^{-1} \end{aligned} \tag{6.10}$$

The new gauge elements $G_{m'}$, $G_{n'}$ tend to 1 at $|\vec{x}| \to \infty$, so each three-dimensional flat space $t = \pm\infty$ can be compactified to the hypersphere S_3 and the analysis (2.28)-(2.38) repeated with the result

$$\begin{aligned} m', n' \;&=\; \text{winding numbers for } S_3 \to S_3 \\ &=\; \text{integers} \end{aligned} \tag{6.11}$$

Of course, the value of ν is not changed by the gauge transformation (6.10), so it equals $(n' - m')$.

The various $t = \pm\infty$ configurations (6.5) and linear combinations thereof represent part of a large vector space of states, V. This space involves non-trivial operators for gauge transformations, so it is larger than the physical Hilbert space H, i.e. H has to be projected out of V. A state in V is represented by a gauge-equivalent class of configurations. The size of these classes, and hence the size of V, depend on the restrictions placed on the gauge transformations relating members of a given class, e.g. restrictive gauge conditions produce small classes and hence a large V-space. For the present discussion, it is sufficient to label equivalence classes by the value of the integral (6.6) for a particular boundary condition (6.5c). Thus the in state $|m\rangle_-$ is represented by the set of all configurations (6.5a) generated by gauge elements G_m with the same value of m and with $G_m(\vec{x}) \to g(\Omega)$ for $|\vec{x}| \to \infty$. Similarly, the out state $|n\rangle_+$ is given by the class (6.5b). In general, a given initial state $|m\rangle_-$ can evolve via the time evolution operator and external operator insertions into any out state $|m+\nu\rangle_+$. The gauge-invariance of ν implies that effects due to this multiplicity of states cannot be gauge-transformed away. Perturbation theory contributes to the transition

$$|m\rangle_- \longrightarrow |m\rangle_+ \tag{6.12}$$

while the instanton is one of the dominant configurations for

$$|m\rangle_- \longrightarrow |m+1\rangle_+ \tag{6.13}$$

Let us consider a special set of time-independent gauge transformations $\tilde{G}(\vec{x})$ with the properties

$$K[\tilde{G}] = 1 \tag{6.14}$$

$$\tilde{G}(\vec{x}) \longrightarrow 1 \quad , \quad (|\vec{x}| \to \infty) \tag{6.15}$$

For example, the gauge element (2.37) on S_3 can be mapped onto flat three-dimensional space; a stereographic mapping $S_3 \leftrightarrow R_3$ [similar to Eq. (2.43) for $S_4 \leftrightarrow R_4$] yields the result

$$\tilde{G}(\vec{x}) = (i\vec{\sigma}.\vec{x} - a)(i\vec{\sigma}.\vec{x} + a)^{-1} \tag{6.16}$$

where a is a real positive constant. The integral (6.6) obeys the relation*)

$$K[GH] = K[G] + K[H] + (8\pi^2)^{-1}\int d^3x\, \nabla_i \operatorname{Tr} \varepsilon_{ijk}[G^{-1}\nabla_j G \nabla_k H H^{-1}] \tag{6.17}$$

for arbitrary smooth gauge-group elements G, H, so if G and H are chosen to be $\tilde{G}$ and G_m, the condition (6.15) ensures that the last term in Eq. (6.17) (written as a surface integral) does not contribute:

$$\tilde{G}\, G_m = G_{m+1} \tag{6.18}$$

Therefore the operator U on V-space which induces gauge transformations $\tilde{G}$ acts as a raising operator:

---

*) This is a non-compact version of the well-known formula[55,66,67] $\nu[GH] = \nu[G] + \nu[H]$, where $\nu[G]$ is given by Eq. (2.27) and G, H are defined on S_3. The usual derivation relies on deformation invariance [Eq. (2.38)] with G and H deformed to non-overlapping regions on S_3. Alternatively, compactification of (6.17) is possible if the gauge-group elements tend to 1 as $\vec{x} \to \infty$ on R_3. Then the last term in (6.17) vanishes and $K[G]$ becomes $\nu[G]$.

$$U|m\rangle_\pm = |m+1\rangle_\pm \tag{6.19}$$

As a check, we note that Eq. (2.15) implies the formula

$$U^{-1}\int d^3x\, K_0(x)\, U = \int d^3x\, K_0(x) + 1 \tag{6.20}$$

Equation (6.20) is consistent with Eq. (6.19) and the eigenvalue equations

$$\int d^3x\, K_0(\pm\infty, \vec{x})\,|m\rangle_\pm = m\,|m\rangle_\pm \tag{6.21}$$

The sets $\{|m\rangle_\pm\}$ can be chosen to be separately orthonormal,

$$_-\langle m|n\rangle_- = \delta_{mn} = {}_+\langle m|n\rangle_+ \tag{6.22}$$

but in $\rightarrow$ out transitions need not be diagonal,

$$_+\langle m|n\rangle_- \neq \delta_{mn} \tag{6.23}$$

unless det ($\not{D}$ + M) vanishes for $\nu \neq 0$ due to one of the quarks being massless.

Physical states must correspond to eigenstates of gauge transformations in general and of U in particular, because the Hamiltonian is gauge-invariant. The operator U is unitary, so its eigenvalues are $e^{-i\theta}$, where θ is a parameter running from 0 to 2π:

$$U|\theta \text{ state}\rangle = e^{-i\theta}|\theta \text{ state}\rangle \tag{6.24}$$

Each value of θ labels a separate theory with a Hilbert space H_θ projected out of V. The pure-gauge states $|m\rangle$ span a subspace of V from which θ-vacua can be constructed using Eq. (6.19):

$$|\theta\rangle_\pm = \sum_m e^{im\theta}|m\rangle_\pm \tag{6.25}$$

The choice of normalization (6.22) for $|m\rangle$ states implies the following continuum normalization for θ-vacua:

$$\begin{aligned} _+\langle\theta'|\theta\rangle_+ = {}_-\langle\theta'|\theta\rangle_- &= \sum_m e^{im(\theta - \theta')} \\ &= 2\pi\,\delta(\theta' - \theta) \end{aligned} \tag{6.26}$$

The out/in notation $\pm$ has been retained because Eq. (6.23) implies that $|\theta\rangle_+$ is not necessarily equal to $|\theta\rangle_-$.

The phases in Eq. (6.1) can be deduced by considering the construction of Green's functions of gauge-invariant operators O_k from θ-vacua. Equation (6.19) and the condition

$$U^{-1} O_k\, U = O_k \tag{6.27}$$

imply that T-products of O_k between $|m\rangle_-$ and $|n\rangle_+$ depend on the difference $\nu = n - m$ but not on $(m + n)$:

$$ {}_+\langle n | \prod_k O_k | m \rangle_- = T\langle \prod_k O_k \rangle_\nu \tag{6.28} $$

The result for the θ-vacuum expectation value is

$$ \begin{aligned} T_+\langle \theta' | \prod_k O_k | \theta \rangle_- &= \sum_{mn} e^{i(m\theta - n\theta')} T_+\langle n | \prod_k O_k | m \rangle_- \\ &= 2\pi\delta(\theta - \theta') \sum_\nu e^{-i\nu\theta} T\langle \prod_k O_k \rangle_\nu \end{aligned} \tag{6.29} $$

so θ is not changed by gauge-invariant perturbations. Now the real vacuum $|\mathrm{vac}\rangle$ in H_θ must be unique and have unit normalization. It is represented in V-space by the equivalence class of states through which $|\theta\rangle_-$ evolves to $|\theta\rangle_+$:

$$ T_+\langle \theta' | \prod_k O_k | \theta \rangle_- = {}_+\langle \theta' | \theta \rangle_- \; T\langle \mathrm{vac} | \prod_k O_k | \mathrm{vac} \rangle_\theta \tag{6.30} $$

Here O_k is an observable operator, i.e. it acts within H_θ. Combining Eqs. (6.29), (6.30), and (6.29) with $\Pi_k O_k$ replaced by the unit operator, we find:

$$ T\langle \mathrm{vac} | \prod_k O_k | \mathrm{vac} \rangle_\theta = \frac{\sum_\nu e^{-i\nu\theta} T\langle \prod_k O_k \rangle_\nu}{\sum_\nu e^{-i\nu\theta} \langle 1 \rangle_\nu} \tag{6.31} $$

The Euclidean version of (6.31) corresponds to the Green's function in Eq. (4.2), so the phases in Eq. (6.1) are given by

$$ e^{i(\mathrm{Phase})_\nu} = e^{-i\nu\theta} \tag{6.32} $$

The phase $e^{-i\nu\theta}$ can be absorbed into the Euclidean action in Eq. (4.2)[4,5],

$$ \sum_\nu e^{-i\nu\theta} \int [dA]_\nu \prod_k O_k \, e^{-(S + S_g)} = \sum_\nu \int [dA]_\nu \prod_k O_k \, e^{-(S(\theta) + S_g)} \tag{6.33} $$

with

$$ S(\theta) = S + i\theta\nu \tag{6.34} $$

This corresponds to a change

$$ \phi \longrightarrow \phi + \theta \tag{6.35} $$

in the Minkowskian action (2.8). In other words, the multiplicity of QCD theories generated by the existence of the θ-parameter corresponds to the freedom of choice of the coupling constant ϕ in Eq. (2.8). In particular, WKB amplitudes generated by the action (2.8) have periodicity 2π in ϕ. Note that the term $-\phi\nu$ in (2.8) can be written in Lagrangian form

$$S_{QCD}(\phi) = \int d^4x\, \mathcal{L}_{QCD}(x, \phi)$$
$$\mathcal{L}_{QCD}(x, \phi) = \mathcal{L}_{QCD}(x) - \phi(g^2/32\pi^2)\, F.F^* \tag{6.36}$$

provided that $\int d^4x\, F.F^*$ is interpreted as follows [for the gauge (6.2)]:

$$(g^2/32\pi^2)\int d^4x\, F.F^*(x) \equiv \int d^3x\, [K_0(\infty, \vec{x}) - K_0(-\infty, \vec{x})] \tag{6.37}$$

An important feature[5] of the analysis is that it is only in the θ basis that vacuum expectation values of observable operators are diagonal [i.e. proportional to $\delta(\theta' - \theta)$] and hence consistent with the cluster property (4.38). For θ-vacuum expectation values, the cluster property is

$$T_+\langle\theta'|\Big(\prod_k{}' O_k(x_k)\Big)\Big(\prod_\ell{}'' O_\ell(y_\ell)\Big)|\theta\rangle_-$$
$$\longrightarrow \int_0^{2\pi} d\theta''\, \big[2\pi\sum_\nu e^{-i\nu\theta''}\langle 1\rangle_\nu\big]^{-1}\, T_+\langle\theta'|\prod_k{}' O_k|\theta''\rangle_-\, T_+\langle\theta''|\prod_\ell{}'' O_\ell|\theta\rangle_- \quad ,$$
$$(x_k - y_\ell)^2 \longrightarrow -\infty \tag{6.38}$$

It is clear that the states $|m\rangle_\pm$ do not correspond to vacua: e.g., for a massless quark theory with only one flavour ($|m\rangle_+ = |m\rangle_-$), the cluster limit $x^2 \to -\infty$ for the transition

$$T\langle m|\,\bar{q}_R q_L(x)\, \bar{q}_L q_R(0)|m\rangle$$
$$\longrightarrow \langle m|\,\bar{q}_R q_L(x)|m+1\rangle\langle m+1|\,\bar{q}_L q_R(0)|m\rangle \tag{6.39}$$

is not diagonal in m. Also, the cluster property need not be valid for gauge-dependent operators. In particular, θ-matrix elements of $\int d^3x\, K_0$ involve derivatives of $\delta(\theta - \theta')$. Operators of this type act on the large space V but cannot be projected as operators on the physical space H_θ.

The discussion above is based on the identification of $|m\rangle_\pm$ as states functionally represented by the pure-gauge configurations (6.5). This is legitimate for Gaussian integrals about classical configurations mappable onto S_4, i.e. for the WKB approximation. However, vacuum structure in the real world of strong interactions is complicated by the spontaneous breaking of chiral SU(L) × SU(L) symmetry ($2 \le L \le N$). Candidates for vacuum states contain virtual quark-antiquark pairs, so the pure-gauge configurations (6.5) are irrelevant -- a linear combination of them will certainly not produce an acceptable vacuum state. Instead, generalized m-states have to be introduced as eigenstates of the operators

$$K_\pm = \int_\pm d\sigma^\mu\, K_\mu(x) \tag{6.40}$$

without reference to configurations in functional space*). Here the notation $\pm$ refers to positive and negative time surfaces which together form a large closed surface in Minkowski space (e.g. $t \gtrless 0$ $\frac{1}{2}$-hyperspheres of S_3):

$$\text{Topological Charge Operator} \quad = \quad K_+ - K_- \tag{6.41}$$

The operator (6.41) may take fractional as well as integer values[12] ν. The eigenstates of $K_\pm$ can also be eigenstates of operators which commute with $K_\pm$, or arbitrary linear combinations thereof. For example, arbitrary generalized m-states can have non-zero baryon number, non-trivial SU(3) properties, etc. In particular, axial SU(L) × SU(L) rotations relate $K_\pm$ eigenstates with the same eigenvalue m because these rotations commute with $K_\pm$ [at equal times, if SU(L) × SU(L) is not exactly conserved]. Obviously, not all of these m-states are needed for the construction of the correct vacua. The subspace in V from which vacua are constructed is specified by imposing the cluster property for observable operators, and a θ basis of the form (6.25)

$$|\theta\rangle_\pm \quad = \quad \sum_m e^{im\theta} |\text{gen. } m \text{ state}\rangle_\pm \tag{6.42}$$

$$K_\pm |\text{gen. } m \text{ state}\rangle_\pm \quad = \quad m |\text{gen. } m \text{ state}\rangle_\pm \tag{6.43}$$

is obtained from Eq. (6.35). The periodicity in θ is either a multiple of 2π or infinite. Generalized m-states need not be connected by gauge transformations, but Eq. (6.24) can be justified by using the generalization of Eq. (6.20) to $K_\pm$. Not all gauge-invariant operators are diagonal in the θ basis: some of them can induce transitions

$$\theta \quad \longrightarrow \quad \theta \ + \ 2\pi(\text{integer}) \tag{6.44}$$

A gauge-invariant operator which is also observable has to be diagonal in θ and satisfy (6.29) in order that the cluster property be preserved.

Generalized θ-vacua arise whenever the role of π, K, η, ... mesons as Nambu-Goldstone bosons becomes important [U(1) problem, axions]. A full treatment requires a separate review of the U(1) problem. However, the following sections make sense for the general case as long as it is remembered that the periodicity is not necessarily 2π.

---

*) This point should be understood when reading Ref. 12. There is *no* use of the WKB approximation, except for a digression [Eqs. (22)-(23)] designed to counter the widespread impression that an inspection of the chiral properties of WKB amplitudes suffices to solve the U(1) problem.

7. γ_5 ROTATIONS, ANOMALIES, SELECTION RULES

Let us ignore the quark mass parameters M_i in the QCD Lagrangian (2.10), and consider the group $U(1)_{ax}$ of global Abelian transformations

$$q_i \longrightarrow e^{i\gamma_5\alpha} q_i \quad , \quad (i = 1, \dots , N) \tag{7.1}$$

Because of the absence of mass parameters, this group is an invariance of the Lagrangian. However, the anomaly[68-72] associated with the triangle diagram in Fig. 5 complicates the formulation of this symmetry. In particular, the construction of the conserved Noether current $J^{sym}_{\mu 5}$ requires some care[19,69].

The anomaly is a phenomenon caused by *renormalization*. For renormalized amplitudes, one cannot assume canonical γ_5 symmetry at vertex A and maintain gauge invariance at vertices B and C at the same time. This means that renormalization cannot be ignored during the construction of $J^{sym}_{\mu 5}$ and that the result cannot be gauge-invariant:

$$\begin{aligned} J_{\mu 5}{}^{sym} &= \text{gauge-dependent} \quad , \\ \partial^\mu J_{\mu 5}{}^{sym} &= 0 \end{aligned} \tag{7.2}$$

To obtain a local description of $U(1)_{ax}$ transformations, it is necessary to pick a local gauge (such as $\partial . A = 0$ or $A_0 = 0$) in which fields commute at space-like separations. (The Coulomb gauge $\vec{\nabla}.\vec{A} = 0$ is not local because there are instantaneous Coulomb forces acting at a distance.) Then the conserved generator of $U(1)_{ax}$ transformations is

$$Q_5 = \int d^3x \; J_{05}(x)^{sym} \tag{7.3}$$

Even though Q_5 is invariant under infinitesimal gauge transformations[19,69], it is not generally gauge-invariant[4,5]. However, it is the commutators*)

$$[Q_5 , O_k]_- = - \chi_k O_k \tag{7.4}$$

which matter. The chirality χ_k of the operator O_k given by Eq. (7.4) appears in gauge-invariant Ward identities, so it is gauge-invariant, despite the gauge-dependence of Q_5.

Of course, we can always decide to renormalize the vertex $\gamma_\mu\gamma_5$ in a gauge-invariant way. The result is a different operator $J_{\mu 5}$ with the well-known formula[19,69]

$$\partial^\mu J_{\mu 5} = 2N\left(g^2/32\pi^2\right) F.F^* \tag{7.5}$$

for its divergence. Because of its gauge-invariance, $J_{\mu 5}$ has nothing to do with γ_5 transformations (apart from its coincidence with $J^{sym}_{\mu 5}$ in the limit of free

*) Equal-time, if the parameters M_i are not zero.

quark theory). It is not a current in the Noether sense. Rather, it is a very good example of a twist-two Wilson operator. Indeed, there is a non-trivial wave-function renormalization[69,71]

$$J_{\mu 5} = Z_5 (J_{\mu 5})_{bare} \tag{7.6}$$

where the leading $O(g^4)$ contribution to $(Z_5 - 1)$ is generated by the diagrams in Fig. 6:

$$Z_5 = 1 + (g^4 N/16\pi^4) \ln\Lambda + O(g^6) \quad , \quad (\Lambda = \text{cutoff mass}) \tag{7.7}$$

The relation between the two $\gamma_\mu\gamma_5$ operators is[19,69]

$$J_{\mu 5} = J_{\mu 5}^{sym} + 2N K_\mu \tag{7.8}$$

where K_μ is a renormalized version of the composite operator (2.14). In Lorentz-invariant gauges, $J^{sym}_{\mu 5}$ is not renormalized because of its conservation, so K_μ is renormalized as follows:

$$K_\mu = \left[K_\mu + (2N)^{-1}(Z_5 - 1) J_{\mu 5}\right]_{bare} \tag{7.9}$$

The $U(1)_{ax}$ Ward identity corresponding to Eqs. (7.2) and (7.4) is

$$\begin{aligned} &\int d^4x\, \partial_x^\mu T_+\langle\theta'|\, J_{\mu 5}(x)^{sym} \prod_k O_k(x_k)|\theta\rangle_- \\ &\quad = - \sum_\ell \chi_\ell\; T_+\langle\theta'|\prod_k O_k(x_k)|\theta\rangle_- \end{aligned} \tag{7.10}$$

In Lorentz-invariant gauges, an equivalent form of this equation is

$$\begin{aligned} &\int d^4x\, e^{iq\cdot x}\, T_+\langle\theta'|\, J_{\mu 5}(x)^{sym} \prod_k O_k(x_k)|\theta\rangle_- \\ &\sim - (iq_\mu/q^2) \sum_\ell \chi_\ell\; T_+\langle\theta'|\prod_k O_k(x_k)|\theta\rangle_- \quad , \quad q \to 0 \end{aligned} \tag{7.11}$$

For vacuum expectation values, Eq. (7.11) shows that there is a zero-mass pole coupled to $J^{sym}_{\mu 5}$ if there exists a non-zero Green's function $T\langle\Pi_k\, O_k\rangle$ with

$$\Sigma\,(\text{right-handed chiralities}) = \sum_\ell \chi_\ell \neq 0 \tag{7.12}$$

An example is the Green's function (5.13) computed by 't Hooft.

The singularity at $q^2 = 0$ in Eq. (7.11) is known as the Kogut-Susskind pole[73]. It is the Nambu-Goldstone boson caused by the *spontaneous* breaking of $U(1)_{ax}$ symmetry; i.e. the Hamiltonian is $U(1)_{ax}$ invariant, but a given θ-vacuum state is not. Indeed, Eqs. (6.20), (7.3), and (7.8) imply the result*)

$$U^{-1} Q_5 U = Q_5 - 2N \tag{7.13}$$

---

*) This formula was obtained in Ref. 5 (up to a sign to be discussed in Section 8) and in Ref. 4 (where the notation $\tilde{Q}_5$ is used).

and hence the identity

$$U\left\{\exp(-iQ_5\phi/2N)|\theta\rangle_\pm\right\} \tag{7.14}$$
$$= \quad e^{-i(\theta+\phi)}\left\{\exp(-iQ_5\phi/2N)|\theta\rangle_\pm\right\}$$

Comparison of Eqs. (6.24) and (7.14) shows that the $U(1)_{ax}$ rotation $\exp(-iQ_5\phi/2N)$ effects the transformation[4,5)]

$$|\theta\rangle_\pm \quad\longrightarrow\quad |\theta+\phi\rangle_\pm \tag{7.15}$$

apart from a phase factor or a unitary gauge-invariant operator:

$$\exp(-iQ_5\phi/2N)|\theta\rangle_\pm \quad = \quad \exp i\beta_\pm\phi \;\; |\theta+\phi\rangle_\pm \tag{7.16}$$

Obviously it is important to check whether the pole at $q^2 = 0$ in Eq. (7.11) survives the projection from the large vector space V to the physical Hilbert space H_θ or disappears completely. The pole is a gauge-dependent phenomenon because it couples to the gauge-dependent operator $J^{sym}_{\mu 5}$, but that is not a sufficient basis for concluding that it cannot contain physical as well as unphysical components. [As an analogy, consider the gauge-dependent tensor

$$T^{ab}_{\alpha\beta\gamma\delta} \quad = \quad F^{a}_{\alpha\beta}(x)\,F^{b}_{\gamma\delta}(x) \tag{7.17}$$

which can have a non-zero physical component given by the colour trace $T^{aa}_{\alpha\beta\gamma\delta}$.]

In order to examine the projection from V to H_θ, it is necessary to rewrite the $U(1)_{ax}$ Ward identity in a gauge-invariant form. We substitute Eq. (7.8) into (7.10) to obtain another identity, the anomalous Ward identity:

$$\int d^4x\;\partial^\mu_x\, T_+\langle\theta'|\,J_{\mu 5}(x)\,\prod_k O_k(x_k)|\theta\rangle_- \tag{7.18}$$
$$= \quad 2N\int d^4x\;\partial^\mu_x\, T_+\langle\theta'|\,K_\mu(x)\,\prod_k O_k(x_k)|\theta\rangle_-$$
$$- \quad \sum_\ell \chi_\ell\;\; T_+\langle\theta'|\prod_k O_k(x_k)|\theta\rangle_-$$

Note that the anomalous term involves a special prescription for the time-ordering of operator products which contain $F.F^*$:

$$T_+\langle\theta'|\,(g^2/32\pi^2)\,F.F^*(x)\,\prod_k O_k(x_k)|\theta\rangle_- \tag{7.19}$$
$$\equiv \quad \partial^\mu_x\; T_+\langle\theta'|\,K_\mu(x)\,\prod_k O_k(x_k)\,|\theta\rangle_-$$

If we had failed to recognize this point, the identity would have been meaningless: an arbitrary time-ordering of the hard (dimension 4) operator $F.F^*(x)$ would have induced ambiguities proportional to $\delta(x - x_k)$ which would compete with the equal-time commutator terms. There is no such ambiguity in (7.19). A T-ordering ambiguity in $T\langle K_\mu \Pi_k O_k\rangle$ proportional $\delta(x - x_k)$ or its derivatives becomes $\partial_x\delta(x - x_k)$, etc., in $T\langle F.F^* \Pi_k O_k\rangle$, and the latter ambiguity is annihilated by the volume integral $\int d^4x$.

The next step is to write the anomalous term as a surface integral and separate contributions from positive- and negative-time portions of the surface:

$$\begin{aligned}&\int d^4x\ \partial_x^\mu\ {}_+\langle\theta'|\,K_\mu(x)\prod_k O_k(x_k)|\theta\rangle_- \\ &= {}_+\langle\theta'|\left\{K_+ T\left[\prod_k O_k(x_k)\right] - T\left[\prod_k O_k(x_k)\right]K_-\right\}|\theta\rangle_-\end{aligned} \tag{7.20}$$

The operators $K_\pm$ are defined by Eq. (6.40); they reduce to $\int d^3x\ K_0(\pm\infty, \vec{x})$ for the special case of WKB calculations in the gauge $A_0 = 0$. Note the ordering of $K_\pm$ with respect to $O_k(x_k)$ in (7.20). This is required by the time-ordering of $K_\mu(x)$ with respect to $O_k(x_k)$ on the left-hand side. There may be points x_μ on the surface with

$$(x_k)_0 < x_0 < (x_\ell)_0 \tag{7.21}$$

but then $x - x_k$ is spacelike for all k (because the surface is sufficiently large) and in that case K_μ commutes with all O_k.

If the operators O_k are observable, Eqs. (6.41)-(6.43) [or (6.8), (6.21), and (6.25) for the WKB case with $A_0 = 0$] imply that the anomalous term is proportional to $\delta(\theta - \theta')$, so it can be written as a matrix element in H_θ space:

$$\begin{aligned}&\int d^4x\ \partial_x^\mu\ T\,{}_+\langle\theta'|\,K_\mu(x)\prod_k O_k(x_k)|\theta\rangle_- \\ &= {}_-\langle\theta'|\theta\rangle_-\ \sum_\nu e^{-i\nu\theta}\,\nu\ T\left\langle\prod_k O_k(x_k)\right\rangle_\nu \\ &= {}_+\langle\theta'|\theta\rangle_-\ T\left\langle vac\right|\left(g^2/32\pi^2\right)\int d^4x\ F.F^*(x)\prod_k O_k(x_k)\left|vac\right\rangle_\theta\end{aligned} \tag{7.22}$$

Thus the complete expression for the anomalous Ward identity in the physical space H_θ is

$$\begin{aligned}&\int d^4x\ \partial_x^\mu\ T\left\langle vac\right|J_{\mu5}(x)\prod_k O_k(x_k)\left|vac\right\rangle_\theta \\ &= 2N\ T\left\langle vac\right|\left(g^2/32\pi^2\right)\int d^4x\ F.F^*(x)\prod_k O_k(x_k)\left|vac\right\rangle_\theta \\ &\qquad - \sum_\ell \chi_\ell\ T\left\langle vac\right|\prod_k O_k(x_k)\left|vac\right\rangle_\theta\end{aligned} \tag{7.23}$$

where $|vac\rangle$ is normalized to unity.

The left-hand side of Eq. (7.23) is the amplitude for a zero-mass particle to couple to the *gauge-invariant* operator $J_{\mu 5}$. Clearly, a necessary condition for the absence of such a particle from the physical spectrum of the θ-world is that this amplitude should vanish, and this will happen *only* if the anomalous term cancels the equal-time commutators[12]. The best way to analyse this problem is to return to the identity (7.18) for θ-vacuum expectation values, substitute the second line of (7.22) and Fourier transform in θ. The result is a special case of identities obtained in Ref. 12:

$$\int d^4x \; \partial_x^\mu \, T\left\langle J_{\mu 5}(x) \prod_k O_k(x_k)\right\rangle_\nu = \left(2N\nu - \sum_\ell \chi_\ell\right) T\left\langle \prod_k O_k(x_k)\right\rangle_\nu \tag{7.24}$$

Thus Goldstone's theorem for conserved currents is replaced by Eq. (7.24) for the non-conserved current $J_{\mu 5}(x)$. The following situations should be distinguished:

i) In the conventional WKB approximation, the relevant functions (the classical solutions A^{cl} and vector and spinor eigenfunctions of the kinetic energy operators) are mappable onto S_4. Thus gauge-invariant quantities such as $J_{\mu 5}$ are completely smooth on S_4 -- no gauge-patching is needed. Since the surface area of S_4 is finite, infrared divergences associated with the $x_\mu \to \infty$ region of R_4 are actually not present, so the left-hand side of (7.24) vanishes:

$$\left(2N\nu - \sum_\ell \chi_\ell\right) T\left\langle \prod_k O_k\right\rangle_\nu^{WKB} = 0 \quad , \qquad (\nu = \text{integer}) \tag{7.25}$$

This is the selection rule mentioned previously [Eq. (5.15)]. Note that S_4 mappability is absolutely essential in this context*). If there should be configurations which cannot be compactified [as discussed in Section 2, Eqs. (2.39)-(2.41)], infrared singularities might cause the left-hand side of (7.24) to be non-zero. This corresponds to what is called a boundary correction in the mathematical literature[75]. Such terms are known to exist for gravitational metrics[15,25] and monopole solutions[26].

ii) In the real world of strong interactions, S_4 mappability is irrelevant, ν need not be an integer, and no viable analysis of functional integration is available. Thus there is no purely theoretical argument to indicate whether the left-hand side of (7.24) vanishes or not. Of course, one may have phenomenological reasons for wanting it to vanish -- that is what the U(1) problem is about[12].

---

*) I must emphasize that the assumption[73,74] that positive- and negative-metric components of the Kogut-Susskind pole cancel when coupled to gauge-invariant operators is *not* ensured by gauge invariance alone in four-dimensional theories.

Some additional remarks are in order:

a) When applying (7.24), it should be remembered that the amplitudes $\langle\Pi_k\ O_k\rangle_\nu$ are individually not physical. Physical amplitudes have to be reconstructed as follows:

$$\begin{aligned} \mathcal{A}_{\text{phys.}}(\theta) &= T\ \langle \text{vac}|\prod_k O_k|\text{vac}\rangle_\theta \quad , \\ {}_+\langle\theta'|\theta\rangle_-\ \mathcal{A}_{\text{phys.}}(\theta) &= {}_-\langle\theta'|\theta\rangle_-\ \sum_\nu e^{-i\nu\theta}\ T\ \langle\prod_k O_k\rangle_\nu \end{aligned} \tag{7.26}$$

b) Note that $g^2 F.F^*$ is *not* multiplicatively renormalized. The correct mode of renormalization can be determined from Eqs. (7.9) and (7.19), with the result[76]

$$\left(g^2/32\pi^2\right)F.F^* = \partial^\mu\left[K_\mu + (2N)^{-1}(Z_5-1)\ J_{\mu5}\right]_{\text{bare},T\text{-ord.}} \tag{7.27}$$

The derivative ∂^μ acts outside time-ordering operations. If there are no zero-mass particles coupled to $J_{\mu5}$, the term $\partial^\mu[J_{\mu5}]_{T\text{-ord}}$ does not contribute at zero momentum and topological charge ν is renormalization-group invariant.

8. IMPORTANT MINUS SIGN

The fact that the sum of the chiralities of operators in Green's functions is related to topological charge was first recognized by 't Hooft[2]. However, the results of various derivations[2,4,5,12] are not identical:

i) There is general agreement that the rule for WKB amplitudes is given by Eq. (5.15) or more explicitly, by Eq. (7.25). However, for γ-matrix conventions consistent with the anomalous divergence equation (7.5), there is a minus sign in the formula (7.4) for right-handed chiralities χ_k, so there has to be a minus sign in Eq. (1.1). As noted in the Introduction, some derivations involving ΔQ_5 have failed to produce this minus sign.

ii) In Ref. 12 and Section 7, the existence of a selection rule was connected with the absence of zero-mass particles coupled to the gauge-invariant operator $J_{\mu5}$; see Eq. (7.24).

In Ref. 5, the axial charge Q_5 for operator chiralities χ_k is correctly identified: it is gauge-dependent and given by Eq. (7.3). In their version of the selection rule, the minus sign is missing only because it was dropped in the arithmetic -- if the gauge transformation operator U takes $|n\rangle$ into $|n+1\rangle$, $\sum_n e^{in\theta}|n\rangle$ is an eigenstate of U with eigenvalue $e^{-i\theta}$ (not $e^{i\theta}$), and $U^{-1}Q_5U$ appears in Eq. (7.13) instead of UQ_5U^{-1}. The derivation in Ref. 5 is based on the formula

$$Q_5|n=0\rangle_\pm = 0 \tag{8.1}$$

Beyond a purely classical context, the conditions under which this equation is true are unclear to me, so I have always avoided its use. We shall see that the requirement that zero-mass particles should be absent includes Eq. (8.1) as a special case.

There is an entirely different derivation which has unfortunately become much more popular. It is based on the incorrect assumption that operator chirality χ_k has something to do with the gauge-invariant operator

$$X(t) \quad = \quad \int d^3x \; J_{05}(x) \tag{8.2}$$

even though anomalies are present. The idea is to apply $\int d^4x$ directly to the anomalous divergence equation (7.5) and use the definition (2.6) of topological charge, with the result

$$\Delta X \quad \stackrel{?}{=} \quad 2N\nu \tag{8.3}$$

In (8.3), ΔX is taken to be the out eigenvalue of $X(+\infty)$ minus the in eigenvalue of $X(-\infty)$, and this is assumed to be the same as the sum of the operator chiralities [i.e. $X(t)$ is confused with Q_5]. Here the problem is not simply a matter of arithmetic: the derivation is genuinely illegitimate.

Apart from troubles with sign, Eq. (8.3) is obviously inconsistent with renormalization-group invariance. The topological charge ν is invariant to changes in renormalization procedure, but $J_{\mu5}$ [see Eq. (7.6)] and hence X suffers a multiplicative renormalization

$$X \quad \longrightarrow \quad (\text{constant})\,X \tag{8.4}$$

This is a very good way of testing the validity of formulas in the literature.

What went wrong?

An obvious point is that the difference between out and in eigenvalues of an operator equals the sum of coefficients of equal-time commutators only if that operator is conserved. Hence the equation

$$\Delta Q_5 \quad = \quad - \sum_k \chi_k \qquad , \quad (\dot{Q}_5 = 0) \tag{8.5}$$

is certainly a consequence of Eq. (7.4), but there is no analogue of it for $X(t)$.

More importantly (as indicated above), it has to be understood that $X(t)$ has absolutely no connection whatever with operator chirality χ_k when anomalies are present. This point was recognized from the beginning by Adler[69,71] who observed that $X(t)$ does not commute with derivatives of the gluon field, whereas Q_5 does*). The reason is that the definition of $J_{\mu5}$ involves an implicit dependence

*) See Eqs. (214) and (234) of Adler's Brandeis lectures[71]; his notation for Q_5 is $\tilde{Q}_5$.

on the gluon field; e.g. it can be recovered from the short-distance limit of the gauge-invariant combination

$$B_{\mu 5}(x,y) \quad = \quad \bar{q}(x)\left\{\exp\, ig\int_y^x dz^\lambda A^a_\lambda(z)\tau^a\right\}_{\text{ord.}}\gamma_\mu\gamma_5\, q(y) \tag{8.6}$$

where the exponentiated line integral is ordered. In higher orders, the lack of renormalization-group invariance results in divergent equal-time limits

$$[X(t'), O(\vec{x},t)]_- \quad \sim \quad \ln^p(\mu|t-t'|)\{\text{operator}\} \quad , \quad (t'\sim t) \tag{8.7}$$

where μ denotes a renormalization subtraction point.

The origin of the confusion is a misunderstanding of the notation

$$\text{"current"} \quad = \quad \sum_{i=1}^{N} \bar{q}_i\gamma_\mu\gamma_5 q_i \tag{8.8}$$

This is not an ordinary product of quark fields in free-field theory -- it is usually understood to be a gauge-invariant renormalized normal product[77]. A more explicit notation for the $\gamma_\mu\gamma_5$ operators would be

$$\begin{aligned} J_{\mu 5} \quad &= \quad N_{gi}\left[\sum_i \bar{q}_i\gamma_\mu\gamma_5 q_i\right] \quad , \\ J_{\mu 5}^{\,sym} \quad &= \quad N_{can}\left[\sum_i \bar{q}_i\gamma_\mu\gamma_5 q_i\right] \end{aligned} \tag{8.9}$$

where N_{gi} and N_{can} denote gauge-invariant and canonically behaved normal products, respectively. It is *convention*, and nothing else, which decrees that the notation (8.8) should refer to N_{gi} rather than N_{can}. Most people follow this convention, which was introduced by Adler[69]. However, Kogut and Susskind[73] prefer to think of (8.8) as the gauge-dependent product $J^{sym}_{\mu 5}$. Their notation is as logical as the usual one.

It is obvious that the choice of current associated with chirality is not decided by somebody's convention. We want the chiralities and hence the relevant current to be renormalization-group invariant. In practice, this means that the current should be conserved or partially conserved, as in Eqs. (7.2)-(7.4). It is not correct to argue that the gauge-dependence of $J^{sym}_{\mu 5}$ means that χ_k is unphysical, because we have seen that χ_k appears in the gauge-invariant anomalous Ward identity (7.23). Since χ_k does not depend on the mass parameters M_i or μ, it must be independent of the coupling constant g. [The same argument implies the validity of Gell-Mann's commutation relations

$$\begin{aligned} [F^i(t), \mathcal{F}^j_\mu(t,\vec{x})]_- \quad &= \quad i f^{ijk}\mathcal{F}^k_\mu(t,\vec{x}) \quad , \\ F^i(t) \quad &= \quad \int d^3x\, \mathcal{F}^i_0(t,\vec{x}) \end{aligned} \tag{8.10}$$

in QCD. Because of the partial conservation of the SU(N) × SU(N) current $\mathcal{F}^i_\mu$, the constants f^{ijk} are renormalization-group invariant and do not depend on the quark mass parameters. Hence f^{ijk} are g-independent and equal to the structure constants of SU(N) × SU(N).]

Another error in the argument leading to Eq. (8.3) is the assumption that the application of $\int d^4x$ to an operator equation for a current divergence yields a result applicable to a T-product or a Euclidean Green's function. To illustrate the problem, let us consider an example where anomalies play no role. Let $\vec{\mathcal{F}}_{\mu 5}$ be the conserved axial-vector current in an SU(2) × SU(2) symmetric σ-model with fields $(\sigma, \vec{\pi})$ forming a $(\frac{1}{2}, \frac{1}{2})$ representation. If $\int d^4x$ is applied to the equation

$$\partial^\mu \mathcal{F}^3_{\mu 5} = 0 \tag{8.11}$$

we get the analogue of Eq. (8.3):

$$\Delta F_5^3 \overset{?}{=} 0 \tag{8.12}$$

Of course, this is nonsense because there are transitions in which F_5^3 chirality changes. The simplest case is the vacuum expectation value

$$\langle \sigma \rangle = \tfrac{1}{2}\langle \sigma - i\pi_3 \rangle + \tfrac{1}{2}\langle \sigma + i\pi_3 \rangle \neq 0 \tag{8.13}$$

where the F_5^3 chirality changes by ±1:

$$[\sigma \mp i\pi_3 , F_5^3]_- = \pm (\sigma \mp i\pi_3) \tag{8.14}$$

The correct procedure for dealing with this problem is well-known to current algebraicists. One needs Ward identities for T-products or Euclidean functions which involve contact terms caused by short-distance singularities such as x_α / x^4:

$$\partial_\alpha [x_\alpha / x^4] = 2\pi^2 \delta^4(x) \quad , \quad (\text{Euclidean } x) \tag{8.15}$$

The procedure leading to (8.12) fails because it ignores contact terms. Instead of (8.12), we consider

$$\int d^4x \; \partial_x^\mu \, T \langle \mathcal{F}^3_{\mu 5}(x) (\sigma + i\pi_3)(0) \rangle = \langle \sigma + i\pi_3 \rangle \tag{8.16}$$

and conclude that $\langle \sigma + i\pi_3 \rangle$ is given by the non-zero amplitude for a pion to couple to $\mathcal{F}^3_{\mu 5}$. The existence of transitions with

$$\begin{aligned} \dot{F}_5^3 &= 0 \\ \Delta F_5^3 &\neq 0 \end{aligned} \tag{8.17}$$

is a signal for *spontaneous* breaking of SU(2) × SU(2).

What is the correct equation which replaces Eq. (8.3)? Consider an amplitude

$$\mathcal{A} = T \langle \text{out} | \prod_k O_k(x_k) | \text{in} \rangle \tag{8.18}$$

where $|\text{out}\rangle$ and $|\text{in}\rangle$ are eigenstates of $X(\infty)$ and $X(-\infty)$, respectively. The identity

$$\mathcal{A}\Delta X = \langle \text{out} | \{ X(\infty) T[\prod_k O_k] - T[\prod_k O_k] X(-\infty) \} | \text{in} \rangle \tag{8.19}$$

can be converted into a volume integral via Gauss's Theorem:

$$\mathcal{A}\Delta X = \int d^4x \, \partial_x^\mu T \langle \text{out} | J_{\mu 5}(x) \prod_k O_k(x_k) | \text{in} \rangle \tag{8.20}$$

The latter amplitude measures the coupling of a zero-mass particle to the gauge-invariant operator $J_{\mu 5}$. Hence the condition for this particle to be absent is simply

$$\Delta X = 0 \tag{8.21}$$

Note that the fact that $J_{\mu 5}$ is not conserved plays no role in the derivation of (8.21) -- the value of ΔX is controlled by the coupling of zero-mass particles to $J_{\mu 5}$. Equation (8.21) replaces the incorrect Eq. (8.3).

It is not difficult to obtain the selection rule (1.1) directly from Eq. (8.21). Equation (7.8) implies the formula

$$X(\pm\infty) = Q_5 + 2NK_\pm \tag{8.22}$$

Since Q_5 and $K_\pm$ commute, it is possible to construct simultaneous eigenstates $|q, n\rangle_\pm$, where q is the eigenvalue of Q_5. These states are also eigenstates of $X_\pm$:

$$X(\pm\infty)|q, n\rangle_\pm = (q + 2Nn)|q, n\rangle_\pm \tag{8.23}$$

The selection rule is obtained by forming the difference between the out and in eigenvalues of Eq. (8.22). Remembering that ΔK is the topological charge ν, we find that Eq. (8.21) becomes

$$0 = \Delta Q_5 + 2N\Delta K = \Delta Q_5 + 2N\nu \tag{8.24}$$

in agreement with Eq. (1.1). This derivation makes the origin of the minus sign obvious.

Note that Eq. (8.1) corresponds to the assumption that $X_\pm$ annihilate θ vacua. Hence (8.1) is a special case of Eq. (8.21).

These remarks are designed to counter the widespread misconception [based on (8.3)] that instantons somehow cause $U(1)_{ax}$ invariance to be explicitly broken by the Lagrangian. It is very important to recognize that gauge-invariant chiralities are derived from the gauge-dependent charge Q_5. The Lagrangian is *always* γ_5 invariant in the absence of mass parameters M_i. Instantons induce a *spontaneous* breaking of $U(1)_{ax}$ by generating a continuum of θ vacua. That is the origin of the non-zero changes of chirality observed in WKB calculations. The analogue of Eq. (8.17)

$$\begin{aligned} \dot{Q}_5 &= 0 \quad , \\ \Delta Q_5 &\neq 0 \end{aligned} \tag{8.25}$$

shows that we have a classic case of spontaneous symmetry breaking.

9. P AND T CONSERVATION IN STRONG INTERACTIONS

There is an old argument[78] [reviewed by Fritzsch[79] at this School] that a renormalizable theory of strong interactions with local non-chiral colour symmetry and no spin-0 fields necessarily conserves P, C, and T separately. The main step is to reduce an arbitrary non-Hermitean, γ_5-dependent quark mass matrix to a real, diagonal, γ_5-independent matrix

$$M = \begin{pmatrix} M_1 & & & \\ & M_2 & & \\ & & \ddots & \\ & & & M_N \end{pmatrix} \tag{9.1}$$

by a γ_5-dependent unitary transformation on the quark fields.

However, the effects of topological charge upset this argument. Even in the WKB approximation, one can easily see that in general, $F.F^*$ has a non-zero vacuum expectation value for non-zero mass parameters M_i. The only exceptional case is that in which the value of θ associated with states $|\theta\rangle$ accidentally cancels the coupling constant ϕ in Eq. (2.8) via Eq. (6.35) (modulo the periodicity). If we choose the convention $\phi = 0$, with P and T violation contained in the states,

$$P|\theta\rangle_\pm = |-\theta\rangle_\pm \quad , \qquad T|\theta\rangle_\pm = |-\theta\rangle_\mp \tag{9.2}$$

the problem is to explain why θ takes one of the values

$$\theta \simeq 0 \ , \ \text{or} \ \ \tfrac{1}{2}[\text{Periodicity}] \tag{9.3}$$

in the real world.

The natural explanation is that the theory actually possesses some sort of chiral U(1) invariance. There are two alternatives:

i) Peccei and Quinn[9] have proposed that the Lagrangian of the unified theory of strong and weak interactions should possess a global chiral symmetry $U(1)_{PQ}$. The theory contains a set of Higgs fields $\{\phi_h\}$ whose vacuum expectation values generate the quark mass matrix with all $M_i \neq 0$:

$$M = M\{\langle\phi_h\rangle\} \tag{9.4}$$

The symmetry $U(1)_{PQ}$ induces a phase rotation

$$\phi_h \longrightarrow \exp(ie_h\alpha)\ \phi_h \tag{9.5}$$

as well as transforming some or all of the quarks according to Eq. (7.1). This corresponds to a conserved current

$$\left(J_{\mu 5}{}^{sym}\right)_{PQ} = \left(J_{\mu 5}{}^{sym}\right)_{QCD} + \sum_h e_h \phi_h^* i\overleftrightarrow{\partial}_\mu \phi_h \tag{9.6}$$

In general, other weak fields (leptons, weak gauge fields, etc.) may also be transformed by $U(1)_{PQ}$.

ii) Otherwise, the up quark q_1 can be given zero mass[10,11]. This means that the unified Lagrangian is supposed to be invariant under transformations which rotate the up quark,

$$q_1 \longrightarrow \exp(i\gamma_5\alpha)\ q_1 \tag{9.7}$$

but leave the other fields (q_2, ..., q_N, ϕ_h, etc.) unchanged.

The observed CP violation is included as a purely weak effect. One possibility is the Kobayashi-Maskawa model[80].

The connection between exact chiral U(1) symmetry and P and T conservation was noticed in the original analyses[4,5] of θ-vacua. Roughly speaking, the $U(1)_{ax}$ transformation changes $|\theta\rangle$ into $|\theta + \text{const.}\rangle$ but leaves the Hamiltonian (and hence the time-evolution operator) invariant, so S-matrix amplitudes cannot depend on θ. If a term $-\phi\nu$ appears in the action as in Eq. (2.8), one first absorbs it into the definition of θ-states via Eq. (6.35) and then carries through the argument. [It is *not* correct to say that γ_5 transformations change the action operator by adding a term proportional to ν. The action is *really* $U(1)_{ax}$ invariant. A discussion of the effect of the chiral charge on *states* is essential if one is to derive conclusions about S-matrix amplitudes. This issue is related to the analysis of selection rules in Section 8.]

In fact, this analysis contains an implicit assumption about the symbols $\beta_\pm$ in Eq. (7.16). If θ-states are constructed from eigenstates of $K_\pm$ with phases as in Eq. (6.42), Eq. (8.22) implies that $\beta_\pm$ is proportional to $X_\pm$. In order to obtain the desired result, it is necessary to use the condition (8.21) that zero-mass particles are absent.

I prefer the following version of the analysis. Consider θ-states $|\theta\ \text{out/in}\rangle$, which are now not just generalized θ-vacua $|\theta\rangle_\pm$ picked out by the cluster property, but can be particle states such as $|\pi p\rangle$, $|p\bar{p}\rangle$, etc.:

$$K_-|\theta\ \text{in}\rangle = -i\,\partial/\partial\theta\,|\theta\ \text{in}\rangle \tag{9.8}$$

The S-matrix $\Sigma(\theta)$ is given by

$$\langle\theta'\text{out}|\theta\ \text{in}\rangle = {}_+\langle\theta'|\theta\rangle_-\ \Sigma(\theta) \tag{9.9}$$

The anomalous divergence equation implied by the $U(1)_{ax}$ symmetry of the Lagrangian,

$$\int d^4x\ \partial_x^\mu \langle\theta'\ \text{out}|J_{\mu5}(x)|\theta\ \text{in}\rangle = 2N\langle\theta'\text{out}|(K_+ - K_-)|\theta\ \text{in}\rangle \tag{9.10}$$

and the absence of zero-mass particles coupled to $J_{\mu5}$ imply the constraint

$$\left(\frac{\partial}{\partial\theta'} - \frac{\partial}{\partial\theta}\right)\langle\theta'\ \text{out}|\theta\ \text{in}\rangle = 0 \tag{9.11}$$

Equation (9.11) is valid for the special case ${}_+\langle\theta'|\theta\rangle_-$, so from Eq. (9.9) we obtain the result

$$\partial/\partial\theta\ \Sigma(\theta) = 0 \tag{9.12}$$

In other words, the S-matrix does not depend on θ. If ϕ is not zero in Eq. (2.8), it can be absorbed into the definition of θ which can then be put equal to zero without changing the S-matrix. Then the S-matrix can be generated by a theory in which both the ground state and the action respect P and T invariance. Consequently, S-matrix amplitudes are P,T symmetric.

Weinberg[10] and Wilczek[11] have observed that the Peccei-Quinn alternative (i) implies the existence of a light pseudoscalar boson, called the axion. The Higgs sector couples to strong interactions via $G_F^{1/2}$, where G_F is the Fermi weak coupling constant, so $v \neq 0$ effects do not contribute to the lowest order in $G_F^{1/2}$. However, $U(1)_{PQ}$ symmetry is valid order-by-order in $G_F^{1/2}$, so in zeroth order there has to be a massless Goldstone boson which causes $U(1)_{PQ}$ symmetry to be spontaneously broken (thus avoiding parity doubling in the weak spectrum). In higher orders in $G_F^{1/2}$, $v \neq 0$ effects are important. They provide an additional spontaneous breaking of $U(1)_{PQ}$ due to the extra $|\theta_{PQ}\rangle$ degeneracy. If the strong sector produces a Kogut-Susskind pole for $(J_{\mu5}^{sym})_{PQ}$ which decouples from the real world, the axion can acquire a small mass:

$$m_a = O[G_F^{1/2}] \tag{9.13}$$

The phenomenology of the axion[10,11,81-93] is somewhat model-dependent. The most general predictions concern the mass m_a and amplitudes for the transitions $a \leftrightarrow \pi^0$, $a \leftrightarrow \eta$. The derivation depends on the observation[10,81] that the limit $M_{1,2} \to 0$ produces a Lagrangian with two $U(1)_{ax}$ symmetries: $U(1)_{PQ}$, and the Abelian symmetry $U_2(1)$ (in the notation of Ref. 12) associated with γ_5 transformations on the up and down quarks alone. To each $U(1)_{ax}$ symmetry, there corresponds a gauge-invariant operator $J_{\mu 5}$ with an anomalous divergence. However, these anomalies can be cancelled by taking the appropriate linear combination of $(J_{\mu 5})_{PQ}$ and $(J_{\mu 5})_2$. The result is a gauge-invariant U(1) current with no anomaly; it becomes exactly conserved in the limit $M_{1,2} \to 0$, so the axion has to be its Goldstone boson, with[10,81]

$$m_a = O\left[G_F^{1/2} m_\pi\right] \tag{9.14}$$

for M_1, $M_2 \neq 0$. This trick enables one to analyse the axion without having to worry about a possible $U(1)_{PQ}$ problem -- the $U(1)_{PQ}$ problem is "cancelled off" against the $U_2(1)$ problem.

The results are complicated by the appearance of an angle α associated with neutral components of complex Higgs multiplets ϕ_1, ϕ_2 coupled to up and down quarks:

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= -\tfrac{1}{2}\Big[M_1 \bar{q}_1(1+\gamma_5)q_1 \phi_1^0/\langle\phi_1^0\rangle + M_2 \bar{q}_2(1+\gamma_5)q_2 \phi_2^{0*}/\langle\phi_2^0\rangle^* \\ &\qquad + \text{other quarks} + \text{Hermitean conjugate}\Big], \\ \tan\alpha &= \left|\langle\phi_2^0\rangle/\langle\phi_1^0\rangle\right| \end{aligned} \tag{9.15}$$

This angle determines the linear combination of neutral Higgs fields which represents the axion, but its value is not fixed by $U(1)_{PQ}$, so the predictions are α-dependent. A typical $\left[SU(2) \times U(1)\right]_{weak}$ model yields the result[10,81]

$$m_a \simeq \left(N/\sin 2\alpha\right)\left\{25 \text{ KeV}\right\} \tag{9.16}$$

where numerical values have been substituted for M_i, G_F, and the pion-decay constant $F_\pi \simeq 94$ MeV, and N is the number of quark flavours transformed by $U(1)_{PQ}$. In practice, N is the same as the total number of quark flavours, so Eq. (9.16) provides a lower bound of roughly 100 keV for the mass. Axions couple to strongly interacting matter (not involving new particles) with strength of order $G_F^{1/2}F_\pi \simeq 3 \times 10^{-4}$ by mixing with π^0 and η. Again there is α-dependence: the couplings for $a \leftrightarrow \pi^0, \eta$ are linear combinations of cot α and tan α. Finally, the coupling to leptons seems to be completely model-dependent -- a prediction for $a \to \ell^+\ell^-$ (α-dependent) is obtained only if the behaviour of the lepton fields under $U(1)_{PQ}$ is specified as an initial condition. An indirect way of doing this is to allow only two Higgs doublets in SU(2) × U(1) models[10,11]. If this condition is relaxed, it is possible to construct models with no direct ae^+e^- vertex[90].

Despite this model dependence, there is general agreement that the experimental outlook for axions is unpromising. The original suggestion

$$m_a < 2m_e \tag{9.17}$$

is excluded by a nuclear reactor experiment[94] sensitive to

$$a + d \rightarrow n + p$$

and by various beam-dump experiments[95]. These experiments establish bounds which are two to four orders of magnitude below theoretical expectations[18,85-89]. For larger masses

$$m_a = O[\text{few MeV}] \tag{9.18}$$

the absence of axions in the above experiments can be explained if a couples to e^+e^- directly, because then the axion decays quickly and does not reach the detector. However, this conflicts with another reactor experiment[96] sensitive to

$$a + e \rightarrow \gamma + e$$

and with a recent analysis[97] of old Gargamelle data. If couplings of the axion to leptons are excluded, m_a would have to be even larger to energetically forbid axion production in reactors and to permit it to decay sufficiently rapidly by other modes in beam-dump experiments. For very large m_a, the coupling $a \leftrightarrow \pi^0$, η is enhanced by a factor

$$(\sin 2\alpha)^{-1} = O[m_a/(100\ \text{KeV})] \quad , \quad (\alpha \simeq 0, \pi/2) \tag{9.19}$$

so a would be seen in ordinary strong-interaction experiments. This phenomenology is not completely exhaustive yet, but it does indicate strongly that alternative (i) is not realized.

Alternative (ii) avoids axions but is difficult[10,11] to reconcile with the conventional picture[79,98] of $K^+ - K^0$, $n - p$, and $D^+ - D^0$ mass differences and $\eta \rightarrow 3\pi$ decay. In this scheme, isospin-violating amplitudes are assumed to receive contributions from two independent sources:

a) a finite photon loop to which the usual soft-pion or dispersive methods can be applied;

b) a finite isospin-breaking mass term in an effective quark Lagrangian, with

$$(M_2/M_1)_{\text{eff.}} \simeq 1.8 \tag{9.20}$$

The consistency of dropping Eq. (9.20) in favour of zero up-quark mass has been the subject of some debate[99].

It is not obvious that the mass $(M_1)_{eff}$ used in the conventional picture is the same as the up-quark mass obtained from the unified Lagrangian. A related problem arose some time ago in a short-distance analysis[100] of the current algebra of electromagnetic corrections. The conclusion was that it is not always possible to commute soft-pion limits with the photon loop integral. [This is easily shown if the photon loop requires an infinite counterterm proportional to $(\bar{q}_1 q_1 - \bar{q}_2 q_2)$, as in asymptotically free theories. The axial charge for the soft π^0 limit commutes with the electromagnetic current but not with $(\bar{q}_1 q_1 - \bar{q}_2 q_2)$, so the naive procedure yields an infinite answer for a finite amplitude.] If true, this implies that the relevant ratio $(M_2/M_1)_{eff}$ is different for different processes. It was suggested that this might explain why the conventional picture does not explain the rate for $\eta \to 3\pi$ decay very well. Perhaps these theoretical uncertainties permit the choice $M_1 = 0$ and P and T conservation in strong interactions need not be regarded as accidental.

Alternatively[11,101,102], one may adopt less natural criteria which do not involve a U(1) invariance but are sufficiently restrictive to ensure P and T invariance.

10. OTHER EFFECTS

Topological charge is claimed to play an important role in a number of applications. We conclude with a brief summary of them.

The effects of instantons for the weak gauge group were discussed by 't Hooft[2]. In this case, the vertices B and C in Fig. 5 involve both γ_α and $\gamma_\alpha\gamma_5$, so it is possible for gauge invariance at B and C to clash with the conservation of a current at A constructed from γ_μ. This means that baryon and lepton currents introduced externally and not gauged can have anomalies which, together with the weak instanton, result in a violation of baryon and lepton number conservation. In SU(2) × U(1) models, amplitudes for this are proportional to[2]

$$\exp - 8\pi^2/g^2_{weak} \quad = \quad \exp - 2\pi\alpha^{-1}\sin^2\theta_W \tag{10.1}$$

where $\alpha \simeq 1/137$ is the fine-structure constant and θ_W is the Weinberg angle.

In strong interactions, it has been shown by current algebraic methods[12] that the existence of topological charge is essential for the generation of a non-zero pion-decay constant F_π in the SU(2) × SU(2) limit. Ward identities originally considered by Glashow[103] acquire an extra anomalous term which provides the leading PCAC contribution

$$\langle\langle \nu^2 \rangle\rangle \quad = \quad - \; i\left(g^2/32\pi^2\right)^2 \; T\langle vac|\int d^4x \, F.\overset{*}{F}(x) \, F.\overset{*}{F}(0)|vac\rangle \qquad (10.2)$$

to the result[12])

$$m_\pi^{\,2} F_\pi^{\,2} \;=\; 4 \; \langle\langle \nu^2 \rangle\rangle \;+\; O\left(m_\pi^4\right) \qquad (10.3)$$

Specific calculations with instantons in strong interactions depend on small instanton sizes being dominant. The simplest process emphasizing short distances is the total cross-section for e^+e^- annihilation at large energies. This has been considered by several authors[104-107]). Although the size-integration problem arises for the full T-product

$$T\,\langle vac|\,J_\mu(x)\,J_\nu(0)\,|vac\rangle$$

it was noticed[106,107]) that the one-instanton contribution to the absorptive part is dominated by small-size instantons. This permits an estimate of the asymptotic behaviour of the one-instanton correction to the cross-section relative to the leading scale-invariant term. The relative correction dies off as[106]) $O(q^{-11-N/3}\ \ln^p q)$ as the energy q becomes large, where p is calculable. Obviously, the next problem is to attempt to compute instanton corrections to moments of the structure functions for deep inelastic leptoproduction. In other applications, the connection with the short-distance region is less clear or is simply assumed:

a) The effects of instantons on the hadronic spectrum have been considered from various points of view[108-112]). The most convincing calculations involve the heavy quarks c, t, b, ..., because it can be argued that distances $O(M^{-1}_{c,t,b})$ dominate, as in charmonium models. In particular, it has been suggested that the relatively large splitting between $\eta_c(2830)$ and $J/\psi(3100)$, which is hard to explain in the usual charmonium picture, should be attributed to instanton effects[108]).

b) It is an important problem[14,113]) to demonstrate spontaneous breaking of the chiral SU(L) × SU(L) symmetry of QCD in a self-consistent calculation. [This is related to Eq. (10.3) and the U(1) problem.] Here there seems to be little connection with the short-distance region: the size integral has to be cut off for purely pragmatic reasons. The main difficulty is to satisfy the self-consistency requirement that the coupled QCD equations of motion be preserved.

The concluding remark is related to comments at the end of Section 5: the temptation to rely on diagrammatic analogies as a source of intuition should be resisted. Of course, nobody supposes that one-instanton amplitudes are given by a sum of Feynman diagrams, since exp $- 8\pi^2/g^2$ has no Taylor expansion about $g = 0$.

Nevertheless, there is a tendency to adopt a similar diagrammatic approach to dilute-gas amplitudes which are assumed to be given by products of amplitudes for individual instantons and free-particle propagators. In general, this assumption is not correct. For example, consider the propagator

$$S_{gas} = \left(\not{D}_{gas} + M \right)^{-1} \delta(x-y) \tag{10.4}$$

for a quark with a small mass M passing through a dilute gas which contains both instantons and anti-instantons. The mass singularity of this propagator is at most $O(M^{-1})$, irrespective of the (finite) number of zero eigenvalues of $\not{D}_{gas}$:

$$S_{gas} = O\left(M^{-1}\right) \tag{10.5}$$

It is therefore not possible to suppose that S_{gas} is the sum of products of individual propagators $(\not{D}_i^{\pm} + M)^{-1}\delta(x - y)$ for the i^{th} instanton or anti-instanton, because there are terms in which instantons and anti-instantons alternate. These terms produce strong singularities at $M = 0$:

$$\begin{aligned} &(\not{D}_1^{+} + M)^{-1}(\not{\partial} + M)(\not{D}_2^{-} + M)(\not{\partial} + M) \;\ldots\; \ldots (\not{\partial} + M)(\not{D}_{2j}^{-} + M)\,\delta(x-y) \\ &\qquad \sim \quad (\text{constant} \neq 0)\; M^{-2j} \end{aligned} \tag{10.6}$$

Wherever possible, one should analytically expand expressions systematically derived from the appropriate functional integrals.

Acknowledgements

I thank N. Christ, M.K. Gaillard, P. Goddard, H. Römer and A. Ukawa for comments. Part of this work was prepared while visiting the Rockefeller University. I thank Professor M.A.B. Bég for his hospitality and the Rockefeller theory group for discussions.

Note Added

For an extensive discussion of the U(1) problem, see: R.J. Crewther, CERN preprint TH 2546 (1978).

APPENDIX

The analysis of Eq. (2.27) in the text is restricted to the gauge group SU(2). This Appendix contains a non-rigorous discussion*) of the extension to any simple, compact Lie group $\mathcal{G}$. Simple Lie groups have no connected invariant subgroups. They consist of the classical groups

$$\mathcal{G} \quad = \quad SO(n)\ ,\ (n \geqslant 5)\ ;\ SU(n)\ ,\ (n \geqslant 2)\ ;\ Sp(n)\ ,\ (n \geqslant 1) \tag{A.1}$$

and the exceptional groups

$$\mathcal{G} \quad = \quad E_6\ ,\ E_7\ ,\ E_8\ ,\ F_4\ ,\ G_2 \tag{A.2}$$

The orthogonal group SO(n + 1) preserves the length of real (n + 1)-dimensional vectors

$$x \quad = \quad \left(x_1, \ldots , x_{n+1}\right) \tag{A.3}$$

so x can be confined to the hypersphere S_n. Let U_x be an SO(n + 1) transformation which rotates x to the N-pole vector

$$\begin{aligned} x_{N\ pole} &= \text{column vector } \left(0, 0, \ldots , 0, 1\right) \\ &= U_x\, x \end{aligned} \tag{A.4}$$

The maximal set of SO(n + 1) matrices which leaves the N pole invariant is given by the SO(n) subgroup {M} formed by matrices

$$M \quad = \quad \begin{pmatrix} SO(n)\ \text{matrix} & 0 \\ 0 & 1 \end{pmatrix} \tag{A.5}$$

This means that the point x is left invariant by the SO(n) subgroup $\{U_x^T M U_x;\ \text{all } M\}$. It also means that the set of all SO(n + 1) matrices can be divided into cosets

$$C_x \quad = \quad \left\{ U_x^T M\ ;\ \text{all } M \right\} \tag{A.6}$$

Since x labels cosets and also labels points on the n-hypersphere, we find:

$$SO(n+1)/SO(n) \quad = \quad S_n \tag{A.7}$$

The same argument works for the other classical groups SU(n), Sp(n). The group SU(n + 1) rotates the complex vector

---

*) Which is trivial for mathematicians[114]. I thank H. Römer for providing essential instructions.

$$z = (z_1, \ldots, z_{n+1}) \qquad (A.8)$$

on the hypersphere S_{2n+1} in (n + 1) complex dimensions, i.e. in (2n + 2) real dimensions. Each point z is left invariant by a subgroup SU(n), so the coset formula becomes

$$SU(n+1)/SU(n) = S_{2n+1} \qquad (A.9)$$

Similarly, the symplectic group Sp(n + 1) of rotations of quaternionic vectors

$$q = (q_1, \ldots, q_{n+1}) \qquad (A.10)$$

on the (4n + 3)-hypersphere contains an invariance subgroup Sp(n) for each point q:

$$Sp(n+1)/Sp(n) = S_{4n+3} \qquad (A.11)$$

Consideration of the exceptional groups is postponed for the moment.

The mapping of S_3 into group space defines a mapping of S_3 into the corresponding coset space and hence into a hypersphere given by Eqs. (A.7), (A.9), or (A.11). The latter mapping

$$S_3 \to S_m \quad , \quad (m > 3) \qquad (A.12)$$

has the property that it can be continuously deformed to a single point P on S_m: there has to be some point Q on S_m not covered by the mapping (A.12) (because S_m has more dimensions than S_3), so the "hole" in S_m created by removing Q can be continuously expanded over the surface until only P remains. [Compare this with examples 2,5 in Sections 3.4, 3.5 of Coleman's 1975 lectures[115].]

The deformation of the coset mapping (A.12) to the trivial mapping onto the point P corresponds to a continuous deformation of

$$S_3 \to SO(n+1)\ ,\ SU(n+1)\ ,\ \text{or}\ \ Sp(n+1) \qquad (A.13)$$

to the mapping

$$S_3 \to \text{subgroup } (SO(n), SU(n), \text{ or } Sp(n)) \text{ leaving P invariant} \qquad (A.14)$$

Successive deformations result in the value of n being reduced as far as permitted by the constraint m > 3 in Eq. (A.12). For any group $\mathcal{G}$ in Eq. (A.1), there is a suitable sequence of subgroups which ends with SU(2):

$$\begin{aligned} &SU(n) \supset SU(n-1) \supset \ldots \supset SU(2) \\ &Sp(n) \supset Sp(n-1) \supset \ldots \supset Sp(1) \equiv SU(2) \\ &SO(n) \supset SO(n-1) \supset \ldots \supset SO(5) \equiv Sp(2) \supset Sp(1) \equiv SU(2) \end{aligned} \qquad (A.15)$$

Hence a mapping of S_3 to a group $\mathcal{G}$ in Eq. (A.1) is related by continuous deformation to a mapping of S_3 to an SU(2) subgroup. It follows from the deformation invariance equation (2.38) that the result (2.36) is still valid.

In Eq. (A.15), the use of Eq. (A.7) for orthogonal groups stops at SO(5), even though Eq. (A.12) permits a further deformation to SO(4). It is necessary to avoid SO(4) because its direct-product structure

$$SO(4) \quad = \quad SU(2) \times SU(2) \tag{A.16}$$

does not (in general) permit deformation to an SU(2) subgroup.

The deformation result is also true for the exceptional groups (A.2), but there is no need to prove it here. All of these groups are subgroups of SU(n) for n sufficiently large. Given that G in Eq. (2.27) is an element of an exceptional group, we simply reinterpret it as an element of SU(n) and conclude from the preceding analysis that the topological charge ν must be an integer.

REFERENCES

1) A.A. Belavin, A.M. Polyakov, A.S. Schwarz and Yu.S. Tyupkin, Phys. Letters 59B, 85 (1975).

2) G. 't Hooft, Phys. Rev. Letters 37, 8 (1976).

3) G. 't Hooft, Phys. Rev. D 14, 3432 (1976).

4) R. Jackiw and C. Rebbi, Phys. Rev. Letters 37, 172 (1976).

5) C.G. Callan, Jr., R.F. Dashen and D.J. Gross, Phys. Letters 63B, 334 (1976).

6) V.N. Gribov (unpublished).

7) V.G. Drinfeld and Yu.I. Manin, Funk. Anal. i. Prilozen (to be published), and Uspehi Mat. Nauk. (to be published).

8) M.F. Atiyah, N.J. Hitchin, V.G. Drinfeld and Yu.I. Manin, Phys. Letters 65A, 185 (1978).

9) R.D. Peccei and H.R. Quinn, Phys. Rev. Letters 38, 1440 (1977); Phys. Rev. D 16, 1791 (1977).

10) S. Weinberg, Phys. Rev. Letters 40, 223 (1978).

11) F. Wilczek, Phys. Rev. Letters 40, 279 (1978).

12) R.J. Crewther, Phys. Letters 70B, 349 (1977).

13) V. de Alfaro, S. Fubini and G. Furlan, Phys. Letters 65B, 163 (1977); C.G. Callan, Jr., R. Dashen and D.J. Gross, Phys. Letters 66B, 375 (1977).

14) C.G. Callan, Jr., R. Dashen and D.J. Gross, Phys. Rev. D 17, 2717 (1978).

15) B. Schroer, Lectures at this School (1978).

16) W. Thirring, Lectures at this School (1978).

17) E.S. Abers and B.W. Lee, Phys. Reports 9C, 1 (1973).

18) W. Marciano and H. Pagels, Phys. Reports 36C, 137 (1978).

19) W.A. Bardeen, Nuclear Phys. B75, 246 (1974).

20) M.F. Atiyah and R.S. Ward, Comm. Math. Phys. 55, 117 (1977).

21) S.L. Adler, Phys. Rev. D 6, 3445 (1972); D 8, 2400 (1973); R. Jackiw and C. Rebbi, Phys. Rev. D 14, 517 (1976).

22) M.F. Atiyah and I.M. Singer, Bull. Amer. Math. Soc. 69, 422 (1963); Ann. Math. 87, 484, 546 (1968).

23) S. Coleman (unpublished); J. Kiskis, Phys. Rev. D 15, 2329 (1977).

24) A.S. Schwarz, Proc. Symposium in Alushta, USSR (1976), JINR preprint D2-9788, p. 224; Phys. Letters 67B, 172 (1977).

25) S.W. Hawking, Phys. Letters 60A, 81 (1977).

26) R. Jackiw and C. Rebbi, Phys. Rev. D 16, 1052 (1977).

27) N.K. Nielsen and B. Schroer, Nuclear Phys. B127, 493 (1977).

28) E.F. Corrigan and D.B. Fairlie, Phys. Letters 67B, 69 (1977);
F. Wilczek, *in* Quark confinement and field theory, (eds. D. Stump and D. Weingarten) (New York, J. Wiley and Sons, 1977).

29) R. Jackiw, C. Nohl and C. Rebbi, Phys. Rev. D 15, 1642 (1977), and Proc. 1977 Banff School (Plenum Press).

30) E. Witten, Phys. Rev. Letters 38, 121 (1977).

31) G. 't Hooft, unpublished.

32) S. Sciuto, Riv. Nuovo Cimento (to be published);
J.J. Giambiagi and K.D. Rothe, Nuclear Phys. B129, 111 (1977).

33) M.F. Atiyah, N.J. Hitchin and I.M. Singer, Proc. Nat. Acad. Sci. 74, 262 (1977);
C.W. Bernard, N.H. Christ, A.H. Guth and E.J. Weinberg, Phys. Rev. D 16, 2967 (1977).

34) R. Jackiw and C. Rebbi, Phys. Letters 67B, 189 (1977).

35) L.S. Brown, R.D. Carlitz and C. Lee, Phys. Rev. D 16, 417 (1977).

36) E.F. Corrigan, D.B. Fairlie, S. Templeton and P. Goddard, ENS (Paris) preprint 78/13 (1978).

37) H. Osborn, DAMTP preprint 78-13 (1978).

38) N.H. Christ, E.J. Weinberg and N.K. Stanton, Columbia Univ. preprint CU-TP-119 (1978).

39) L.S. Brown, R.D. Carlitz, D.B. Creamer and C. Lee, Phys. Letters 70B, 180 (1977) or 71B, 103 (1977); Phys. Rev. D 17, 1583 (1978);
D.B. Creamer, Phys. Rev. D 16, 3496 (1977);
L.S. Brown and C. Lee, Univ. Washington preprint RLO 1388-756 (1978).

40) R. Flume, Saclay preprint (1978);
M. Daniel, P.K. Mitter and C. Viallet (unpublished).

41) D. Förster, Phys. Letters 66B, 279 (1977).

42) E. Witten, Harvard preprint (1978).

43) R.P. Feynman, Rev. Mod. Phys. 20, 267 (1948); Phys. Rev. 80, 440 (1950); 84, 108 (1951);
R.P. Feynman and A.R. Hibbs, Quantum mechanics and path integrals (McGraw-Hill, New York, 1965);
I.M. Gel'fand and A.M. Yaglom, J. Math. Phys. 1, 48 (1960);
L.D. Faddeev, *in* Methods in field theory, Les Houches Session XXVIII (1975) (eds. R. Balian and J. Zinn-Justin) (North-Holland, 1976);
V.N. Popov, CERN preprint TH 2424 (1977).

44) G.W. Gibbons, Phys. Letters 61A, 3 (1977).

45) A.A. Belavin and A.M. Polyakov, Nuclear Phys. B123, 429 (1977).

46) F. Ore, Phys. Rev. D 16, 2577 (1977).

47) S. Chadha, A. d'Adda, P. di Vecchia and F. Nicodemi, Phys. Letters 72B, 103 (1977).

48) R.P. Feynman, Acta Phys. Polon. 24, 697 (1963);
B. de Witt, Phys. Rev. Letters 12, 742 (1964) and Phys. Rev. 162, 1195, 1239 (1967);
L.D. Faddeev and V.N. Popov, Phys. Letters 25B, 29 (1967);
L.D. Faddeev, Kiev ITP Report 67-36 [Trans. FNAL-THY-57 (1973)];
S. Mandelstam, Phys. Rev. 175, 1580 (1968).

49) V.N. Gribov, Lecture at the 12th Winter School of the Leningrad Nuclear Physics Institute, Vol. 1, p. 147 (1977) (in Russian); Quantization of non-Abelian gauge theories, LNPI preprint (1977).
See also S. Mandelstam, talk presented at APS Meeting, April (1977).

50) S. Sciuto, Phys. Letters 71B, 129 (1977);
M. Ademollo, E. Napolitano and S. Sciuto, Nuclear Phys. B134, 477 (1978);
L.F. Abbott and T. Eguchi, Phys. Letters 72B, 215 (1977);
C. Bender, T. Eguchi and H. Pagels, Phys. Rev. D 17, 1086 (1978);
R. Jackiw, I. Muzinich and C. Rebbi, Phys. Rev. D 17, 1576 (1978);
R.D. Peccei, Phys. Rev. D 17, 1097 (1978);
D. Amati and A. Rouet, Phys. Letters 73B, 39 (1978), and CERN preprint TH 2468 (1978);
I.M. Singer, Comm. Math. Phys. 60, 7 (1978);
M.B. Halpern and J. Koplik, Nuclear Phys. B132, 239 (1978);
H.D.I. Abarbanel and J. Bartels, Nuclear Phys. B136, 237 (1978);
A.P. Balachandran, H.S. Mani, R. Ramachandran and P. Saran, Syracuse Univ. preprint COO-3533-108 (1978);
C. Montonen, CERN preprint TH 2477 (1978).

51) S.W. Hawking, Comm. Math. Phys. 55, 133 (1977);
G.W. Gibbons, Phys. Letters 60A, 385 (1977);
A.S. Schwarz, Lett. Math. Phys. 2, 201 (1978).

52) F.A. Berezin, Method of second quantization (Academic Press, New York and London, 1966).

53) S.L. Adler, Institute for Advanced Study (Princeton) preprints (1978).

54) A.M. Polyakov, Phys. Letters 59B, 82 (1975); Nuclear Phys. B120, 429 (1977).

55) S. Coleman, 1977 Erice Lectures, Harvard Univ. preprint (1978).

56) J.L. Gervais and B. Sakita, Phys. Rev. D 11, 2943 (1975);
J.L. Gervais, A. Jevicki and B. Sakita, Phys. Rev. D 12, 1038 (1975);
C. Callan and D. Gross, Nuclear Phys. B93, 29 (1975);
E. Tomboulis, Phys. Rev. D 12, 1678 (1975);
N.H. Christ and T.D. Lee, Phys. Rev. D 12, 1606 (1975).

57) A. Jevicki, Nuclear Phys. B117, 365 (1976).

58) K. Lane, Phys. Rev. D 10, 1353 (1974);
D.J. Gross and A. Neveu, Phys. Rev. D 10, 3235 (1974);
T. Banks, Phys. Rev. D 11, 2876 (1975).

59) A. Patrascioiu, Phys. Rev. D 17, 2764 (1978).

60) J. Hietarinta, W.F. Palmer and S.S. Pinsky, Phys. Rev. Letters 38, 103 (1977);
R.D. Peccei and H.R. Quinn, Nuovo Cimento 41A, 309 (1977).

61) T. Appelquist and R. Shankar, Yale Univ. preprint (1978).

62) B. Grossmann, Phys. Letters 61A, 86 (1977).

63) P. Carruthers, S.S. Pinsky and F. Zachariasen, Nuclear Phys. B135, 486 (1978).

64) S. Wadia and T. Yoneya, Phys. Letters 66B, 341 (1977).

65) J. Lowenstein and J.A. Swieca, Ann. Phys. (N.Y.) 68, 172 (1971);
S. Coleman, R. Jackiw and L. Susskind, Ann. Phys. (N.Y.) 93, 267 (1975);
S. Coleman, Ann. Phys. (N.Y.) 101, 239 (1976).

66) L. Faddeev, Proc. Symposium in Alushta, USSR (1976).

67) D. Olive, Riv. Nuovo Cimento (to be published).

68) J.S. Bell and R. Jackiw, Nuovo Cimento 60A, 47 (1969).

69) S.L. Adler, Phys. Rev. 177, 2426 (1969).

70) J. Schwinger, Phys. Rev. 82, 664 (1951).

71) S.L. Adler, Lectures on Elementary Particles and Quantum Field Theory, Brandeis University Summer Institute (Massachusetts Institute of Technology Press, Cambridge, Mass., 1970), Vol. I.

72) R. Jackiw, Lectures in Current Algebra and its Applications (Princeton University Press, Princeton, N.J., 1972).

73) J. Kogut and L. Susskind, Phys. Rev. D 11, 3594 (1975).

74) S. Weinberg, Phys. Rev. D 11, 3583 (1975).

75) P.B. Gilkey, Adv. Math. 15, 334 (1975).

76) A.L. Mason, Nuovo Cimento 4A, 749 (1971).

77) W. Zimmermann, Lectures on Elementary Particles and Quantum Field Theory, Brandeis University Summer Institute (Massachusetts Institute of Technology Press, Cambridge, Mass., 1970), Vol. I; Ann. Phys. (N.Y.) 77, 536, 570 (1973).

78) S. Weinberg, Phys. Rev. Letters 31, 494 (1973); Phys. Rev. D 8, 4482 (1973).

79) H. Fritzsch, Lectures at this School, CERN preprint TH 2483 (1978).

80) M. Kobayashi and K. Maskawa, Prog. Th. Phys. 49, 552 (1973).

81) W.A. Bardeen and S.-H.H. Tye, Phys. Letters 74B, 229 (1978).

82) J. Kadaswamy, P. Salomonson and J. Schechter, Syracuse Univ. preprint SU-4211-109 (1978).

83) V. Baluni, Phys. Letters 40, 1358 (1978).

84) T. Goldman and C.M. Hoffmann, Phys. Rev. Letters 40, 220 (1978);
J. Kadaswamy, P. Salomonson and J. Schechter, Phys. Letters 74B, 377 (1978).

85) C. Baltay, G. Feinberg and M. Goldhaber, unpublished.

86) E. Paschos, unpublished.

87) J. Ellis and M.K. Gaillard, Phys. Letters 74B, 374 (1978).

88) G. Michelmacher and B. Pontecorvo, Dubna preprint E1-11318 (1978).

89) T.W. Donnelly, S.J. Freeman, R.S. Lytel, R.D. Peccei and M. Schwartz, Stanford Univ. preprint ITP-598 (1978).

90) S.B. Treiman and F. Wilczek, Phys. Letters 74B, 381 (1978).

91) W.A. Bardeen, S.-H.H. Tye and J.A.M. Vermaseren, Fermilab-Pub 78/20-THY (1978).

92) D.W. McKay and H. Munczek, Kansas Univ. preprint (1978).

93) K.O. Milaelian, Fermilab-Pubs 78-32, 78-34 THY (1978).

94) H.S. Gurr, F. Reines and H.W. Sobel, Phys. Rev. Letters 33, 179 (1974).

95) P. Alibran et al., Phys. Letters 74B, 134 (1978);
T. Hansl et al, Phys. Letters 74B, 139 (1978);
P.C. Bosetti et al., Phys. Letters 74B, 143 (1978).

96) F. Reines, H.S. Gurr and H.W. Sobel, Phys. Rev. Letters 37, 315 (1976).

97) E. Bellotti, E. Fiorini and L. Zanotti, Phys. Letters 76B, 223 (1978).

98) S. Weinberg, *in* A Festschrift for I.I. Rabi (ed. L. Motz) (New York Academy of Sciences, New York, 1977).

99) S. Weinberg, Harvard Univ. preprint HUTP-78/A005 (1978);
A. Zepeda, C.I.E.A.-I.P.N. (Mexico) preprint (1978);
C.A. Domínguez, C.I.E.A.-I.P.N. (Mexico) preprint (1978);
N.G. Deshpande and D.E. Soper, Oregon preprint (1978).

100) R.J. Crewther, Weak and electromagnetic interactions at high energies, Cargèse 1975, Part A (eds. M. Lévy, J.-L. Basdevant, D. Speiser and R. Gastmans) (Plenum Press, New York and London, 1976), p. 345.

101) H. Georgi, Hadronic J. 1, 155 (1978).

102) M.A.B. Bég and H.-S. Tsao, Rockefeller University preprint COO-2232B-154 (1978).

103) S.L. Glashow, Hadrons and their interactions, 1967 Erice Lectures (Academic Press Inc., New York, 1968), p. 83.

104) N. Andrei and D.J. Gross, Princeton Univ. preprint (1978).

105) R.D. Carlitz and C. Lee, Univ. of Pittsburgh preprint PITT-193 (1978).

106) L. Baulieu, J. Ellis, M.K. Gaillard and W.J. Zakrzewski, CERN preprint TH 2482 (1978).

107) S.-S. Shei, New York Univ. preprint NYU-TR 78/3 (1978).

108) F. Wilczek and A. Zee, Phys. Rev. Letters 40, 83 (1978);
M.A. Shifman, A.I. Vainshtein, M.B. Voloshin and V.I. Zakharov, ITEP (Moscow) preprint 10 (1978).

109) A. Duncan, Columbia Univ. preprint CU-TP 116 (1978).

110) D. Horn and S. Yankielowicz, Phys. Letters 76B, 343 (1978).

111) C.G. Callan, R. Dashen, D.J. Gross, F. Wilczek and A. Zee, IAS (Princeton) preprint COO-2220-132 (1978).

112) M. Suzuki, UC Berkeley preprint (1978).

113) D.G. Caldi, Phys. Rev. Letters 39, 121 (1977);
R.D. Carlitz, University of Pittsburgh preprint PITT-186 (1977);
A. Ukawa (unpublished).

114) N. Steenrod, The topology of fibre bundles, No. 14, Princeton Mathematics Series (Princeton Univ. Press, Princeton, N.J., 1951).

115) S. Coleman, *in* New phenomena in subnuclear physics (ed. A. Zichichi) (Plenum Press, New York, 1977), p. 297.

Figure captions

Fig. 1 : Mapping of surface of hypersphere in four dimensions onto another hypersphere. See Eq. (2.31).

Fig. 2 : Mapping of R_4 (flat four-dimensional Euclidean space) onto the surface S_4, the unit hypersphere in five dimensions.

Fig. 3 : Gauge patches for the potential $\hat{A}$ on S_4.

Fig. 4 : Plot of the action S as a functional of the quantum field A^{qu}. The base of the trough is flat due to the existence of a continuous parameter Λ in the set of classical solutions A^{cl}.

Fig. 5 : Anomalous triangle diagram for the γ_5 transformations (7.1).

Fig. 6 : Diagrams responsible for wave-function renormalization of $J_{\mu 5}$.

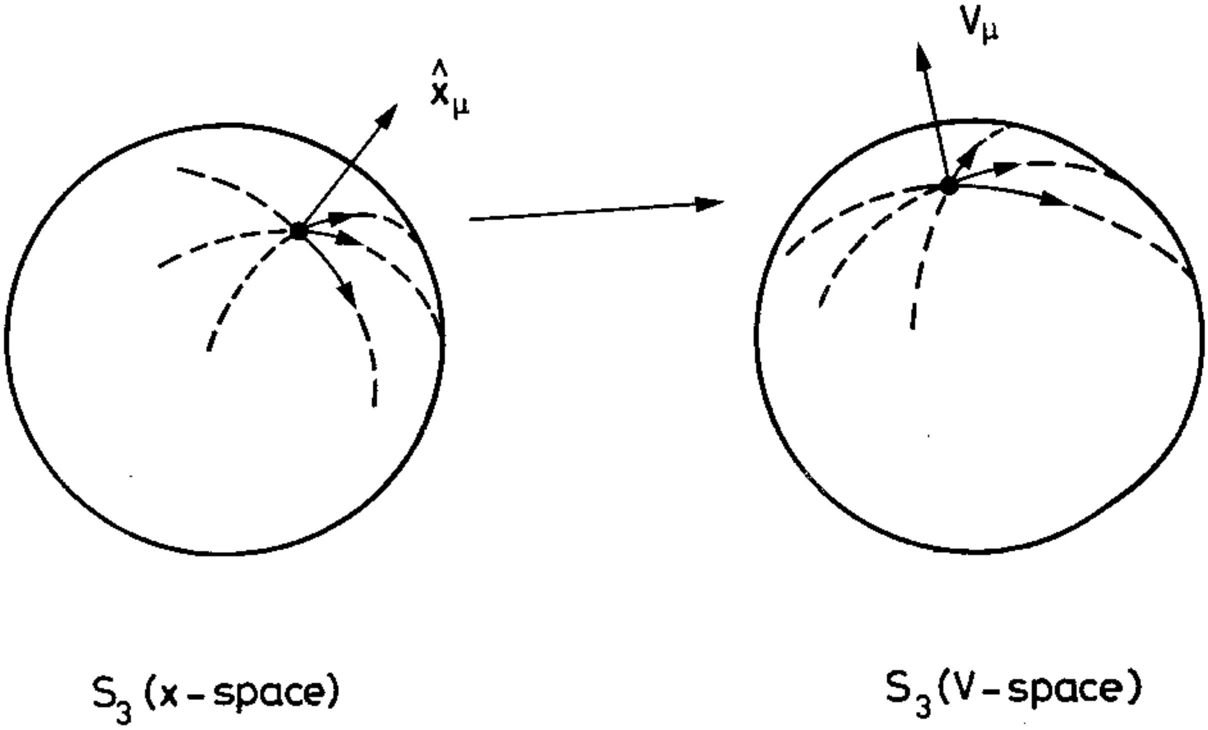

FIG. 1

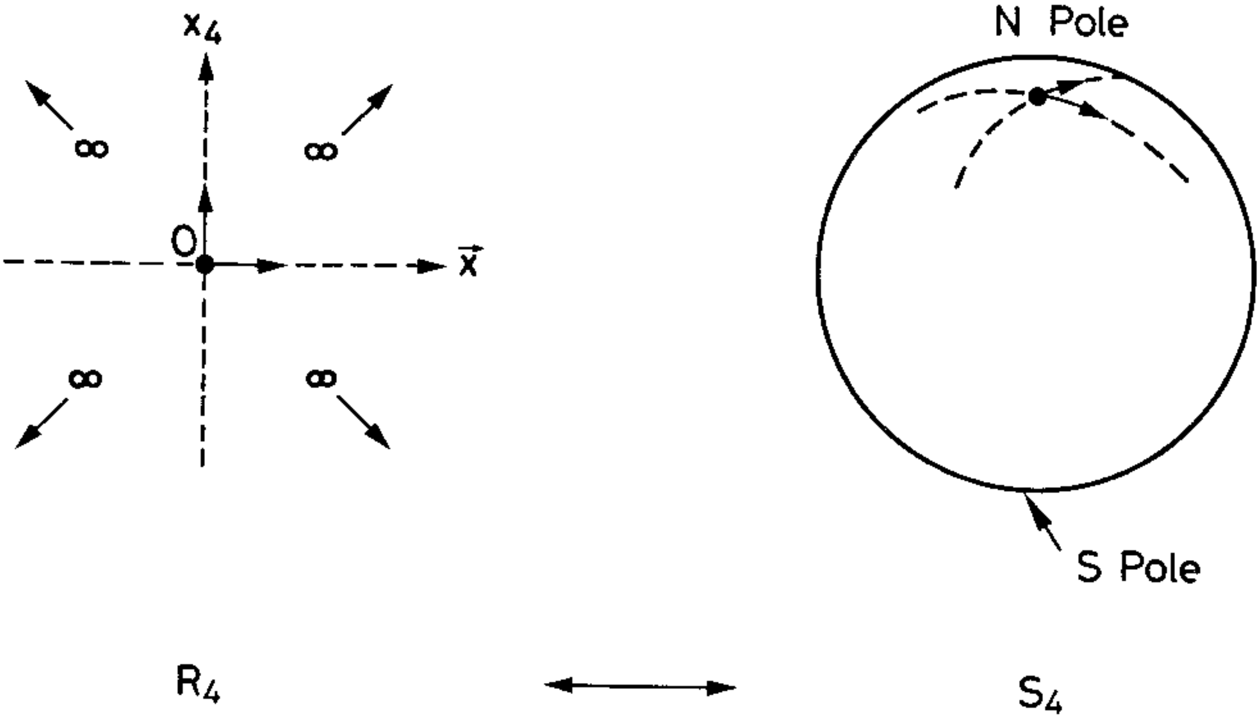

FIG.2

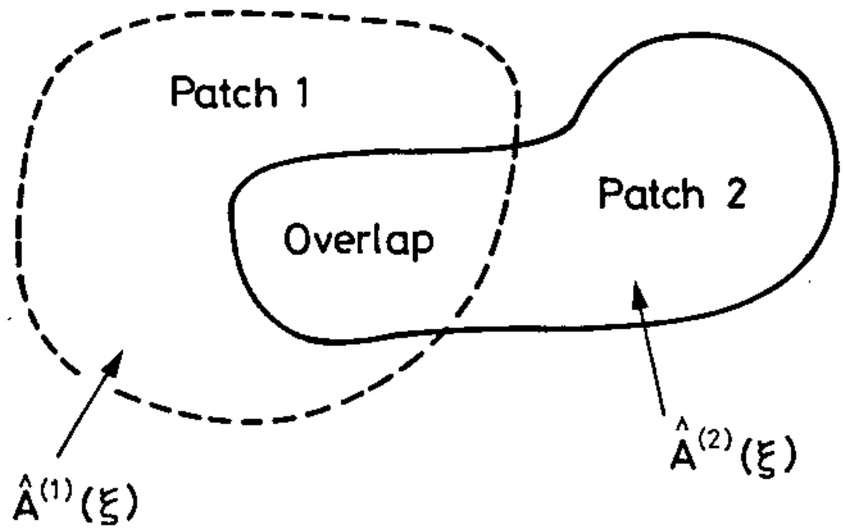

FIG.3

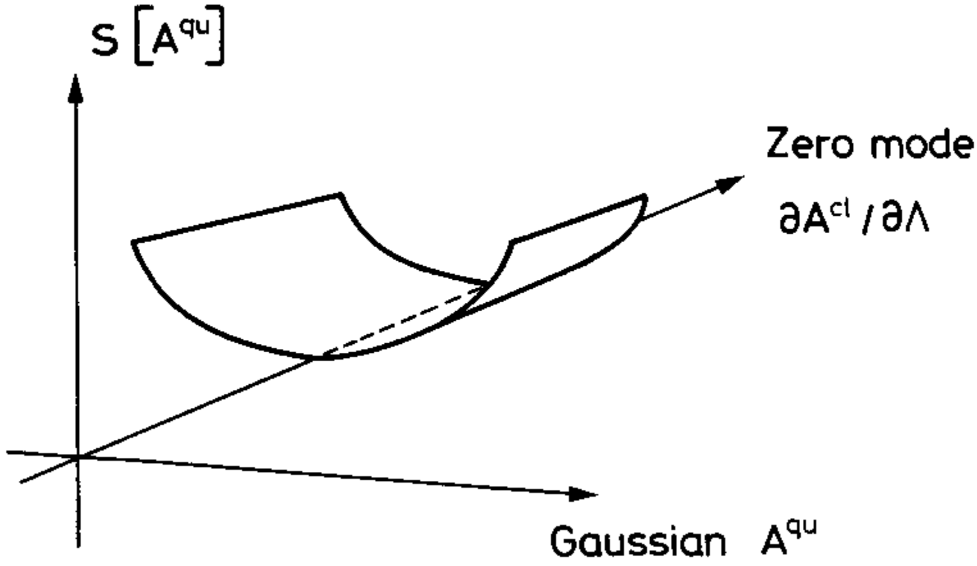

FIG.4

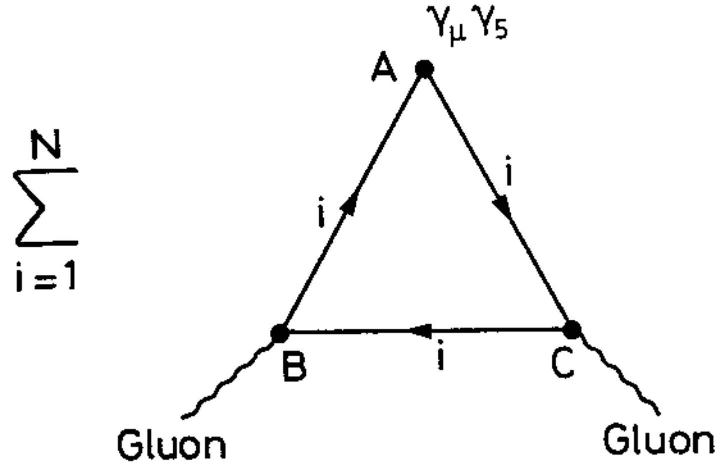

FIG.5

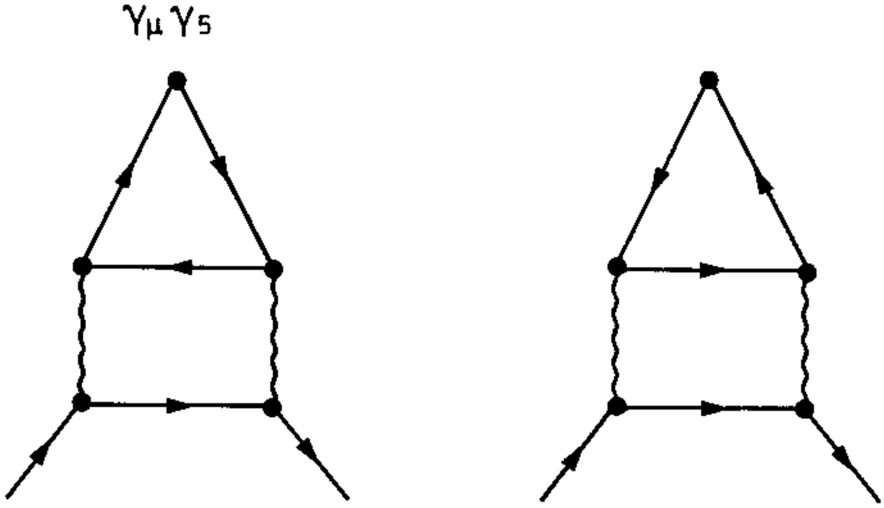

FIG.6