E. Triangle Anomaly

(R. Jackiw, in Lectures on Current Alg. S. Adler, Brandeis 1970)

The symmetries of a classical field theory may fail to survive quantization. This can occur as a consequence of infinities in the theory; there is a need for regularization, and there may be no means of regularization which preserves the symmetry.

We have already encountered one example: A massless renormalizable field theory is scale-invariant at the classical level, but a mass scale enters the quantum theory (dimensional transmutation). The implications were profound.

Now we will consider another example of an "anomaly", a symmetry of a classical field theory which does not survive in the quantum theory. We will find that certain chiral symmetries are afflicted with anomalies.

We begin by considering free field theory

$\mathcal{L} = \bar{\psi} i \partial\!\!\!/ \psi$; ψ is a (massless) 4-component fermion.

This theory has a $U(1)_V \times U(1)_A$ symmetry, with associated conserved currents

$$J_\mu = \bar{\psi}\gamma_\mu\psi \qquad \partial^\mu J_\mu = 0$$
$$J_{5\mu} = \bar{\psi}\gamma_5\gamma_\mu\psi \qquad \partial^\mu J_{5\mu} = 0$$

Let us consider the 3-point function

$$-i\Gamma_{\mu\nu\lambda}(x_1, x_2 x_3) = \langle 0| T[J_{5\mu}(x_3) J_\nu(x_1) J_\lambda(x_2)]|0\rangle$$

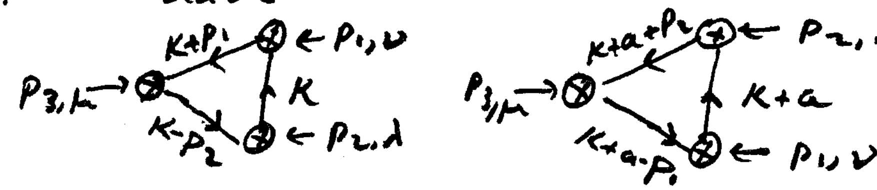
(Note: really a "T^*-product")

This is expected to obey Ward Identities

$$\frac{\partial}{\partial x_{3\mu}}\Gamma_{\mu\nu\lambda} = \frac{\partial}{\partial x_{1\nu}}\Gamma_{\mu\nu\lambda} = \frac{\partial}{\partial x_{2\lambda}}\Gamma_{\mu\nu\lambda} = 0,$$

because the currents are conserved and commute with each other.

But now let us explicitly compute the divergence of Γ. We have

$-i\Gamma =$ 

or $-i\Gamma_{\mu\nu\lambda}(p_1, p_2, p_3)$

$$= \underset{\text{fermi}}{(-1)}(i)^3 \int \frac{d^4K}{(2\pi)^4}\frac{1}{K^2}\frac{1}{(K+p_1)^2}\frac{1}{(K-p_2)^2}\, K\!\!\!/\,\gamma_\nu K\!\!\!/\,\gamma_\lambda (K\!\!\!/-p\!\!\!/_2)\gamma_5\gamma_\mu(K\!\!\!/+p\!\!\!/_1)$$

$$+ \quad \begin{matrix} p_1 \leftrightarrow p_2 \\ \lambda \leftrightarrow \nu \\ K \to K+a \end{matrix}$$

Now we will consider $-ip_3^\mu\Gamma_{\mu\nu\lambda}$. To evaluate it use the identity

$$\gamma_\mu p_3^\mu = \gamma_\mu(-p_1 - p_2)^\mu = (K\!\!\!/ - p\!\!\!/_2) - (K\!\!\!/ + p\!\!\!/_1)$$

so $$-ip_3^\mu \Gamma_{\mu\nu\lambda} = i\int \frac{d^4k}{(2\pi)^4}\left[-\frac{\mathrm{tr}\,\gamma_\nu \not k \gamma_\lambda (\not k - \not p_2)\gamma_5}{k^2(k-p_2)^2}\right.$$
$$\left. - \frac{\mathrm{tr}\,\gamma_\nu \not k \gamma_\lambda \gamma_5 (\not k + \not p_1)}{k^2(k+p_1)^2} + \begin{matrix} p_1 \leftrightarrow p_2 \\ \nu \leftrightarrow \lambda \\ k \to k+a \end{matrix}\right]$$

and $\mathrm{tr}\,\gamma_\nu\gamma_\alpha\gamma_\lambda\gamma_\beta\gamma_5 = 4i\,\epsilon_{\nu\alpha\lambda\beta}$

$$-ip_3^\mu \Gamma_{\mu\nu\lambda} = 4\epsilon_{\nu\alpha\lambda\beta}\int\frac{d^4k}{(2\pi)^4}\left[\frac{k^\alpha(k-p_2)^\beta}{k^2(k-p_2)^2} - \frac{k^\alpha(k+p_1)^\beta}{k^2(k+p_1)^2}\right.$$
$$\left. + \frac{(k+a)^\alpha(k+a+p_2)^\beta}{(k+a)^2(k+a+p_2)^2} - \frac{(k+a)^\alpha(k+a-p_1)^\beta}{(k+a)^2(k+a-p_1)^2}\right]$$

We can now verify the naive Ward identity by the following formal manipulation. In the first term replace k by $k+a+p_2$; in the second term replace k by $k+a-p_1$. Then these are the same as the third and fourth terms, but with α and β interchanged. Thus, the four terms sum to zero.

But the integral is quadratically divergent, and it is not at all obvious that we can justify shifting the origin of the k integration term by term.

In fact, suppose the (Euclidean space) integral
$$I(0) = \int d^4k\, f(k)$$
is linearly divergent. Then the formal manipulation says
$$I(a) - I(0) = \int d^4k\,[f(k+a) - f(k)]$$
vanishes, while actually
$$I(a) - I(0) = a^\mu \int d^4k\, \partial_\mu f = a^\mu \int d\Omega_k\, k^2 k_\mu f\Big|_{k^2=\infty}$$
— = "surface term"

If e.g. $f^\nu(k) = \frac{k^\nu}{(k^2+b^2)^2}$, then

$$I(a) - I(0) = a^\mu \int d\Omega_k \, k^3 \frac{k_\mu k^\nu}{k^4}$$

May replace $k_\mu k^\nu$ by $\frac{1}{4}\delta_\mu^\nu$, and since volume of sphere is $2\pi^2$, we have

$$I(a) - I(0) = \frac{\pi^2}{2} a^\nu \neq 0$$

Exercise 3.2

$$\boxed{\begin{array}{l}\text{Evaluate the relevant surface term in the quadratically divergent expression for } p_3^\mu \Gamma_{\mu\nu\lambda} \text{ to show} \\ \qquad p_3^\mu \Gamma_{\mu\nu\lambda} = \frac{-1}{8\pi^2} \epsilon_{\nu\lambda\alpha\beta} a^\alpha p_3^\beta \\ \text{(and check that the sign is correct!)}\end{array}}$$

It is now easy to evaluate the other divergences of $\Gamma_{\mu\nu\lambda}$. We can commute γ_5 to any vertex without a sign change. And a becomes replaced by difference between value of loop momentum opposite the vertex in the two graphs, so

$$\boxed{\begin{aligned} 8\pi^2 p_1^\nu \Gamma_{\mu\nu\lambda} &= -\epsilon_{\lambda\mu\alpha\beta} (a+2p_2)^\alpha p_1^\beta \\ 8\pi^2 p_2^\lambda \Gamma_{\mu\nu\lambda} &= -\epsilon_{\mu\nu\alpha\beta} (a-2p_1)^\alpha p_2^\beta \end{aligned}}$$

How can the answers depend on the arbitrary four-vector a? Apparently a is ill-defined, because of the very singular structure of the theory at short distances. This shows up as a linear divergence in the perturbative evaluation of Γ. The divergence must be regulated, but there is no method of regularization which preserves both the $U(1)_V$ and $U(1)_A$ symmetries. (Pauli-Villars requires massive fermions; dimensional regularization cannot be applied, because γ_5 has no dimensional continuation)

We must impose some sort of convention to define Γ. E.g. suppose that we decide that the vector current Ward identities should be satisfied. This requires

$$\begin{aligned} a + 2p_2 &= c\,p_1 \\ a - 2p_1 &= d\,p_2 \end{aligned} \Rightarrow 2(p_2+p_1) = c\,p_1 - d\,p_2$$

$$\Rightarrow c = 2,\ d = -2$$

$$\text{or } a = 2(p_1 - p_2)$$

Then we have $(p_3 = -p_1 - p_2)$

$$\boxed{p_3^{\mu}\Gamma_{\mu\nu\lambda} = \frac{1}{2\pi^2}\epsilon_{\nu\lambda\alpha\beta}\,p_1^{\alpha}p_2^{\beta}}$$ — vector Ward identity satisfied

— the axial vector current suffers a Ward-identity anomaly. (Note that this expression is symmetric under crossing, $p_1 \leftrightarrow p_2$, $\lambda \leftrightarrow \nu$, as it should be)

The anomaly is a polynomial in momentum space, or a derivative of a δ function in position space:

$$\frac{\partial}{\partial x_{3\mu}}\Gamma_{\mu\nu\lambda}(x_1, x_2, x_3) \;\propto\; \delta'(x_1 - x_3)\,\delta'(x_2 - x_3)$$

This is what we expect, if the anomaly is due to an ambiguity which arises when the three currents become very close to each other.

Is there any way of restoring the Ward identity, by changing the definition of Γ, adding a purely local terms? Since the anomaly is quadratic in momenta, this term must be linear. Furthermore, if it is a pseudo-tensor (transforms appropriately under Lorentz transformations and parity) and is crossing symmetric under interchange of the vector currents, it must be

$$\Gamma_{\mu\nu\lambda} \to \Gamma_{\mu\nu\lambda} + C\,\epsilon_{\mu\nu\lambda\alpha}(p_1-p_2)^\alpha$$

Now, by choosing C appropriately, we can restore the axial vector Ward identity, but only at the cost of spoiling the vector Ward identity. This just reexpresses our old observation that we cannot choose a_μ so that all the Ward identities are satisfied simultaneously.

We might also wish to consider the object

$$-i\,\tilde{\Gamma}_{\mu\nu\lambda} = \langle 0|\, T[J_{5\mu} J_{5\nu} J_{5\lambda}]\,|0\rangle$$

The graph is the same as before, but now the ambiguity is removed by demanding complete crossing symmetry. This is satisfied if we choose $a = \frac{2}{3}(p_1-p_2)$, so that

$$8\pi^2 p_1{}^\nu \tilde{\Gamma}_{\mu\nu\lambda} = -\tfrac{4}{3}\epsilon_{\lambda\mu\alpha\beta}\,p_2{}^\alpha p_1{}^\beta = \tfrac{4}{3}\epsilon_{\lambda\mu\alpha\beta}\,p_2{}^\alpha p_3{}^\beta$$

$$8\pi^2 p_2{}^\lambda \tilde{\Gamma}_{\mu\nu\lambda} = +\tfrac{4}{3}\epsilon_{\mu\nu\alpha\beta}\,p_1{}^\alpha p_2{}^\beta = \tfrac{4}{3}\epsilon_{\mu\nu\alpha\beta}\,p_3{}^\alpha p_1{}^\beta$$

$$8\pi^2 p_3{}^\mu \tilde{\Gamma}_{\mu\nu\lambda} = +\tfrac{2}{3}\epsilon_{\nu\lambda\alpha\beta}(p_1-p_2)^\alpha(p_1+p_2)^\beta = \tfrac{4}{3}\epsilon_{\nu\lambda\alpha\beta}\,p_1{}^\alpha p_3{}^\beta$$

From the form of the anomaly equation, it is possible to deduce much about the form of Γ or $\tilde{\Gamma}$. It is convenient to consider the symmetric point

$$p_1^2 = p_2^2 = p_3^2 \equiv p^2,$$

since at this point p^2 is the only invariant. Furthermore, since $J_{5\mu}$ is an axial vector, Γ and $\tilde{\Gamma}$ must transform under parity as a pseudo tensor — it must involve an ϵ tensor. And since crossing doesn't interchange any invariants at the symmetric point, just the tensor part of Γ $\tilde{\Gamma}$ must have appropriate crossing properties.

Let's construct the possible tensor structures for $\tilde{\Gamma}$ — pseudo-tensors with complete crossing symmetry. There is no such tensor linear in momentum (i.e. with one index of the ϵ tensor contracted because e.g.

$\epsilon_{\mu\nu\lambda\alpha} p_1^\alpha$ is antisymmetric under $p_2 \leftrightarrow p_3$, $\mu \leftrightarrow \lambda$

there is no term with three indices of ϵ contracted with momenta, because there are only two independent momenta. We are left with

$$\tilde{\Gamma}_{\mu\nu\lambda} = f(p^2)\left[p_{3\mu}\epsilon_{\nu\lambda\alpha\beta}p_1^\alpha p_2^\beta + p_{1\nu}\epsilon_{\lambda\mu\alpha\beta}p_2^\alpha p_3^\beta + p_{2\lambda}\epsilon_{\mu\nu\alpha\beta}p_3^\alpha p_1^\beta\right]$$

as the most general form for $\tilde{\Gamma}_{\mu\nu\lambda}$

Taking the divergence, we have

$$p_3^\mu \tilde{\Gamma}_{\mu\nu\lambda} = p^2 f(p^2)\,\epsilon_{\nu\lambda\alpha\beta}p_1^\alpha p_2^\beta = \frac{1}{6\pi^2}\epsilon_{\nu\lambda\alpha\beta}p_1^\alpha p_2^\beta$$

at the symmetric point $\Rightarrow$ $p^2 f(p^2) = \dfrac{1}{6\pi^2}$

So we conclude

$$\boxed{\tilde{\Gamma}_{\mu\nu\lambda} = \frac{1}{6\pi^2}\left[\frac{p_{3\mu}}{p^2}\epsilon_{\nu\lambda\alpha\beta}p_1^\alpha p_2^\beta + \frac{p_{1\nu}}{p^2}\epsilon_{\lambda\mu\alpha\beta}p_2^\alpha p_3^\beta + \frac{p_{2\lambda}}{p^2}\epsilon_{\mu\nu\alpha\beta}p_3^\alpha p_1^\beta\right]}$$

at the symmetric point.

We see that $\tilde{\Gamma}$ has a singularity at $p^2 = 0$. This is no surprise, since there are massless fermions running around the triangle graph. But the form of the singularity is a bit of a surprise; it is a pole instead of a cut. This is a peculiarity of the kinematics of massless fermions. The remarkable thing is that just the anomaly eqn, and crossing symmetry, determine $\tilde{\Gamma}_{\mu\nu\lambda}$ for all values of p^2.

Having seen that the anomaly implies a singularity at $p^2=0$, one wonders what happens if the fermions are massive.
If $\mathcal{L} = \bar\psi(i\partial\!\!\!/ - m)\psi$ then $J_{5\mu} = \bar\psi\gamma_5\gamma_\mu\psi$

$$\Rightarrow \partial^\mu J_{5\mu} = -2m\bar\psi i\gamma_5\psi$$

and the anomalous ward identity becomes

$$p_3^\mu \tilde{\Gamma}_{\mu\nu\lambda} = -2m\langle 0|T(\bar\psi i\gamma_5\psi)J_{5\nu}J_{5\lambda}|0\rangle + \frac{1}{6\pi^2}\epsilon_{\nu\lambda\alpha\beta}p_1^\alpha p_2^\beta$$

Since a surviving anomaly term for $p\to 0$ would imply a singularity at $p=0$, the first term must cancel the second in the limit $p \ll m$.

<u>External Gauge Fields</u>

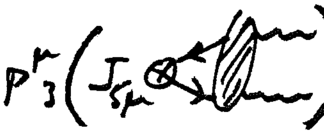
The ward identity anomaly in free field theory does not imply that $J_{5\mu}$ has an anomalous <u>divergence</u>; it is still a conserved current, but the naive manipulations used to derive Ward identities which are consequences of current conservation are not always valid.

Let us now consider the effect of coupling the vector current to an external gauge field.

$$\mathcal{L} = \bar\psi i(\partial\!\!\!/ - ieA\!\!\!/)\psi$$

Now we <u>must</u> choose to have the vector Ward identities satisfied, for otherwise gauge invariance is lost.

Suppose we now consider the divergence of $J_{5\mu}$ in the presence of this external gauge field. We see that $\partial^\mu J_{5\mu}$ has a matrix element between the vacuum and two

$$-ip_3^\mu\Big(J_{5\mu}$$  $$\Big) = \langle 0|\partial^\mu J_{5\mu}|p_1,\epsilon_1;\,p_2,\epsilon_2\rangle$$

$$= -p_3^\mu\,\Gamma_{\mu\nu\lambda}\,(ie)^2\,\epsilon_1^\nu\,\epsilon_2^\lambda$$

(Because only the triangle graph contributes a surface term)

$$= \frac{e^2}{2\pi^2}\,\epsilon_{\nu\lambda\alpha\beta}\,p_1^\alpha p_2^\beta\,\epsilon_1^\nu\epsilon_2^\lambda$$

Let us compare this with the matrix element (in free field theory) of $\epsilon_{\nu\lambda\alpha\beta}F^{\alpha\nu}F^{\beta\lambda} = 4\epsilon_{\nu\lambda\alpha\beta}\partial^\alpha A^\nu\partial^\beta A^\lambda$

$$\langle 0|4\,\epsilon_{\nu\lambda\alpha\beta}\partial^\alpha A^\nu\partial^\beta A^\lambda|p_1,\epsilon_1;\,p_2,\epsilon_2\rangle$$

$$= -8\,\epsilon_{\nu\lambda\alpha\beta}\,p_1^\alpha p_2^\beta\,\epsilon_1^\nu\epsilon_2^\lambda$$

so we find $\displaystyle \partial^\mu J_{5\mu} = \frac{-e^2}{16\pi^2}\,\epsilon_{\nu\lambda\alpha\beta}\,F^{\alpha\nu}F^{\beta\lambda}$

$$= \frac{e^2}{16\pi^2}\,\epsilon_{\nu\lambda\rho\sigma}\,F^{\nu\lambda}F^{\rho\sigma}$$

Is this the correct sign?
[I think I have obtained the conventional sign for the anomolous divergence, but I have defined the axial current with an unconventional sign!] If we now define

$$J_{\mu 5} = \bar\psi\gamma_\mu\gamma_5\psi = -J_{5\mu},$$

we have

$$\boxed{\begin{aligned}\partial^\mu J_{\mu 5} &= \frac{e^2}{16\pi^2}\,\epsilon_{\nu\lambda\rho\sigma}\,F^{\nu\lambda}F^{\rho\sigma}\\ &= \frac{e^2}{8\pi^2}\,F^{\nu\lambda}\tilde F_{\nu\lambda}\end{aligned}}$$

where $\tilde F_{\nu\lambda} = \frac{1}{2}\epsilon_{\nu\lambda\rho\sigma}F^{\rho\sigma}$.

[From the anomolous divergence equation, we can obtain an anomolous ward identity satisfied by

$$\partial^\mu \Lambda_{\mu\nu} = \partial^\mu_x \langle 0| T J_{\mu 5}(x) J_\nu(y) |0\rangle_A$$

$$= \frac{1}{ie} \frac{\delta}{\delta A^\nu(y)} \langle 0| \partial^\mu J_{\mu 5} |0\rangle_A$$

$$= \frac{1}{ie} \frac{e^2}{16\pi^2} \frac{\delta}{\delta A^\nu(y)} 4\epsilon_{\alpha\beta\gamma\delta} \partial^\alpha A^\beta(x) \partial^\gamma A^\delta(x)$$

$$= \frac{-ie}{4\pi^2} \epsilon_{\alpha\beta\gamma\delta} \Big[\delta_\nu{}^\beta \partial^\alpha \delta^4(x-y) \partial^\gamma A^\delta(x) + \delta_\nu{}^\delta \partial^\gamma \delta^4(x-y) \partial^\alpha A^\beta(x)\Big]$$

$$\boxed{\partial^\mu \Lambda_{\mu\nu} = \frac{ie}{2\pi^2} \epsilon_{\nu\alpha\gamma\delta} \partial^\alpha \delta^4(x-y) \partial^\gamma A^\delta(x)}$$

this is called the "anomolous commutator" equation, because the RHS can be interpretted as an anomolous ("Schwinger term") contribution to the Ward identity due to the commutator of J_{05} with J_ν.]

Dynamical Gauge Fields

We have considered the anomolous Ward identities of free field theory, and the anomolous divergence equation in the presence of external gauge fields. Now consider dynamical gauge fields; What sort of radiative corrections to the anomoly equations appear?

There are no radiative corrections.

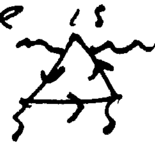
First consider the anomolies of global flavor currents which commute with gauged currents (which may be abelian or nonabelian) We are worried about the effects of graphs such as below. But let's

evaluate anomalies by doing functional integrals in two steps. First we integrate over fermion variables in a fixed external gauge field, generating graphs like.... But the only contribution to the anomaly

in an external field comes from the graph with no external gauge field lines. (All other graphs are insufficiently divergent.) Thus, the anomalies involving the global flavor currents in an external gauge field are <u>independent</u> of the gauge field. When we integrate over the gauge fields, we just get the factor 1. (This argument also indicates ~~that~~ there are no <u>nonperturbative</u> corrections to the anomalies of the global flavor currents.

Now consider the anomalous divergence equation, the anomalies involving global currents and gauge currents.

This is a little more subtle, because the anomalous divergence depends on the gauge field. But the idea is the same, there is no contribution to the anomaly from the



graph shown, nor, therefore, from

if the integral over the internal photon is regulated. The radiative corrections to the two gauge boson matrix element of the current divergence are of the form

But this is just a correction to the matrix element of $F\tilde{F}$.

(To show this in detail takes some work. I do not think it has ever been fully worked out for the case of dynamical nonabelian gauge fields.)

An Operator Derivation

The anomaly is a subtle beast, so I think it is useful to understand its origin in several different ways. One or another way may be the most appropriate in a given context.

Let's next see how the anomaly arises in position space when care is taken in the definition of renormalized current operators. As before, consider free fermions coupled to external gauge fields,

$$\mathcal{L} = \bar{\Psi} i (\not{\partial} - i \not{A}) \Psi . \quad \text{- (coupling constant absorbed in } A_\mu \text{ here)}$$

We will see that there is an ambiguity in the divergence of the axial current $J_{\mu 5} = \bar{\Psi} \gamma_\mu \gamma_5 \Psi$, an ambiguity which can be resolved, as before, by demanding explicit gauge-invariance.

The ambiguity arises because of short distance singularities; these singularities can be controlled by the method of <u>split-point regularization</u>

We thus define the axial current by

$$J_{\mu 5}(x) = \lim_{\epsilon_\mu \to 0} \bar{\Psi}(x + \epsilon/2) \gamma_\mu \gamma_5 \exp\left[i \int_{x-\epsilon/2}^{x+\epsilon/2} dl^\nu A_\nu \right] \Psi(x - \epsilon/2)$$

where an average over the orientation of ϵ_μ is performed before the limit is taken (to ensure covariance), and the (Schwinger) string operator is inserted in the bilocal operator, to ensure gauge-invariance; i.e. we ought to obtain a gauge-invariant, Lorentz-singlet operator when we take the divergence of $J_{\mu 5}$ defined as above.

To calculate $\partial^\mu J_{\mu 5}$, we use the operator equation of motion, derived from the Lagrangian,

$$D_\mu \gamma^\mu \psi \equiv (\partial_\mu - iA_\mu)\gamma^\mu \psi = 0.$$

$$\Rightarrow \gamma^\mu \partial_\mu \psi = iA_\mu \gamma^\mu \psi$$

Thus

$$\partial_\mu \bar\psi \gamma^\mu = -iA_\mu \bar\psi \gamma^\mu$$

$$\partial^\mu J_{\mu 5}(x) = \lim_{\epsilon \to 0} \bar\psi(x+\epsilon/2)\gamma_\mu \gamma_5 \Big[\big(-iA^\mu(x+\epsilon/2) + iA^\mu(x-\epsilon/2)\big) + i\epsilon_\nu \partial^\mu A^\nu(x) + \cdots\Big]\psi(x-\epsilon/2)$$

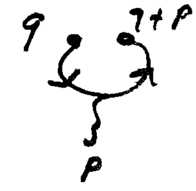
The first term in the brackets is obtain by differentiating ψ and $\bar\psi$, the second by differentiating $\exp[i\int A\cdot dl]$; terms down by more powers of ϵ are dropped. We have

$$\partial^\mu J_{\mu 5}(x) = iF^{\mu\nu} \lim_{\epsilon \to 0} \epsilon_\nu \bar\psi(x+\epsilon/2)\gamma_\mu \gamma_5 \psi(x-\epsilon/2)$$

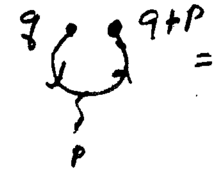
To proceed with the evaluation of the $\epsilon \to 0$ limit, we apply the operator product expansion to the right-hand-side of this equation. Only operators with coefficients as singular as $1/\epsilon$ can survive when the limit is taken. And the only operator of dimension 2 or less with the right symmetries to appear in the OPE of $\bar\psi\psi$ is $F_{\mu\lambda}$.

So we need to find the coefficient of $F_{\mu\lambda}(x)$ in the OPE of $\bar\psi(x+\epsilon/2)\gamma_\mu\gamma_5\psi(x-\epsilon/2)$. It will be easiest to work this out in momentum space, where we can use (momentum space) Feynman rules. So we evaluate

$$\langle 0|\bar\psi(q+p)\gamma_\mu\gamma_5\psi(q)|\gamma, p\rangle \quad = $$



for $|q|$ large



$$= i\epsilon^\lambda \, \mathrm{tr}\, \gamma_\mu\gamma_5 \frac{i(\not q+\not p)}{(q+p)^2}\gamma_\lambda \frac{i\not q}{q^2}$$

and $\mathrm{tr}\,\gamma_\mu\gamma_\alpha\gamma_\lambda\gamma_\beta\gamma_5 = 4i\epsilon_{\mu\alpha\lambda\beta}$

so  $= (i)^3\epsilon^\lambda(-4i)\epsilon_{\mu\alpha\lambda\beta}\frac{p^\alpha q^\beta}{q^4} + \cdots$

$$-4\epsilon_{\mu\beta\alpha\lambda}\frac{q^\beta}{q^4}p^\alpha\epsilon^\lambda = -4i\epsilon_{\mu\beta\alpha\lambda}\frac{q^\beta}{q^4}\langle 0|\partial^\alpha A^\lambda|\gamma,p\rangle$$

So the momentum space coefficient function is given by

$$\bar\psi(q+p)\gamma_\mu\gamma_5\psi(q) = -2i\epsilon_{\mu\beta\alpha\lambda}\frac{q^\beta}{q^4}F^{\alpha\lambda} + \cdots.$$

To Fourier transform back to position space, use

$$\int\frac{d^Dk}{(2\pi)^D}e^{-ik\cdot x}\frac{k_\lambda}{(k^2)^{D/2}} = \frac{2}{(4\pi)^{D/2}\Gamma(\frac{D}{2})}\frac{x_\lambda}{x^2}$$

$$= \frac{x_\lambda}{8\pi^2x^2} \quad \text{for } D=4$$

and we finally have

$$\partial^\mu J_{\mu5}(x) = F^{\mu\nu}\lim_{\epsilon\to 0}\left(\frac{\epsilon_\nu\epsilon^\beta}{4\pi^2\epsilon^2}\right)\epsilon_{\mu\beta\alpha\lambda}F^{\alpha\lambda}$$

Averaging over the orientation of ϵ before taking limit gives $\lim_{\epsilon\to 0}\frac{\epsilon_\nu\epsilon^\beta}{\epsilon^2} = \frac{1}{4}\delta_\nu^\beta$, and we obtain

$$\boxed{\partial^\mu J_{\mu5}(x) = \frac{1}{16\pi^2}\epsilon_{\mu\beta\alpha\lambda}F^{\mu\beta}F^{\alpha\lambda}}$$

, in agreement with our earlier result. (Recall e has been absorbed in A_λ.)

Exercise 3.3

a) In $D=2$ dimensions, define Dirac matrices

$$\gamma_0 = \sigma_x \qquad \gamma_1 = i\sigma_y \qquad \gamma_5 = -\sigma_z$$

so that $\operatorname{tr} \gamma_\mu \gamma_\nu \gamma_5 = 2\epsilon_{\mu\nu}$

Define the current $J_{\mu 5} = \bar{\psi}\gamma_\mu \gamma_5 \psi$ by the gauge-invariant point-splitting method, and evaluate

$$\partial^\mu J_{\mu 5}.$$

b) Find the generalization to $D=2n$ dimensions, n an integer.

Path Integral Derivation (K. Fujikawa, PRL 42, 1195(1979).)

Back in section 1.J, we learned how functional methods could be used to derive Ward identities, given symmetries of the classical action

Suppose the Lagrangian $\mathcal{L}$ is invariant under an infinitesimal linear transformation of fields

$$\delta\mathcal{L} = 0 \qquad \delta\phi = \epsilon A\phi$$

where ϵ is a constant (independent of spacetime position) and A is a constant matrix. To derive Ward identities, consider a change of variable in functional integral of the form

$$\delta\phi(x) = \epsilon(x) A\phi(x)$$

under which

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\mu(\epsilon A\phi) + \frac{\partial\mathcal{L}}{\partial\phi}\epsilon A\phi$$

$$= \partial_\mu\epsilon \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}A\phi + \left[\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}A\partial_\mu\phi + \frac{\partial\mathcal{L}}{\partial\phi}A\phi\right]\epsilon$$

The term in brackets vanishes, since $\delta\mathcal{L}=0$ for $\epsilon=\text{const}$ and we have

$$\delta\mathcal{L} = (\partial_\mu \epsilon) J^\mu \quad \text{where } J^\mu = \frac{\partial \mathcal{L}}{\partial\partial_\mu \phi} A\phi$$

$$\text{and } \delta S = \delta \int d^4x\, \mathcal{L} = -\int d^4x\, \epsilon(x)\, \partial_\mu J^\mu(x)$$

(it is safe to discard surface term; e.g., we can choose ϵ to vanish outside a compact region)

So, performing change of variable

$$e^{iW[J]} = \int (d\phi)\, e^{i[S[\phi]+\int J\phi]}$$
$$= \int (d\phi)\, e^{i[S[\phi]+\int J\phi]}\, e^{i[\int d^4x\, \epsilon(x)(-\partial_\mu J^\mu + JA\phi)]}$$

Now, matching terms linear in ϵ,

$$0 = \int (d\phi)\, e^{i(S[\phi]+\int J\phi)} \int d^4x\, \epsilon(x)\, (-\partial_\mu J^\mu + JA\phi)$$

and, since $\epsilon(x)$ is arbitrary, we have

$$\boxed{\int (d\phi)\, e^{i(S[\phi]+\int J\phi)} [-\partial_\mu J^\mu + JA\phi] = 0}$$

The terms in the expansion of this expression in powers of J are the Ward identities.

E.g.

$$\partial_\mu \langle J^\mu \rangle = 0$$
$$\partial_\mu \langle T J^\mu \phi \rangle = \langle A\phi \rangle$$
etc.

(Derivatives can be pulled outside functional derivative.)

Now, consider the chiral Ward identity

$\mathcal{L} = \bar\psi i(\not\partial - i\not A)\psi$ and apply the above manipulations to

$$\delta\psi = \epsilon i\gamma_5 \psi \qquad e^{iW[\eta,\bar\eta]} = \int d\psi\, d\bar\psi\, e^{i[S[\psi,\bar\psi]+\int(\bar\eta\psi+\bar\psi\eta)]}$$
$$\delta\bar\psi = \epsilon \bar\psi i\gamma_5$$

We find

$$\int (d\psi d\bar\psi)\, e^{i[S[\psi,\bar\psi] + \int(\bar\psi\eta + \bar\eta\psi)]}\left[\partial_\mu J^\mu_5 + \bar\eta i\gamma_5\psi + \bar\psi i\gamma_5\eta\right]$$

where $J^\mu_5 = \bar\psi\gamma^\mu\gamma_5\psi$

Thus $\partial_\mu \langle J^\mu_5 \rangle = 0$, etc.

Where is the anomaly? Apparently, some manipulation performed above is not justified, but what is it?

It is implicit in the above derivation that the path-integral measure is invariant under the $U(1)_A$ chiral transformation, that the Jacobian of the transformation is trivial. It is this assumption which is incorrect.

We will calculate the Jacobian explicitly. First recall what the functional integral

$\int d\psi d\bar\psi$ means.

ψ and $\bar\psi$ are to be treated as anticommuting c-no. functions. If we choose a basis of spinors

$\{\phi_n\}$, $\int d^4x\, \bar\phi_n \phi_m = \delta_{nm}$

we may write

$$\psi(x) = \sum_n a_n \phi_n(x)$$
$$\bar\psi(x) = \sum_n \bar b_n \bar\phi_n(x)$$

where $a_n, \bar b_n$ are anticommuting (Grassman) variables. It is convenient to choose the ϕ_n to be eigenstates of $i\not\partial$ (which is hermitian):

$$i\not\partial\phi_n = \lambda_n\phi_n$$

then the action is $S = \int d^4x\, \bar\psi i\not\partial\psi = \sum_n \lambda_n \bar b_n a_n$

and the functional measure is

$$\prod_n \int d\bar{b}_n\, da_n ,$$

where the integrations are performed according to the rule

$$\int da \begin{Bmatrix} a \\ 1 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} = \int d\bar{b} \begin{Bmatrix} \bar{b} \\ 1 \end{Bmatrix}$$

(and $a^2 = b^2 = 0$, since a, b are anticommuting objects)
Hence, for example

$$\int d\psi\, d\bar{\psi}\; e^{i\int d^4x\, \bar{\psi} i \not{D} \psi} = \prod_n (i\lambda_n) = \det(-\not{D})$$

under the transformation $\delta\psi = \epsilon(x) i\gamma_5 \psi(x)$, we have

$$\sum_n \delta a_n \phi_n = \sum_m \epsilon(x) i\gamma_5 \phi_m a_m$$

$$\Rightarrow \delta a_n = B_{nm} a_m \text{ where } B_{nm} = \int d^4x\, \bar{\phi}_n i\gamma_5 \phi_m \epsilon(x)$$

therefore the Jacobian is

$$\det(\delta_{nm} + B_{nm})^{-1} = \det e^{-B} = e^{-\mathrm{tr}\, B} \quad \text{to order } \epsilon$$

or $$\det(I + B)^{-1} = \exp\left[-i \int d^4x\, \epsilon(x) \sum_n \bar{\phi}_n(x) \gamma_5 \phi_n(x)\right]$$

It is this phase factor which we must evaluate; but it is ambiguous. We must regularize the sum over eigenmodes. To resolve the ambiguity, we again choose a gauge-invariant means of regularization:

$$\int d^4x\, \epsilon(x) \sum_n \bar{\phi}_n \gamma_5 \phi_n = \lim_{M\to\infty} \int d^4x\, \epsilon(x) \sum_n \bar{\phi}_n(x) \gamma_5 \phi_n e^{-\lambda_n^2/M^2}$$

$$= \lim_{M\to\infty} \int d^4x\, \epsilon(x) \sum_n \bar{\phi}_n(x) \gamma_5 e^{-(i\not{D})^2/M^2} \phi_n(x)$$

We need to now evaluate $\sum_n \bar{\phi}_n(x)\gamma_5 f\left[\frac{(i\slashed{D})^2}{M^2}\right]\phi_n(x)$

where f is a smooth regulator,

$f(0)=1 \qquad f(\infty)=f'(\infty)=$

Since, by completeness, $\sum_n \phi_n(x)\bar{\phi}_n(y)=\delta^4(x-y)$, we can choose to evaluate the sum

using any complete basis; it is convenient to use the plane-wave basis (eigenstates of ∂_μ) instead of eigenstates of $i\slashed{D}$,

$$\sum_n \langle\phi_n|x\rangle \gamma_5 f\left[\frac{(i\slashed{D})^2}{M^2}\right]\langle x|\phi_n\rangle = \int \frac{d^4k}{(2\pi)^4}\langle k|x\rangle \mathrm{tr}\left\{\gamma_5 f\left[\frac{(i\slashed{D})^2}{M^2}\right]\right\}\langle x|k\rangle .$$

Furthermore,
$$\slashed{D}^2 = \gamma^\mu\gamma^\nu D_\mu D_\nu = \tfrac{1}{2}\{\gamma^\mu,\gamma^\nu\}D_\mu D_\nu + \tfrac{1}{2}[\gamma^\mu,\gamma^\nu]D_\mu D_\nu$$
$$= D^\mu D_\mu + \tfrac{1}{4}[\gamma^\mu,\gamma^\nu][D_\mu,D_\nu]$$
and $[D_\mu,D_\nu] = -iF_{\mu\nu} \Rightarrow$
$$\slashed{D}^2 = D^\mu D_\mu - \tfrac{i}{2}F_{\mu\nu}\gamma^\mu\gamma^\nu$$

To find leading term in integral as $M^2\to\infty$, expand
$$f(-\slashed{D}^2/M^2) = f(-D^2/M^2) + f'(-D^2/M^2)\left(\tfrac{i}{2}\right)\gamma^\mu\gamma^\nu F_{\mu\nu}\left(\tfrac{1}{M^2}\right)$$
$$+ f''(-D^2/M^2)\left(-\tfrac{1}{8}\right)\gamma^\mu\gamma^\nu F_{\mu\nu}\gamma^\lambda\gamma^\sigma F_{\lambda\sigma}\left(\tfrac{1}{M^2}\right)^2 + \cdots$$

using $\mathrm{tr}\,\gamma_5\gamma^\mu\gamma^\nu = 0 \quad \mathrm{tr}\,\gamma_5\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma = 4i\epsilon^{\mu\nu\lambda\sigma}$, we find
$$\sum_n \bar{\phi}_n(x)\gamma_5 f\left[\frac{(i\slashed{D})^2}{M^2}\right]\phi_n(x) = -\frac{i}{2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}\int\frac{d^4k}{(2\pi)^4}\frac{1}{M^4}f''[(k_\mu-A_\mu)^2/M^2] + \cdots$$

Now, do the integral by shifting the k integral
$$\frac{1}{M^4}\int\frac{d^4k}{(2\pi)^4}f''[(k_\mu-A_\mu)^2/M^2] = \frac{i}{16\pi^2}\int dk^2\,k^2 f''(k^2)$$

and integrating by parts $\int_0^\infty dk^2\,k^2 f''(k^2) = -\int_0^\infty dk^2 f'(k^2) = f(0)-f(\infty) = 1$

So the term which survives as the regulator mass $M\to\infty$ is

$\boxed{\sum_n \bar{\phi}_n(x)\gamma_5\phi_n(x) = \frac{1}{32\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}}$ and the Jacobian is ...

$$\boxed{\det(1+B)^{-1} = \exp\left[-i\int d^4x\,\epsilon(x)\,\frac{1}{32\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}\right]}$$

The same Jacobian factor occurs when we make the change of variable $\delta\bar\Psi(x) = \epsilon(x)\bar\Psi i\gamma_5$. We finally obtain the modified Ward identities

$$\int d\Psi d\bar\Psi\, e^{i[S + S(\bar\Psi\eta + \bar\eta\Psi)]}\Big[\partial_\mu J^\mu{}_5 - \frac{1}{16\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma} + \bar\eta i\gamma_5\Psi + \bar\Psi i\gamma_5\eta\Big],$$

or

$$\boxed{\partial_\mu J^\mu{}_5 = \frac{1}{16\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}}$$

(It is clear that no such anomaly afflicts the vector symmetries, under which Ψ and $\bar\Psi$ are multiplied by opposite phases.)

The need for regularization of the Jacobian is better understood if we further consider the interpretation of the equation

$$\sum_n \bar\phi_n\gamma_5\phi_n = \frac{1}{16\pi^2}F_{\mu\nu}\tilde F^{\mu\nu}\quad , \qquad \text{derived above}$$

Because $(i\not{D})\gamma_5 = -\gamma_5(i\not{D})$, the nonzero eigenvalues of $i\not{D}$ come in pairs of opposite sign:

$$i\not{D}\phi_n = \lambda_n\phi_n \Rightarrow i\not{D}\gamma_5\phi_n = -\lambda_n\gamma_5\phi_n .$$

The eigenstates with zero eigenvalue only can be chosen to be eigenstates of γ_5 also. But, formally, since $\gamma_5\phi_n$ and ϕ_n are orthogonal for $\lambda_n \neq 0$, the contribution due to modes with nonvanishing eigenmodes to

$$\sum_n \int d^4x\, \bar\phi_n\gamma_5\phi_n \qquad \text{vanishes.}$$

So we have $\nu \equiv \frac{1}{16\pi^2}\int d^4x\, F_{\mu\nu}\tilde F^{\mu\nu} = \int d^4x\, \bar\phi_n\gamma_5\phi_n \Big|_{\text{zero mode}}$

Thus $\boxed{n_+ - n_- = \nu \equiv \frac{1}{16\pi^2}\int d^4x\, F_{\mu\nu}\tilde{F}^{\mu\nu}}$ Atiyah-Singer index Theorem

where n_+ and n_- are the no of zeromodes of $i\not{D}$ which are eigenstates of γ_5 with eigenvalues + and − respectively.

This derivation is merely "formal" because the sum in $\sum_n \bar{\phi}_n \gamma_5 \phi_n$ must be carefully defined. But it is clear that any "compactification" of spacetime which makes the spectrum of $i\not{D}$ discrete will suffice.

We need to regularize the sum $\sum_n \bar{\phi}_n \gamma_5 \phi_n$, but it doesn't matter much how we do it. The modes with large eigenvalues tend to give cancelling contributions. Any smooth cutoff should be as good as any other (as we saw).

Heuristic Discussion

We have calculated the anomaly three different ways, but it still seems mysterious. The anomaly tells us (since $F\tilde{F} = 2\vec{E}\cdot\vec{B}$) that chiral fermion pairs can be produced in parallel $\vec{E}$ and $\vec{B}$ fields. Can we understand the physical origin of this effect?

We considered spin-$\frac{1}{2}$ fermions in a constant magnetic field back in section 1.L.
For

$$\vec{B} = B\hat{e}_z \qquad \vec{A} = Bx\hat{e}_y,$$

we can write the "Hamiltonian" as

$$H^2 = (\vec{p} - e\vec{A})^2 + m^2 - ge\vec{B}\cdot\vec{S}$$

$$= p_z^2 + p_x^2 + (p_y - eBx)^2 + m^2 - geBS_z$$

Since p_z, p_y commute with H^2, they are constants of the motion, and

$$p_x^2 + (eB)^2(x-x_0)^2 \text{ has spectrum } (2n+1)eB$$

So

$$\boxed{H^2 = p_z^2 + (2n+1)eB - eB(2S_z) + m^2}$$

where n is an integer and $2S_z = \pm 1$
Notice that, if $m^2 = 0$, there is a zero energy mode with $n=0$ and $2S_z = +1$. The anomalous pair production occurs because parallel electric and magnetic fields can excite the zero energy mode.

(Note also that for $gS_z > 1$ – e.g. vector bosons – there are modes with imaginary energy, which signal an instability. This instability is related to another anomaly, the trace anomaly, about which more will be said later.)

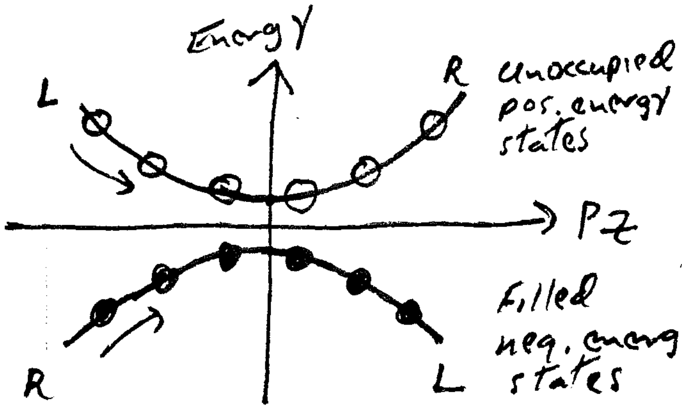
Let us now restrict our attention to the modes with $n=0$ and $S_z = +\frac{1}{2}$, and consider turning on a weak electric field parallel to $\vec{B}$

$$\vec{E} = E\hat{e}_z \quad , \quad A_z = Et \quad (A_0 = 0 \text{ gauge})$$

If E is small, A_z changes slowly and the levels shift adiabatically (n, S_z are unchanged); the energy is

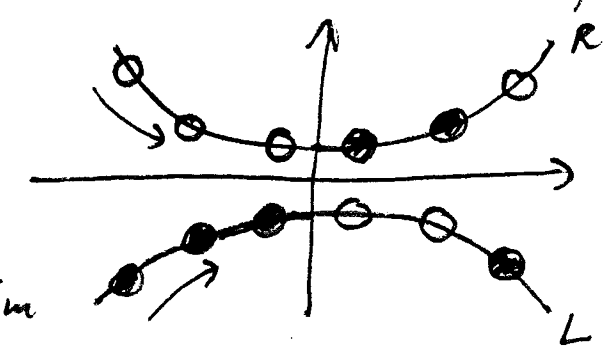
$$H^2 = (p_z - Et)^2 + m^2$$

If we impose a finite volume cutoff, the levels are discrete, and slide along the mass-shell hyperbola like beads on a necklace, as t increases.

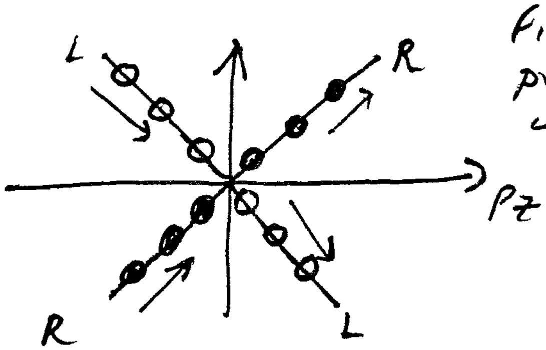


Each mode can be assigned a helicity. All have $S_z = +\frac{1}{2}$, and the mode is moving left or right depending on whether group velocity is − or +.

If $E \gtrsim m^2$, the fermions are able to jump across the $2m$ energy gap, excited by the time varying A_z. The result is the appearance of a chiral pair; fermion and antifermion (i.e. hole) with vacuum quantum numbers — zero momentum and angular momentum.



If $m = 0$, an arbitrarily weak electric field can excite pair production. (The modes do not want to change direction when they cross $p_z = 0$.) The Fermi sea provides an inexhaustible supply of filled states which can be evacuated to produce chiral pairs (RH fermions and LH antifermions)



F. Implications of the Anomaly

π^0 - Decay

The π^0, an isotriplet pseudoscalar, couples to the current

$$J_{\mu A} = \tfrac{1}{2}(\bar{u}\gamma_\mu\gamma_5 u - \bar{d}\gamma_\mu\gamma_5 d)$$

with strength f_π; i.e.

$$\langle 0| J_{\mu A} |\pi\rangle = -i f_\pi P_{\pi\mu}\ , \quad f_\pi = 93\,\text{MeV}$$

We are interested in computing the amplitude $A(\pi^0\to\gamma\gamma)$, so let consider

$$-e^2\langle 0| T[J_{\mu A} J_\nu^{em} J_\lambda^{em}]|0\rangle = -i N_c e^2 \tfrac{1}{2}(Q_u^2 - Q_d^2)\,\Gamma_{\mu\nu\lambda}$$

where $N_c = 3$ comes from summing over all quark colors

This Green's function contains the contribution

which is a pion pole term

$$-i N_c \tfrac{1}{2} e^2 (Q_u^2 - Q_d^2)\,\Gamma_{\mu\nu\lambda}\,\epsilon_1^\nu \epsilon_2^\lambda$$

$$= -i f_\pi P_{3\mu}\,\frac{i}{P_3^2}\, iA(\pi\to\gamma\gamma) + \text{nonpole Terms}$$

Now we may isolate the pole term by taking the divergence and then allowing $P_3 \to 0$. We find

$$A(\pi\to\gamma\gamma) = \frac{-N_c e^2}{2 f_\pi}(Q_u^2 - Q_d^2)\, P_3^\mu \Gamma_{\mu\nu\lambda}\,\epsilon_1^\nu \epsilon_2^\lambda$$

$$\boxed{A(\pi\to\gamma\gamma) = -\frac{N_c e^2}{4\pi^2 f_\pi}(Q_u^2 - Q_d^2)\,\epsilon_{\nu\lambda\alpha\beta}\, p_1^\alpha p_2^\beta\, \epsilon_1^\nu \epsilon_2^\lambda}$$

This is an exact expression in the chiral limit. If not for the anomaly, the amplitude would vanish. We will make the usual extrapolation from $p_3^2 = 0$ to $p_3^2 = m_\pi^2$ to get a current algebra estimate of the decay rate

To calculate the decay rate, we need the spin-summed value of

$$\epsilon_{\nu\lambda\alpha\beta}\, p_1^\alpha p_2^\beta\, \epsilon_1^\nu \epsilon_2^\lambda\; \epsilon_{\rho\sigma\gamma\delta}\, p_1^\gamma p_2^\delta\, \epsilon_1^{*\rho} \epsilon_2^{*\sigma}$$

since $$\sum_r \epsilon_r^\nu \epsilon_r^\rho = -g^{\nu\rho} + \text{gauge terms}$$

we get $$\epsilon_{\nu\lambda\alpha\beta}\,\epsilon^{\nu\lambda}{}_{\gamma\delta}\, p_1^\alpha p_2^\beta p_1^\gamma p_2^\delta$$

$$= -2(\eta_{\alpha\gamma}\eta_{\beta\delta} - \eta_{\alpha\delta}\eta_{\beta\gamma})\, p_1^\alpha p_2^\beta p_1^\gamma p_2^\delta$$

$$= -2\,(p_1^2 p_2^2 - (p_1\cdot p_2)^2)$$

$$= +\tfrac{1}{2}[(p_1+p_2)^2]^2 = +\tfrac{1}{2} m_\pi^2$$

Thus, we have

$$\Gamma = \underbrace{\frac{1}{2}}_{\text{Bose statistics}} \frac{1}{16\pi m_\pi}\Sigma|A|^2 = \frac{m_\pi^3}{64\pi}\frac{1}{f_\pi^2}\left(\frac{\alpha}{\pi}\right)^2 [N_c(Q_u^2 - Q_d^2)]^2$$

Putting in $m_\pi \sim 135\,\text{MeV}$, $f_\pi \sim 93\,\text{MeV}$, $\alpha^{-1} \sim 137$ we find

$$\boxed{\Gamma = 7.6\,\text{eV} \times [N_c(Q_u^2 - Q_d^2)]^2}$$

while experimentally $\Gamma = 7.9\,\text{eV} \pm 7\%$

we conclude that $N_c(Q_u^2 - Q_d^2) \sim 1$

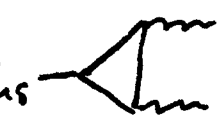
which is the expected value: $3\left(\frac{4}{9} - \frac{1}{9}\right) = 1$

this process has "counted quarks" correctly

(Derived by Schwinger in 1951!)

The U(1) Problem

We have noted several times that the $U(1)_A$ symmetry is not a good symmetry of QCD; there is neither a parity doubled spectrum or a ninth Goldstone boson. Now we can understand why.

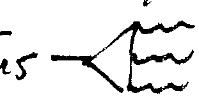
$J_{\mu 5}$ 

The current $J_{\mu 5} = \bar{q}\gamma_\mu \gamma_5 q$ is not conserved because of the anomaly:

$$\partial^\mu J_{\mu 5} = \frac{ng^2}{8\pi^2} \operatorname{tr}(T^a T^b) F^a_{\mu\nu} \tilde{F}^{\mu\nu b}$$

$$= \frac{ng^2}{16\pi^2} F^a_{\mu\nu} \tilde{F}^{\mu\nu a} \qquad (n = \text{no. of flavors})$$

[Note — Gauge invariance apparently requires there to be a square anomaly, with a nonabelian structure.

$J_{\mu 5}$ 

It too arises from a surface term; the square graph becomes linearly divergent when differentiated.

Or, it is more straightforwardly calculated using Fujikawa's path integral method]

However, we saw (Exercise I) that the RHS of the anomalous divergence eqn. is itself a total divergence:

$$\partial^\mu J_{\mu 5} = \frac{ng^2}{4\pi^2} \partial^\mu \epsilon_{\mu\nu\lambda\sigma} \operatorname{tr}\left[A^\nu \partial^\lambda A^\sigma - \tfrac{2}{3} i g A^\nu A^\lambda A^\sigma\right]$$

So there is a conserved (but gauge-variant) current, which has the same commutation relations with fermions as $J_{\mu 5}$. In other words, the RHS of this eqn has no effects in perturbation theory. Can we argue then that the $U(1)_A$ symmetry is spoiled by nonperturbative effects?

Let's derive the Ward identity satisfied by

$$G_\mu(x-y) = \langle 0 | T[J_{\mu 5}(x)\, \mathcal{O}(y)] | 0\rangle|_A$$ where $\mathcal{O}$ is some local operator

(in external gauge field A)

$$\partial_x^\mu G_\mu = \langle 0 | T[D(x)\, \mathcal{O}(y)] | 0\rangle + \delta(x^0-y^0) \langle 0 | [J_{50}(x), \mathcal{O}(y)] | 0\rangle$$

where $D = \partial^\mu J_{\mu 5}$

From this Ward identity, we may derive a low energy theorem

$$\int d^4x\, \partial_x^\mu G_\mu(x-y) = \langle 0 | T\left[\int d^4x\, D(x)\right] \mathcal{O}(y) | 0\rangle + \langle 0 | [Q_5, \mathcal{O}(y)] | 0\rangle$$

where $Q_5 = \int d^3x\, J_{05}$

In momentum space, we have

$$-ip^\mu G_\mu(p)\Big|_{p\to 0} = \langle 0 | T\left[\int d^4x\, D(x)\right] \mathcal{O} | 0\rangle + \langle 0 | [Q_5, \mathcal{O}] | 0\rangle$$

The first term on RHS is a derivative of a total divergence, which is irrelevant in pert. theory. If it is ignored, then we conclude that, if Q_5 does not annihilate the vacuum, then there must be a Goldstone boson pole. (Goldstone's theorem)

But there is another consistent way of satisfying this equation. If no G.B. couples to $J_{\mu 5}$, we have ---

Define chirality of operator $\mathcal{O}$ by

$$[Q_5, \mathcal{O}] = \chi \mathcal{O}$$

(E.g. $\bar{\psi}_L \psi_R$ has $\chi = -2$) Then

$$\boxed{\left[\chi + \int d^4x\, D(x)\right] \langle 0 | \mathcal{O} | 0\rangle = 0}$$ Chirality Selection Rule

The only operators which can have vevs have

$\chi = -\int d^4x\, D(x)$ (in a given external gauge field.)

For what kind of external gauge fields will we have $\int d^4x\, D \neq 0$, and therefore for operators with nonvanishing χ? Lets consider gauge field configurations with finite Euclidean action, i.e.

$$S_E[A] = \int d^4x \left(\tfrac{1}{4} F_{\mu\nu}^a F^{\mu\nu a} \right) < \infty \Rightarrow |F_{\mu\nu}| \xrightarrow[|x|\to\infty]{} \frac{1}{|x|^{2+\epsilon}}$$

Finite action implies A_μ is a "pure gauge" at infinity

$$A_\mu \to \frac{1}{ig} (\partial_\mu \Omega)\Omega^{-1}$$

up to a correction which falls off faster than $1/|x|$

The "pure gauge" long range field can contribute to $\int d^4x\, D$. We may write

$$D = \frac{ng^2}{4\pi^2} \partial^\mu \epsilon_{\mu\nu\lambda\sigma} \mathrm{tr}\left[\tfrac{1}{2} A^\nu F^{\lambda\sigma} + \tfrac{1}{3} ig A^\nu A^\lambda A^\sigma \right]$$

and, therefore, since the first term falls off too rapidly to contribute to the surface integral,

$$-\chi = \int d^4x\, D = \frac{-n}{12\pi^2} \int d^3S^\mu \epsilon_{\mu\nu\lambda\sigma} \mathrm{tr}\left[\partial^\nu\Omega\Omega^{-1} \partial^\lambda\Omega\Omega^{-1} \partial^\sigma\Omega\Omega^{-1} \right]$$

Suppose now that the gauge group is $SU(2)$. Then the integrand in the above surface integral is (up to normalization) just the volume element of the gauge group. We have

$$\chi = \frac{n}{12\pi^2} \int d\theta_1 d\theta_2 d\theta_3\, \epsilon_{ijk} \mathrm{tr}\left[\frac{\partial\Omega}{\partial\theta_i}\Omega^{-1} \frac{\partial\Omega}{\partial\theta_j}\Omega^{-1} \frac{\partial\Omega}{\partial\theta_k}\Omega^{-1} \right]$$

which clearly does not depend on the parametrization $(\theta_1, \theta_2, \theta_3)$ of the three sphere, it is just (in differential form notation)

$$\chi = \frac{n}{12\pi^2} \int \mathrm{tr}\left[d\Omega\Omega^{-1} \wedge d\Omega\Omega^{-1} \wedge d\Omega\Omega^{-1} \right]$$

- the (signed) invariant group volume, up to normalization.

Since $SU(2) = \{ a_0 + i\vec{a}\cdot\vec{\sigma},\ a_0^2 + \vec{a}^2 = 1 \}$ is topologically a three sphere S^3, A_μ at spatial infinity defines a map $\Omega : S^3 \to S^3$

This map has a <u>winding</u> <u>number</u> which specifies how many times the gauge group is wrapped around the sphere at infinity. The winding number is an integer ν, given by

$$\nu = \frac{1}{24\pi^2} \int \mathrm{tr}\left[d\Omega\,\Omega^{-1} \wedge d\Omega\,\Omega^{-1} \wedge d\Omega\,\Omega^{-1} \right]$$

To check the normalization, we calculate ν for the "identity map"

$$\Omega = (x_0 + i\vec{\sigma}\cdot\vec{x}) / (x_0^2 + \vec{x}^2)^{\frac{1}{2}}$$

Since the group measure is invariant under $\Omega \to \Omega\,\Omega_0$ (where Ω_0 is a constant element of the group), it suffices to determine the measure in the vicinity of $\Omega = \mathbb{1}$. If we choose $\theta_i = x_i$, we have

$$\partial_k \Omega\,\Omega^{-1} = i\sigma_k \text{ and } \mathrm{tr}[\partial_k\Omega\,\Omega^{-1}\partial_\ell\Omega\,\Omega^{-1}\partial_m\Omega\,\Omega^{-1}]$$
$$= (i)^3\,\mathrm{tr}\,\sigma_k\sigma_\ell\sigma_m = 2\epsilon_{k\ell m}$$

and $\epsilon_{k\ell m}\,\mathrm{tr}(\quad) = 12$

so $\nu = \frac{1}{2\pi^2}\int d\theta_1 d\theta_2 d\theta_3 = 1$, since $2\pi^2$ is the volume of the sphere S^3

The chiral selection rule becomes

$$\boxed{\begin{array}{l} \chi = 2n\nu \quad \text{where} \\ \nu = \dfrac{g^2}{32\pi^2}\displaystyle\int d^4x\, F^a_{\mu\nu} F^{\mu\nu a} = \text{integer} \end{array}}$$

If the gauge group is $SU(N)$, ν counts the number of times some $SU(2)$ subgroup of $SU(N)$ is wrapped around the sphere at infinity. (A mapping $S^3 \to SU(N)$ can be continuously deformed to a mapping into an $SU(2)$ subgroup)

Gauge field configurations with finite action and $\nu \neq 0$ are called "instantons". We see that, in the presence of instantons, $U(1)_A$ chirality is violated by an amount

$$\chi = 2n\nu$$

if there is no $U(1)_A$ Goldstone boson. Conversely, if the operators which have expectation values in an external gauge field obey the selection rule $\chi = 2n\nu$, then we see from the chiral Ward identity that there is no Goldstone boson which couples to $J_{\mu 5}$

Why must χ depend on n, the number of flavors? It is because the anomaly afflicts only the $U(1)_A$ chirality, not the $SU(n)_L \times SU(n)_R$ chiral symmetry. The operator which has $U(1)_A$ chirality and gets an expectation value must be an $SU(n)_L \times SU(n)_R$ singlet. The lowest dimension operator which meets these requirements is

$$\det_{a,b} \left(\bar{q}_{La}\, q_{Rb} \right) \qquad a,b = 1,2,\ldots,n$$

which has $\chi = 2n$. The most general such operator is a polynomial in $\det(\ )$, with $\chi = 2n\nu$, ν an integer.

The violation of $U(1)_A$ chirality due to instantons is obviously nonperturbative — it requires "large" gauge fields with nonzero winding number, which are distant from the perturbative vacuum, $A_\mu = 0$. We will quantify the extent to which it is nonperturbative below.

More about instantons — the Θ parameter

Let's forget about the quarks for the moment, and consider pure Yang-Mills theory. The quantization of the theory in the $A_0 = 0$ gauge involves a subtlety which we have ignored up to now.

Consider the action on physical states of a gauge transformation $\Omega(\vec{x})$ such that

$$\Omega(\vec{x}) \to \mathbb{1} \quad \text{as} \quad |\vec{x}| \to \infty.$$

The physical states (those obeying the Gauss's Law constraint) in $A_0 = 0$ gauge are invariant under infinitesimal gauge transformations, but not every g.t. which approaches $\mathbb{1}$ at spatial infinity can be built from infinitesimals — i.e. can be continuously deformed to $\Omega = \mathbb{1}$.

Three-dimensional space with spatial infinity identified as a single point is topologically equivalent to the sphere S^3, so the g.t. is a mapping from S^3 into the gauge group

$$\Omega : S^3 \longrightarrow G$$

and if $G = SU(2) \sim S^3$, this is a mapping from S^3 to S^3 which has a winding number. (If $G = SU(N)$, Ω can be deformed so that it maps S^3 to an $SU(2)$ subgroup of $SU(N)$, so it still has a winding number.)
The winding number is an integer, so it cannot change under a continuous deformation of $\Omega(\vec{x})$. If it is nonzero, $\Omega(\vec{x})$ cannot be built up out of infinitesimal gauge transformations, in which case we call it a "big" gauge transformation. If it can be built up out of infinitesimal gauge transformations, we say $\Omega(\vec{x})$ is "small."

The big gauge transformations generate a global Z symmetry of Yang-Mills theory, analogous to the $(Z_N)^3$ symmetry we discovered earlier for gauge fields in a box. An element of this global symmetry group is specified by the winding number p of the big g.t.; we need specify only p because two g.t.'s with the same value of p differ by a small g.t., under which physical states are invariant.

Physical states must furnish irreducible unitary representations of the Z symmetry, which are parametrized by an angle θ.

$$\Omega_p |\psi\rangle = e^{ip\theta} |\psi\rangle$$

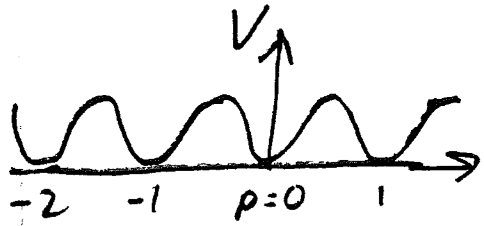
means unitary operator representing Ω_p

Suddenly, a new arbitrary parameter θ has entered Yang-Mills theory. Where did it come from? What is its interpretation?

To better understand the origin of θ, consider the vacuum structure of the theory. The "classical vacuum" in $A_0 = 0$ gauge is the eigenstate of A_i with $A_i = 0$. Of course, any gauge transformation of this state is degenerate with it, but a small gauge transformation of $|A_i = 0\rangle$ is not to be considered a distinct state. However, a large gauge transformation of $|A_i = 0\rangle$ is a distinct state, also degenerate with it. So the classical vacuum is actually infinitely degenerate:

$$\Omega_p |A_i = 0\rangle = |p\rangle$$

are classical vacua, in one-to-one correspondence with the integers.



This classical vacuum structure is analogous to that of a periodic potential in particle mechanics. The "p-vacua" are separated by a potential barrier; i.e., they cannot be deformed continuously into one another inside the vacuum manifold, because Ω_p and Ω_q cannot be continuously deformed into one another if $p \neq q$.

When we carry out semiclassical quantization of the periodic potential problem, we find that the classical vacua are destabilized by quantum mechanical tunneling, and the true eigenstates of the Hamiltonian are Bloch states, which transform as an irreducible representation of the discrete translation symmetry.

Naturally, one wonders whether an analogous phenomenon occurs in Yang-Mills theory. Can a p-vacuum evolve into a q-vacuum $(q \neq p)$ by a quantum-mechanical tunneling event? To answer this question, we attempt to evaluate semiclassically the transition amplitude

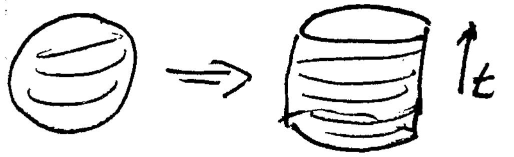
$$\langle q | e^{-HT} | p \rangle = \int (dA)\, e^{-S_{\text{Euclidean}}[A]}$$

where the gauge field obeys the B.C.

$$A_i(t=-T/2) = \frac{1}{ig}(\partial_i \Omega_p)\Omega_p^{-1}$$
$$A_i(t=T/2) = \frac{1}{ig}(\partial_i \Omega_q\ )\Omega_q^{-1}$$

There is a nonzero contribution to this amplitude if there exist finite-action gauge field configurations which interpolate between a p-vacuum configuration at large negative time and a q-vacuum configuration at large positive time.

In fact, such configurations exist, and are precisely the instantons. To see this, deform the "sphere" at infinity in 4-d Euclidean space into a cyllinder with top and bottom faces orthogonal to $\hat{t}$ and walls running along $\hat{t}$



In $A_0=0$ gauge the only contribution to

$$\nu = \frac{1}{24\pi^2}\int d^3S^\mu \epsilon_{\mu\nu\lambda\sigma}\,\mathrm{tr}\left(\partial^\nu\Omega\,\Omega^{-1}\partial^\lambda\Omega\,\Omega^{-1}\partial^\sigma\Omega\,\Omega^{-1}\right)$$

comes from top and bottom of cyllinder, since $\partial_0\Omega=0$, and we see that

$$\nu = q-p$$

(We restricted attention to finite action configurations on page 3.50 because we wanted configurations which could contribute to this tunneling amplitude)

To estimate the tunneling amplitude in the semiclassical approximation, we need to find the leading contribution to it in the $\hbar\to 0$ limit (Being a tunneling amplitude, it vanishes when $\hbar=0$.) For example consider tunneling with $\nu=1$. In the $\hbar\to 0$ limit, we have

$$\langle p+1|e^{-HT}|p\rangle = \int (dA)_{\nu=1}\, e^{-S_E/\hbar}$$

$$\sim (-)\, e^{-S_0/\hbar}$$

where S_0 is the minimum value of the Euclidean action in the gauge field sector with winding number $\nu=1$. The prefactor $(-)$ is generated by integrating over small deviation from the configuration which minimizes S_E. This is just the (Gaussian) steepest-descent approximation to the functional integral, valid in the $\hbar \to 0$ limit.

We can determine S_0 by deriving an inequality satisfied by S_E. In Euclidean space, we have

$$0 \leq (F^a_{\mu\nu} \pm \tilde{F}^a_{\mu\nu})^2 = 2(F^a_{\mu\nu} F^{\mu\nu a} \pm F^a_{\mu\nu} \tilde{F}^{\mu\nu a})$$

(where we have used the identity
$$\tfrac{1}{4} \epsilon_{\mu\nu\lambda\sigma} \epsilon^{\mu\nu}{}_{\rho\xi} = \tfrac{1}{2}(\delta_{\lambda\rho}\delta_{\sigma\xi} - \delta_{\lambda\xi}\delta_{\sigma\rho}))$$

Therefore, $\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} \geq \pm \frac{8\pi^2}{g^2} \frac{g^2}{32\pi^2} F^a_{\mu\nu} \tilde{F}^{\mu\nu a}$

and since $\nu = \frac{g^2}{32\pi^2} \int d^4x\, F^a_{\mu\nu} \tilde{F}^{\mu\nu a}$, we have, by choosing the $\pm$ sign appropriately

$$\boxed{S_E \geq \frac{8\pi^2}{g^2} |\nu|}$$

So the tunneling amplitude is exponentially suppressed for small g^2 (or small $\hbar$) by

$$e^{-8\pi^2/g^2}$$

for $|\nu| = 1$. Tunneling is truly nonperturbative; the amplitude has no power series expansion in g^2

We can saturate the inequality satisfied by S_E by choosing a field configuration with

$$F = \pm \tilde{F}$$

("self dual" or "antiselfdual") Such a configuration is necessarily a solution to the Euclidean eqns of motion; since S_E is a local minimum, it is obviously stationary.

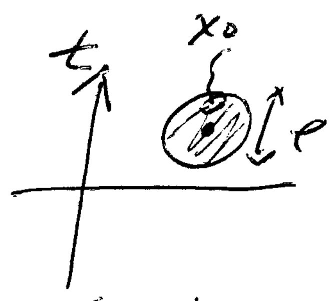
The equation $F = \pm \tilde{F}$ can be solved. I won't describe the solution explicitly, but will mention a few relevant features of the solution.

Consider the $\nu = 1$ case. There is actually a degenerate family of solutions because there are exact symmetries of the classical theory which act nontrivially on the solution. The most general solution is described by eight free parameters (in gauge group $SU(2)$)

center x_0 (4 parameters) due to translation invariance
size ρ (1 parameter) " scale invariance
Group orientation (3 parameters) " gauge invariance

[For $G = SU(N)$, there are
$$N^2-1 - [(N-2)^2-1] - 1 = 4N-5$$
group orientation parameters, or $4N$ parameters all together.
For general ν there are $4N\nu$ parameters, $4N$ for each instanton.]

For $\nu = 1$, we have
$$\mathrm{tr}\, F^2 \sim \left[\frac{\rho}{(x-x_0)^2+\rho^2}\right]^4$$



the action density is localized in a region in Euclidean spacetime with center x_0 and radius of order ρ. It is because the action density is localized in time as well as space that this object is called an "instanton" (or "pseudoparticle")

Returning now to the semiclassical evaluation of the tunneling amplitude, we note that it is necessary to sum over all possible instanton positions and sizes ---

$$\langle p+1 | e^{-HT} | p \rangle = \underbrace{VT}_{} \int \frac{d\rho}{\rho^5} e^{-8\pi^2/g_\mu^2} f(\rho\mu)$$

from integrating over spacetime position

integral over size. ρ^{-5} from dimensional analysis

from integrating over small fluctuations around instanton soln. Has a divergence which is removed by same subtractions as need in vacuum; hence μ dependence

The factor VT appears because there is a constant tunneling probability per unit time and volume.

To avoid large logs in f, we choose $\rho\mu \sim 1$. Then tunneling rate is

$$\Gamma/V \sim \int \frac{d\rho}{\rho^5} e^{-8\pi^2/g_\rho^2} \sim \Lambda^4 \int \frac{d\rho}{\rho} (\rho\Lambda)^{\frac{11}{3}N-4}$$

from 1 loop β func of $SU(N)$ YM theory

The ρ integral diverges for large ρ $(N \geq 2)$. The divergence is a signal that the semiclassical approximation breaks down for large instantons. It also tells us that the tunneling probability is unsuppressed, because the coupling constant runs, and Yang-Mills theory has no small parameter. It is

$$\Gamma/V \sim \Lambda^4$$

as we would guess on dimensional grounds

To construct semiclassical vacuum states which do not suffer from a tunneling instability, we recall the remarks on p 3.54; we define states which transform as an irreducible unitary rep of the global $\mathbb{Z}$ symmetry. (Just as we would construct Bloch states, eigenstates of the discrete translation symmetry, in a one-dimensional solid.)

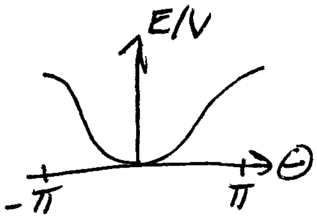
These states are
$$|\theta\rangle = \frac{1}{\sqrt{2\pi}} \sum_p e^{-ip\theta} |p\rangle ,$$
since
$$\Omega_q |\theta\rangle = \frac{1}{\sqrt{2\pi}} \sum_p e^{ip\theta} |p+q\rangle = e^{iq\theta} |\theta\rangle .$$
These states are orthogonal:
$$\langle \theta' | \theta \rangle = \frac{1}{2\pi} \sum_p e^{-ip(\theta-\theta')} = \delta(\theta'-\theta)$$
And they are also stable under semiclassical tunneling:
$$\langle \theta' | e^{-HT} | \theta \rangle = \frac{1}{2\pi} \sum_{pq} e^{ip\theta'} e^{-iq\theta} \langle p | e^{-HT} | q \rangle$$
and
$$\langle p | e^{-HT} | q \rangle = A(p-q)$$ is a function of only $\nu = p-q$
so
$$\langle \theta' | e^{-HT} | \theta \rangle = \frac{1}{2\pi} \sum_p e^{ip(\theta'-\theta)} e^{i(p-q)\theta} A(p-q)$$
$$= \delta(\theta'-\theta) \sum_\nu e^{i\nu\theta} A(\nu)$$

Now θ is a free parameter in Yang-Mills theory; the θ vacua are physically distinguishable states, and each is a suitable vacuum upon which to build a theory. All local physics in diagonal in the θ basis.

The vacuum energy depends on θ. In the leading semiclassical approximation
$$\langle e^{-HT} \rangle_\theta = e^{-E(\theta)T} \sim O(1) - E(\theta)T$$
$$\cong O(1) + VT \int \frac{d\rho}{\rho^5} e^{-8\pi^2/g_\mu^2} f(\mu\rho) \left[e^{i\theta} + e^{-i\theta} \right]$$
from $\nu = \pm 1$ sectors

+ corrections suppressed by more powers of $e^{-8\pi^2/g^2}$

so
$$E(\theta)/V = (\theta\text{-independent}) - 2\cos\theta \int \frac{d\rho}{\rho^5} f(\mu\rho) e^{-8\pi^2/g_\mu^2}$$



In the leading semiclassical approximation, the vacuum energy density is proportional to $-\cos\Theta$ (and of order Λ^4).

We can now explain why we were interested in gauge transformations of the trivial vacuum such that $\Omega \to \mathbb{1}$ as $|\vec{x}| \to \infty$, back on page (3.53). A constant term in Ω is irrelevant in

$$A_i = \frac{1}{ig}(\partial_i \Omega)\Omega^{-1},$$

so the significant thing is that $\Omega \to$ a constant, independent of angles. We are always free to choose $\Omega \to \mathbb{1}$ at large negative time $-T/2$ because we can do a time-independent gauge transformation. But if Ω at $|\vec{x}| = \infty$ depends on angle ϕ at large positive time $T/2$, then

$$A_\phi(t=\tfrac{T}{2}) = \frac{1}{ig}\left(\frac{1}{r}\partial_\phi \Omega\right)\Omega^{-1} \neq 0$$

And the minimal possible value of the contribution to the Euclidean action due to the term

$$\int_{-\frac{T}{2}}^{T/2} dt \int d^3x \, (\dot{A}_\phi)^2 \quad \text{is} \quad \leq \; T\int d^3x \left(\frac{A_\phi(\frac{T}{2})}{T}\right)^2$$

$$\sim \frac{1}{T}\int d^3x \, \frac{1}{|\vec{x}|^2} = \infty$$

In the semiclassical approximation, the trivial vacuum can evolve only to states $|A_i\rangle$ such that $A_i = \frac{1}{ig}(\partial_i \Omega)\Omega^{-1}$ where $\Omega \Rightarrow$ const as $|\vec{x}| \to \infty$.

That is why we can restrict our attention to such states.

That instantons can be important in nonperturbative strong interaction physics was first appreciated by A. M. Polyakov. For a review of instanton physics, see

S. Coleman, Erice Summer School, 1977

Θ as a coupling constant

In a "θ-vacuum", expectation values are computed by

$$\langle \mathcal{O} \rangle_\theta = \frac{\sum_\nu e^{i\theta\nu} \int (dA)_\nu \, e^{-S_E[A]} \, \mathcal{O}}{\sum_\nu e^{i\theta\nu} \int (dA)_\nu \, e^{-S_E[A]}}$$

where $\int (dA)_\nu$ denotes a sum over all gauge fields with winding number ν. But, since

$$\nu = \frac{g^2}{32\pi^2} \int d^4x \, F_{\mu\nu}^a \tilde{F}^{\mu\nu a}$$

we see that it is equivalent to include in the Euclidean action the term

$$S_{E,\theta} = -\frac{i\theta g^2}{32\pi^2} \int d^4x \, F_{\mu\nu}^a \tilde{F}^{\mu\nu a}$$

When we continue to Minkowski space, this becomes a term in the Lagrangian

$$\mathcal{L}_{M,\theta} = \frac{\theta g^2}{32\pi^2} F_{\mu\nu}^a \tilde{F}^{\mu\nu a}$$

Although we will see that including the effects of instantons resolves the $U(1)$ problem, in the process we have created the strong CP problem. The new term in the Lagrangian is CP-odd, so YM theory is CP violating unless $\theta = 0$ or π (Since CP: $\theta \to -\theta$, and θ is an angle, it is unchanged by CP when $\theta = \pi$; i.e., it changes by 2π)

In QCD, we know that θ is quite small, because the strong interactions conserve CP to high accuracy (from the bound on the electric dipole moment of the neutron, we conclude $\theta \lesssim 10^{-9}$) That the "arbitrary" parameter θ chooses to take such an "unnaturally" small value is a puzzle.

The U(1) Problem again — G. 't Hooft, Phys. Rev. Lett. 37, 8 (1976); S. Coleman, Erice, 1977

The vacuum structure of QCD becomes drastically modified when we reintroduce massless quarks.

To begin to see why, consider QCD with n flavors of massive quarks, which has the Lagrangian

$$\mathcal{L} = -\tfrac{1}{4}F^a_{\mu\nu}F^{\mu\nu a} + \bar{q}\, i\not{D} q - \sum_b m_b \left(e^{i\theta_b}\bar{q}_{Lb} q_{Rb} + e^{-i\theta_b}\bar{q}_{Rb} q_{Lb}\right)$$
$$+ \frac{\theta g^2}{32\pi^2} F^a_{\mu\nu}\tilde{F}^{\mu\nu a}$$

Here m_b is chosen to be real and positive; the phases of the quark masses are shown explicitly. The Lagrangian can be put in this form by the manipulations described in Section 1.B.

If we now try to remove the phases from the quark masses by performing a change of variable in the functional integral, we saw on (3.38) ff that the change of variable also changes the parameter θ.

$$\delta q_{R1} = -i\frac{\epsilon}{2} q_{R1} \qquad \delta q_{L1} = i\frac{\epsilon}{2} q_{L1}$$

$$\theta \to \theta - \epsilon$$

The observable parameter, which cannot be changed by such a redefinition is

$$\bar{\theta} \equiv \theta + \sum_b \theta_b = \theta + \operatorname{Arg}\det \mathcal{M}$$

We can push the angle $\bar{\theta}$ into the phase of one or another quark mass, or into the coefficient of $F\tilde{F}$, but we cannot get rid of it.

If there is a massless quark, the phase of its mass is ill-defined, and so is Θ. We can "rotate Θ away" by redefining the massless quark field. Thus, we see that, because of the axial anomaly, the parameter Θ is unobservable if there is a massless quark, and there is no CP violation.

Apparently, the vacuum tunneling is suppressed when there are massless fermions (and so physics becomes Θ-independent). We must understand why this is so. Three explanations are given, each instructive in a different way.

i) Recall the Atiyah-Singer index theorem (p 3.43)
In QCD we have (one flavor)

$$\sum_n \int d^4x\, \bar{\phi}_n \gamma_5 \phi_n = \frac{g^2}{32\pi^2}\int d^4x\, F^a_{\mu\nu}\tilde{F}^{\mu\nu a} = \nu$$

(recall factor of $\frac{1}{2}$ from $\mathrm{tr}(T^aT^b) = \frac{1}{2}\delta^{ab}$) where the ϕ_n's are eigenstates of $i\not{D}$

$$i\not{D}\phi_n = \lambda_n \phi_n$$

Since γ_5 anticommutes with $\not{D}$, $i\not{D}\gamma_5\phi_n = -\lambda_n\gamma_5\phi_n$, and $\gamma_5\phi_n$ is orthogonal to ϕ_n for $\lambda_n \neq 0$. Therefore, only the zero modes contribute to the sum over n above, and, since the zero modes can be chosen to be γ_5 eigenstates, we have

$$\nu = n_+ - n_-$$

where $n_\pm$ is no. of zero modes with $\gamma_5 = \pm 1$. We see that $i\not{D}$ has zero modes whenever ν is nonzero. (If there are n flavors, we have $\nu = n_+ - n_-$ for each flavor.)

The tunneling amplitude, in the presence of massless fermions is...

$$\langle \rho+\nu | e^{-HT} | \rho \rangle = \int (dA)_\nu \int (dq)(d\bar{q})\, e^{-S_E}$$

$$S_E = \tfrac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} - \bar{\psi} i \not{D} \psi$$

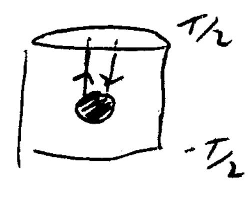
$$= \int (dA)_\nu \det(i\not{D})\, e^{-S_{YM}}$$

But we know that, for $\nu \neq 0$, $i\not{D}$ has a zero mode and therefore $\det(i\not{D}) = 0$, so

$$\langle \rho+\nu | e^{-HT} | \rho \rangle = 0 \text{ for } \nu \neq 0,$$

and the Θ-vacua are degenerate.

ii) From the anomoly equation



$$\partial^\mu J_{\mu 5} = \frac{n g^2}{16\pi^2} F^a_{\mu\nu} \tilde{F}^{\mu\nu a},$$

we see that the change in the axial charge,

$$Q_5 = \int d^3x\, J_{05},$$

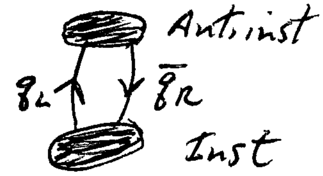
between large neg and pos times, is

$$\Delta Q_5 = \int dt\, \dot{Q}_5 = \int d^4x\, \partial^0 J_{05} = \int d^4x\, \partial^\mu J_{\mu 5} = 2n\nu$$

charge localized in space

So when the gauge field vacuum tunnels, chiral charge is created; the final state contains a chiral pair and is orthogonal to the fermi vacuum.

Only "virtual" tunneling can occur: An antiinstanton is needed to soak up the chirality created by the instanton



iii) Recall that since the divergence of $J_{\mu 5}$ is itself a total divergence, we can construct a conserved current.

$$\partial^\mu J_{\mu 5} = \partial^\mu K_\mu \ , \quad K_\mu = \frac{n g^2}{4\pi^2}\epsilon_{\mu\nu\lambda\sigma}\,\mathrm{tr}\left[A^\nu \partial^\lambda A^\sigma - \tfrac{2}{3} i g A^\nu A^\lambda A^\sigma\right].$$

So $J_{\mu 5} - K_\mu = \tilde{J}_{\mu 5}$ is conserved, but it is not gauge-invariant.

In fact, the conserved charge

$\tilde{Q}_5 = \int d^3x\, \tilde{J}_{05}$ is changed by a big gauge transformation,

$$\Omega_p \tilde{Q}_5 \Omega_p^{-1} = \tilde{Q}_5 - 2np .$$

We know this is so because Q_5 changes by $2np$ when the vacuum tunnels from $|0\rangle$ to $|p\rangle$, and

$\tilde{Q}_5 = Q_5 + Q_5'$ is unchanged

where $Q_5' = \int d^3x K_0 = 2np$ in a $|p\rangle$ vac.

Since $\tilde{Q}_5$ is conserved, it commutes with H, and

$$0 = \langle p | [\tilde{Q}_5, e^{-HT}] | q\rangle$$

Also $\tilde{Q}_5 |p\rangle = \tilde{Q}_5 \Omega_p |0\rangle = (\Omega_p \tilde{Q}_5 + 2np\, \Omega_p)|0\rangle$
$= 2np\, |p\rangle$,

So $$0 = 2n(q-p)\langle p | e^{-HT} | q\rangle$$

Thus $\langle p | e^{-HT} | q\rangle = 0$ for $p \neq q$, and we conclude again that tunneling is suppressed.

The chiral selection rule

Back on page 349 we derived a Ward identity satisfied by the expectation value

$$G_\mu(p) = \text{F.T.} \; \langle J_{\mu 5} \, \mathcal{O}_\chi \rangle_\nu$$

in a background gauge field with winding number ν. We found

$$-ip^\mu G_\mu(p)\big|_{p=0} = (\chi - 2n\nu) \, \langle \mathcal{O}_\chi \rangle_\nu .$$

As we noted then, to show that no Goldstone boson couples to the $U(1)_A$ current, we must verify the selection rule

$$\langle \mathcal{O}_\chi \rangle \neq 0 \quad \text{only if} \quad \chi = 2n\nu .$$

(χ is the chirality of the operator $\mathcal{O}_\chi$.)

We can now show that this selection rule is a consequence of the index theorem, and hence verify that there is no $U(1)_A$ Goldstone boson.

First, to keep the discussion simple, consider the one flavor case, $n=1$, and the operator

$$\mathcal{O} = \bar{q}_R q_L \qquad \text{which has } \chi = 2$$

the expectation value of $\mathcal{O}$ in an external gauge field is

$$\langle \mathcal{O} \rangle_A = \underbrace{N}_{\text{normalization factor}} \int (dq)(d\bar{q}) \, e^{\int d^4x \, \bar{q} i \not{D} q} \, \mathcal{O}$$

We expand

$$q = \sum_n a_n \phi_n \qquad \text{where} \quad i\not{D}\phi_n = \lambda_n \phi_n$$

$$\bar{q} = \sum_n b_n \bar{\phi}_n$$

then

$$\int d^4x \, \bar{q} i \not{D} q = \sum_n b_n a_n \lambda_n$$

and

$$\mathcal{O} = \bar{q}_R q_L = \sum_{m\ell} b_\ell a_m \bar{\phi}_\ell \tfrac{1}{2}(1+\gamma_5)\phi_m$$

we have

$$\langle \bar{q}_R q_L \rangle \propto \int (da)(db) \prod_n (1 + \lambda_n \bar{b}_n a_n) \sum_{m\ell} \bar{b}_\ell a_m \bar{\phi}_\ell \tfrac{1}{2}(1+\gamma_5)\phi_m$$

Here the a's and $\bar{b}$'s are anticommuting variables, and integrals are performed according to the rules

$$\int da \begin{pmatrix} 1 \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \int d\bar{b} \begin{pmatrix} 1 \\ \bar{b} \end{pmatrix}$$

Thus, the only terms which survive when we integrate are those in which each a_n and $\bar{b}_n$ appears exactly once

$$\langle \bar{q}_R q_L \rangle \propto \left(\prod_n \lambda_n \right) \sum_m \frac{1}{\lambda_m} \bar{\phi}_m \tfrac{1}{2}(1+\gamma_5)\phi_m$$

Now consider the following cases:

i) $i\not{D}$ has no zero modes
Since γ_5 and $i\not{D}$ anticommute, all eigenmodes of $i\not{D}$ come in pairs

$$i\not{D}\phi_n = \lambda_n \phi_n \qquad i\not{D}\gamma_5\phi_n = -\lambda_n \phi_n$$

with opposite eigenvalues. These modes give cancelling contributions to the sum over m, which therefore vanishes

ii) one + chirality zero mode, $i\not{D}\phi_0 = 0$; $\gamma_5\phi_0 = \phi_0$
Only the term in the sum with $m=0$ is not killed by the zero mode:

$$\langle \bar{q}_R q_L \rangle \propto \underbrace{\det{}'(i\not{D})}_{\text{product of nonzero eigenvalues}} \bar{\phi}_0 \phi_0 \neq 0$$

iii) one − chirality zero mode $i\not{D}\phi_0 = 0$, $\gamma_5\phi_0 = -\phi_0$
the $m=0$ term is not killed by the zero mode, but

$$\bar{\phi}_0 \tfrac{1}{2}(1+\gamma_5)\phi_0 = 0$$

so $\langle \bar{q}_R q_L \rangle = 0$

iv) More than one zero mode
Now every term in sum is killed by zero modes
$$\langle \bar{q}_R q_L \rangle = 0$$

We conclude that $\langle \bar{q}_R q_L \rangle$ can be nonzero only if
$n_+ = 1$
$n_- = 0$

The index theorem says $\nu = n_+ - n_- = 1$
So we have verified that
$$\langle \bar{q}_R q_L \rangle_{\chi=2} \neq 0 \implies \nu = 1,$$
as required

The generalizations of this argument are obvious
- $\langle \bar{q}_L q_R \rangle \neq 0$ requires $\nu = n_+ - n_- = 0 - 1 = -1$
- $\langle \bar{q}_R q_L \bar{q}_R q_L \rangle \neq 0$ requires $\nu = 2 - 0 = 2$
- $\langle \bar{q}_R q_L \bar{q}_R q_L \bar{q}_L q_R \rangle \neq 0$ requires $\nu = 2 - 1 = 1$

And, in the n flavor case

- $\langle \bar{q}_{L1} q_{R1} \bar{q}_{L2} q_{R2} \cdots \bar{q}_{Ln} q_{Rn} \rangle \neq 0$ $\quad (\chi = -2n)$
requires one − chirality zero mode for each flavor, or $\nu = -1$
It is clear from Fermi statistics that the operator which gets an expectation value is actually
$$\det(\bar{q}_{La} q_{Rb}), \text{ which is an } SU(n)_L \times SU(n)_R \text{ invariant.}$$

Instanton physics breaks $U(1)_A$ symmetry, but not $SU(n) \times SU(n)$ chiral symmetry

We see that the chiral selection rule holds in general, so there is no Goldstone boson coupling to $J_{\mu 5}$.

It is also clear that $U(1)_A$ symmetry is explicitly broken, since there are operators with nonzero $U(1)_A$ charge which acquire vevs.

In the semiclassical approximation, and the one flavor case,

$$\langle \bar{q}_R(x) q_L(x) \rangle_\Theta = \underbrace{\int \frac{d\rho}{\rho^5} \int d^4x_0}_{\substack{\text{integral over}\\ \text{instanton size}\\ \text{and position}}} \frac{\det'(i\not{D})}{\underbrace{\det(i\not{\partial})}_{\substack{\text{denominator}\\ \text{comes from}\\ \text{dividing by}\\ \text{vac amplitude}\\ \text{in } \nu=0 \text{ sector}}}} \bar{\phi}_0(x-x_0)\phi_0(x-x_0) e^{-8\pi^2/g^2} e^{i\Theta} f(\rho\mu)$$

Now $\int d^4x_0 \, \bar{\phi}_0 \phi_0 = 1$ (Normalization of zero mode).

So

$$\langle \bar{q}_R q_L \rangle_\Theta = e^{i\Theta} \int \frac{d\rho}{\rho^5} \frac{\det'(i\not{D})}{\det(i\not{\partial})} e^{-8\pi^2/g_\mu^2} f(\rho\mu).$$

The $e^{i\Theta}$ is expected; rotating Θ is equivalent to redefining the phase of $\bar{q}_R q_L$. The dimensions are ok because $\det'(i\not{D})/\det(i\not{\partial})$ has dimensions of ρ — one less eigenvalue appears in $\det'$, because the zero mode is removed.

We note that the $U(1)_A$ symmetry breaking, like spontaneous $SU(n)_L \times SU(n)_R$ symmetry breaking is <u>soft</u>. The zero mode has the form

$$\phi_0 \sim \frac{\rho}{[\rho^2 + (x-x_0)^2]^{3/2}} \, u \leftarrow \text{spinor};$$

it is localized within radius ρ of the instanton center.

If we calculate

$$\langle \bar{q}_R(x)\, q_L(0)\rangle \propto \int d^4x_0\, \bar{\phi}_0(x-x_0)\, \phi_0(-x_0)$$

$$= \int d^4p\, d^4k\, d^4x_0\, e^{ip(x-x_0)} e^{ikx_0}\, \bar{\phi}_0(p)\, \phi_0(k)$$

$$= \int d^4p\;\; \bar{\phi}_0(p)\, \phi_0(p)\, e^{ipx}$$

or $\quad F.T.\, \langle \bar{q}_R(x)\, q_L(0)\rangle \sim \bar{\phi}_0(p)\, \phi_0(p)$

$\phi_0(p)$ cuts off exponentially $\quad \phi_0(p) \sim e^{-p\rho}$

$$F.T.\, \langle \bar{q}_R\, q_L\rangle \sim e^{i\theta} \int d\rho\, (\rho\Lambda)^{(\frac{11}{3}N - \underbrace{\tfrac{2}{3}}_{\text{one flavor}})} e^{-2p\rho}$$

$$\sim e^{i\theta} \Lambda^{-1} \left(\frac{\Lambda}{p}\right)^{\frac{11}{3}N + \frac{1}{3}}$$

Very rapid fall off at large momentum transfer.

The symmetry breaking is very soft because only small instantons, $\rho \ll \Lambda^{-1}$, are important at large momentum transfer, $p \gg \Lambda$. And small instantons are suppressed by the nonperturbative factor $\exp[-8\pi^2/g_\rho^2]$.

Remarks:

- Chiral $SU(n)_L \times SU(n)_R$ symmetry:

  We saw that instanton effects do not spoil the $SU(n)_L \times SU(n)_R$ symmetry. We have seen no evidence of spontaneous breakdown of this symmetry.

  Nor did we expect to. Instanton calculations are semiclassical calculations which reveal behavior in the $\hbar \to 0$ or $g^2 \to 0$ limit. Since spontaneous breaking of chiral symmetry can occur only at sufficiently strong coupling, we do not expect to see evidence of it in the $\hbar \to 0$ limit.

While the chiral selection rule does not <u>require</u> breaking of $SU(n)_L \times SU(n)_R$ symmetry, one worries about its <u>consistency</u> with spontaneous chiral symmetry breaking:

θ is surely a periodic variable with period 2π, but the chiral selection rule implies

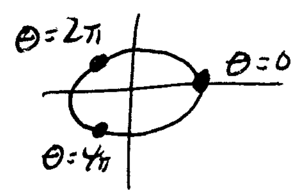
$\langle \mathcal{O}_\chi \rangle_\theta \sim e^{i(\chi/2n)\theta}$ if there are n flavors

But then $\langle \bar{q} q \rangle_\theta \sim e^{i\theta/n}$ has period $2\pi n$ instead of 2π

Is $\langle \bar{q} q \rangle \neq 0$, i.e., spontaneous chiral symmetry breaking, thus inconsistent with the chiral selection rule?

NO! It is true that $\langle \bar{q}_{La} q_{Rb} \rangle_\theta \sim \Delta e^{i\theta/n} U$

where U is a unitary matrix of determinant one.
But $e^{i2\pi/n} U$ also has det 1.
It is just a chiral $SU(n)_L \times SU(n)_R$ rotation of U



Physics really is periodic in θ with period 2π because a $U(1)_A$ rotation by 2π is an $SU(n)_L \times SU(n)_R$ rotation, which is a good symmetry.

- θ-vacua and n-vacua:
  In pure Yang-Mills theory, and QCD with nonzero quark masses, we have seen why the θ-vacua are physically preferable to the n-vacua — they are stable under quantum mechanical tunneling

  But in QCD with massless quarks, tunneling is suppressed. So is there any reason to prefer the θ vacua to the n-vacua?

Yes, there is a reason, already implicit in the discussion on p 3.65.

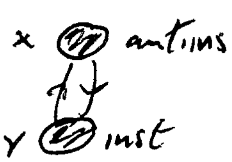
The chiral selection rule implies that

$$\langle p|\, \bar q_R q_L \,|p\rangle = \langle p|\, \bar q_L q_R \,|p\rangle = 0$$

But it does not require

$$\langle p|\, \bar q_R q_L(x)\, \bar q_L q_R(y) \,|p\rangle$$

to vanish, even for $|x-y|$ large.



In fact there is a contribution due to a widely separated instanton and anti-instanton:

$$0 \neq \lim_{|x-y|\to\infty} \langle p|\, \bar q_R q_L(x)\, \bar q_L q_R(y) \,|p\rangle$$
$$\neq \langle p|\, \bar q_L q_R \,|p\rangle \langle p|\, \bar q_R q_L \,|p\rangle$$

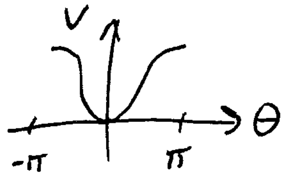
The cluster decomposition <u>fails</u> because of "virtual" tunneling. But the θ-vacua do not suffer from this defect.

- What happened to the Goldstone boson?

The resolution of the $U(1)_A$ problem is rather odd. We have discovered degenerate vacua, labeled by θ, but there is no Goldstone boson.

The reason is that θ is a <u>parameter</u> in the theory, not a dynamical variable. The Goldstone boson would correspond to a long wavelength fluctuation in θ, but this not allowed; θ is frozen in place. There are degenerate vacua, and it is possible to rotate the vacuum without changing the vacuum, but we can only perform the rotation globally, not locally. (This is called "seizing of the vacuum.")

It is possible, by a trick known as the Peccei-Quinn mechanism, to promote θ from the status of a coupling constant to that of a dynamical variable. Then there is a Goldstone boson, known as the "axion".



If there are nonzero quark masses, physics depends on θ, and the axion becomes a massive "pseudo-Goldstone" boson. θ is able to relax dynamically to the value $\theta \sim 0$ which minimizes the vacuum energy, and the CP conservation of the strong interactions is naturally explained.

<u>One final note:</u>
We saw (p. 3.66) that it is possible to define a current

$$\tilde{J}_{\mu 5} = J_{\mu 5} - K_\mu$$

which is conserved, but not gauge-invariant. Since operators with nonzero $\tilde{Q}_5 = \int d^3x\, \tilde{J}_{05}$ get vevs, the Ward identity satisfied by $\tilde{J}_{\mu 5}$ implies that there <u>is</u> a Goldstone boson which couples to $\tilde{J}_{\mu 5}$!

This Goldstone boson is spurious — it appears only in Green's functions which are not gauge-invariant. We know this; we have already shown that there is no Goldstone boson in the gauge-invariant Green's functions with insertions of $J_{\mu 5}$.