K. Other Anomalies

Trace Anomaly

Aside from the triangle anomaly, are there other anomalies which afflict four-dimensional gauge theories?

Another anomaly is the trace anomaly. It comes as no real surprise, because it simply reflects the breaking of scale invariance in a quantum field theory which is scale-invariant at the classical level.

Consider pure Yang-Mills theory, with Lagrangian

$$\mathcal{L} = -\tfrac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} \quad , \quad F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + C^{abc} A^b_\mu A^c_\nu$$

"Classical scale invariance" means that the action is unchanged by the scale transformation

$$A^a_\mu(x) \to \lambda A^a_\mu(\lambda x)$$

under which $\mathcal{L}(x) \to \lambda^4 \mathcal{L}(\lambda x)$

and $S = \int d^4x\, \mathcal{L}(x) \to \int d^4x'\, \mathcal{L}(x') = S$

Classical scale invariance also formally implies that the energy-momentum tensor is traceless. Actually, this statement has to be formulated carefully, because the energy momentum tensor is not uniquely defined, even in the classical theory. We can add a total derivative to $\Theta_{\mu\nu}$ which does not change the integrated charges P_μ but does change the trace of $\Theta_{\mu\nu}$.

A good way to specify $\Theta_{\mu\nu}$ is to consider coupling the theory to gravity. The action of Yang-Mills theory in a curved background metric (which we will treat here as a classical background field) must be a scalar under general coordinate transformations; it is

$$S = \int d^4x \sqrt{g} \left[-\frac{1}{4} F_{\mu\nu}^a F_{\lambda\sigma}^a g^{\mu\lambda} g^{\nu\sigma} \right].$$

Now, we define the energy-momentum tensor $\Theta_{\mu\nu}$ by considering the infinitesimal variation

$$g^{\mu\nu}(x) \to g^{\mu\nu}(x) + \delta g^{\mu\nu}(x)$$

under which

$$\delta S = \frac{1}{2} \int d^4x \sqrt{g}\; \Theta_{\mu\nu} \delta g^{\mu\nu}(x)$$

This is a good definition of $\Theta_{\mu\nu}$ because it gaurantees that $\Theta_{\mu\nu}$ is covariantly conserved

To apply this definition in Yang-Mills theory, note that

$g = -\det g_{\mu\nu}$, and

$$\ln \det(M + \delta M) = \mathrm{tr} \ln(M+\delta M) = \mathrm{tr} \ln M + \mathrm{tr} \ln(1 + M^{-1}\delta M)$$
$$= \mathrm{tr} \ln M + \mathrm{tr}\, M^{-1}\delta M$$

$$\frac{1}{g}\delta g = g_{\mu\nu}\delta g^{\mu\nu}$$

$$\text{or} \quad \sqrt{g} \to \sqrt{g}\left(1 + \tfrac{1}{2} g_{\mu\nu}\delta g^{\mu\nu}\right)$$

thus ...

$$S \to \int d^4x \sqrt{g}\left(1 + \tfrac{1}{2} g_{\mu\nu}\delta g^{\mu\nu}\right)\left[-\tfrac{1}{4} F_{\mu\nu}^a F_{\lambda\sigma}^a (g^{\mu\lambda} + \delta g^{\mu\lambda})(g^{\nu\sigma} + \delta g^{\nu\sigma})\right]$$

or $$S \to S + \frac{1}{2}\int d^4x \sqrt{g}\left(-F_{\mu\rho}^{\ a} F_{\nu}^{\ \rho a} - \frac{1}{4} g_{\mu\nu} F^{\lambda\sigma a} F_{\lambda\sigma}^{a}\right) \delta g^{\mu\nu}$$

and therefore

$$\boxed{\Theta_{\mu\nu} = F_{\mu\rho}^{\ a} F^{\rho}_{\ \nu}{}^{a} - \frac{1}{4} g_{\mu\nu} F^{\lambda\sigma a} F_{\lambda\sigma}^{\ a}}$$

The absence of any scale of length in the theory is reflected by invariance of the action under a conformal transformation of the metric

$$g^{\mu\nu}(x) \to \lambda^2 g^{\mu\nu}(x)$$

or an infinitesimal transformation such that $\delta g^{\mu\nu} \propto g^{\mu\nu}$ then $\delta S = 0 \Rightarrow$

$$\int d^4x \sqrt{g}\; \Theta^{\mu}_{\ \mu}(x) = 0 \quad \text{or} \quad \boxed{\Theta^{\mu}_{\ \mu} = 0.}$$

Classical scale variance implies that the energy-momentum tensor is traceless.

The equation $0 = \Theta^{\mu}_{\ \mu}$ is a trivial identity in the classical theory. But, since we know that scale invariance is broken by quantum effects, one suspects that there will be corrections to this equation in higher order.

A statement can be proved about the form of the "trace anomaly" to all orders of perturbation theory. The proof requires a careful definition of $\Theta_{\mu\nu}$ to all orders. Rather than go into all the subtleties, we will just give a heuristic derivation. (For the careful treatment, see S. Adler, J. Collins, A. Duncan, Phys. Rev. D15, 1712 (1977).)

Let's rescale the gauge field, so the action is

$$S = \int d^4x \sqrt{g}\left[-\frac{1}{4g_0^2} F_{\mu\nu}^a F^{\mu\nu a} + \text{gauge fix} + \text{ghost}\right]$$

(Sorry about the two different g's.) Suppose the theory is defined by introducing a gauge-invariant cutoff (e.g., a lattice cutoff), so the bare coupling g_0^2 is a function of the cutoff Λ. The action is independent of scale only if we rescale Λ along with the metric.

The action is invariant under

$$\eta^{\mu\nu} \to \lambda^2 \eta^{\mu\nu}$$

$$g^2_{0,\Lambda} \to g^2_{0,\Lambda/\lambda}$$

or--

$$\delta\eta^{\mu\nu} = 2\delta\lambda\, \eta^{\mu\nu}$$

$$\delta \frac{1}{g_0^2} = \frac{-2}{g_0^3}(-\Lambda\delta\lambda)\frac{d}{d\Lambda} g_0$$

Thus

$$0 = \delta S = \delta\lambda\left[\int d^4x \left(\Theta^\mu{}_\mu\left(\frac{1}{g_0^2}\right) - \frac{2}{4g_0^3}\left(\Lambda\frac{d}{d\Lambda}g_0\right) F_{\mu\nu}^a F^{\mu\nu a}\right)\right]$$

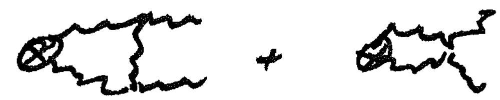
or

$$(\Theta^\mu{}_\mu)_{\text{bare}} = \frac{2}{g_0}\beta(g_0)\frac{1}{4}(F_{\mu\nu}^a F^{\mu\nu a})_{\text{bare}}$$

It is plausible - we won't prove it - that a similar relation holds for properly defined renormalized operators

$$\boxed{\Theta^\mu{}_\mu = \frac{2}{g}\beta(g)\,\frac{1}{4} F_{\mu\nu}^a F^{\mu\nu a}}$$

In terms of graphs, the leading (order g^2) contribution to the trace anomaly comes from



Note that $\Theta^\mu{}_\mu = 0$ at a fixed point of the renormalization group, $\beta = 0$. At a fixed point, the quantum theory is scale-invariant.

What do we learn from the trace anomaly? It is sometimes said that the trace anomaly implies that gluons must "condense", i.e. that F^2 acquires a vacuum expectation value.

To see the connection, let's construct a crude phenomenological Lagrangian for the 0^+ "glueball", i.e., the state which couples to the operator F^2. We will write

$$F^a_{\mu\nu} F^{\mu\nu a} = h^4$$

so h is a conventionally normalized scalar field, and write down an effective Lagrangian for the self interactions of the field h; i.e. the glueball. This approach is crude because we neglect the coupling of h to other channels, and neglect nonlocal effects in the self-interaction:

$$\mathcal{L} = \tfrac{1}{2} \partial_\mu h \partial^\mu h - V(h)$$

We want this effective theory to correctly take into account the effect of the trace anomaly, $\Theta^\mu{}_\mu \propto h$.

That is, under a scale transformation, we want

$$\delta\mathcal{L} = -ch^4,$$

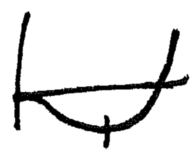
where c is a numerical constant.

Under $h(x) \to \lambda h(\lambda x)$, we have

$$S = \int d^4x\, \mathcal{L}(x) \to \int d^4x'\, \lambda^{-4}\left[\lambda^4 \partial_\mu h \partial^\mu h - V(\lambda h)\right]$$

so $$\delta\mathcal{L} = 4V - h\frac{dV}{dh} = -ch^4$$

or $$\boxed{V = 4c\, h^4 \ln(h/\Lambda)}$$



This potential is minimized at $4h^3\ln(h/\Lambda) + h^3 = 0$ or $\ln h/\Lambda = \frac{1}{4}$

$h = \Lambda e^{1/4}$

(Note: in abelian theory $c<0$, and potential not bounded below; apparently, phenomenological theory makes no sense.)

What is the significance of "gluon condensation"? $\langle F^2\rangle \neq 0$ breaks no symmetry (except scale-invariance) so it is not so significant as the condensation of quark pairs which breaks chiral symmetry spontaneously.

Gluon condensation means that the 0^+ "glueball" mixes strongly with the vacuum, so its properties are hard to calculate. And, actually, instanton physics already provided a dynamical basis for $\langle F^2\rangle \neq 0$.

If we introduce quarks into the theory, then we must also rescale the bare quark mass under a scale transformation to leave physics unchanged. We obtain

$$\Theta^{\mu}{}_{\mu} = \frac{2\beta(g)}{g}\frac{1}{4}F_{\mu\nu}^{a}F^{\mu\nu a} + (1+\gamma_m)\, m\, \bar{q}q$$

where $\gamma_m m = \mu \frac{d}{d\mu} m$

<u>Canonical approach to trace identity</u>

The connection between $\Theta^{\mu}{}_{\mu} = 0$ and scale invariance can also be derived by canonical methods. We can show that $\Theta^{\mu}{}_{\mu}$ is the divergence of a symmetry current associated scale-invariance

Consider a scalar field theory with Lagrange density $\mathcal{L}(\phi, \partial_\mu \phi)$. A scale transformation is

$$\phi(x) \to \lambda\, \phi(\lambda x)$$

and an infinitesimal scale transformation is

$$\delta\phi = (1 + x_\lambda \partial^\lambda)\phi$$

under which the change in the Lagrangian is

$$\delta \mathcal{L}_d = (d + x_\lambda \partial^\lambda)\mathcal{L}_d$$ where $\mathcal{L}_d$ is the term in $\mathcal{L}$ with dimension d

$$= \partial^\lambda (x_\lambda \mathcal{L}_d) + (d-4)\mathcal{L}_d$$

If the Lagrangian contains dimension-four terms only, it changes by a total derivative.

According to Noether's theorem, we have

$$\partial^\mu\left[\frac{\partial \mathcal{L}}{\partial \partial^\mu \phi}\,\delta\phi\right] = \delta\mathcal{L}$$

or $$\partial^\mu\left[\frac{\partial \mathcal{L}}{\partial \partial^\mu \phi}\,(x_\lambda \partial^\lambda \phi + \phi)\right] = \partial^\mu (x_\mu \mathcal{L}) + \sum_d (d-4)\,\mathcal{L}_d$$

And the canonical energy momentum tensor is

$$\Theta_{\mu\nu} = \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi}\,\partial_\nu \phi - \eta_{\mu\nu}\mathcal{L}$$

So we have

$$\partial^\mu D_\mu = \sum_d (d-4)\,\mathcal{L}_d \quad \text{where } D_\mu = x^\nu \Theta_{\mu\nu} + \pi_\mu \phi$$

D_μ is the scale current, or "dilatation" current

In the quantum theory, we have the freedom to do a redefinition of $\Theta_{\mu\nu}$ which preserving the properties

$\Theta_{\mu\nu}$ is symmetric
conserved

and such that the change in Θ is a total derivative; then the integrated charges P_μ, $M_{\mu\nu}$ which are the generator of the Poincare algebra, are unchanged.

The redefinition

$$\Theta_{\mu\nu} \to \Theta_{\mu\nu} + \alpha(\partial_\mu \partial_\nu - \eta_{\mu\nu}\partial^2)\phi^2$$

has the desired properties. But this refinition changes the relation between $\Theta_{\mu\nu}$ and the scale current D_μ. For a particular choice of α we will have

$$D_\mu = x^\nu \Theta_{\mu\nu}, \text{ and then}$$

$$\partial^\mu D_\mu = \eta^{\mu\nu}\Theta_{\mu\nu} + x^\nu \partial^\mu \Theta_{\mu\nu} = \Theta^\mu{}_\mu .$$

So conservation of D is equivalent to tracelessness of Θ

The "field virial" $\Pi_\mu \phi$ can be write $\phi \partial_\mu \phi = \partial_\mu(\frac{1}{2}\phi^2)$ in a theory with a canonical kinetic term.

We want to define $\tilde{\Theta}_{\mu\nu} = \Theta_{\mu\nu} + f_{\mu\nu}$ where f is a conserved, symmetric, total derivative, and

$$\partial^\mu x^\nu f_{\mu\nu} = \partial^\mu \partial_\mu (\tfrac{1}{2}\phi^2)$$
$$\text{or} \quad f^\mu{}_\mu = \tfrac{1}{2}\partial^2 \phi^2$$

and if $f_{\mu\nu} = \alpha(\partial_\mu \partial_\nu - \eta_{\mu\nu}\partial^2)\phi^2$, we have $-3\alpha = \frac{1}{2} \Rightarrow \alpha = -\frac{1}{6}$. So the "new improved" energy, momentum Tensor in the classical theory is

$$\tilde{\Theta}_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \eta_{\mu\nu} \mathcal{L} - \tfrac{1}{6}(\partial_\mu \partial_\nu - \eta_{\mu\nu} \quad)\phi^2$$

and we have verified that $\tilde{\Theta}^\mu{}_\mu = 0$ follows from the classical eqns of motion.

In a gauge theory, the "field virial" term is trivial, so we do not need to do a redefinition of the classical energy momentum tensor.

see ---

Coleman, Jackiw, Ann. Phys. 67:552 (1971)

Callan, Coleman, Jackiw, Ann. Phys. 59:42 (1970)

The SU(2) Anomaly

Consider a gauge theory with gauge group

$$G = SU(2)$$

and a single two-component fermion which is an SU(2) doublet

$$\psi_{i\alpha}$$

color index $i=1,2$ — spinor index $\alpha=1,2$

This remarkably simple theory has some puzzling features.

What happens when the interaction gets strong? We expect a pairing instability, but in what channel? The attractive channel is

$$2 \times 2 \supset 1,$$

but the SU(2)-singlet Lorentz-singlet bilinear does not exist, because of Fermi statistics:

$\epsilon^{ij}\epsilon^{\alpha\beta}\psi_{i\alpha}\psi_{j\beta}$ is symmetric under interchange

Condensates which are spin-1 or Lorentz triplet are allowed, but it is not plausible that condensation occurs in these channels.

The same oddity occurs if...

- there are an odd no. of doublets, $2n+1$. Fermi statistics will allow $2n$ fermions to pair up, but one is left over, massless
- the gauge group is $Sp(2n)$, with odd no. of fermions in defining rep. This theory is anomaly free.

and a gauge singlet is an antisymmetric combination of two $(2n)$'s

$$\underbrace{\eta^{ij}}_{\text{antisym.}} \psi_{i\alpha} \psi_{j\beta}$$

A further odd property: consider zero modes in this theory in an instanton background. What effective operator is generated by the fermion integration?

We have
$$\mathcal{L}_{\substack{\text{ferm.}\\ \text{Eucl.}}} = \underset{\substack{\uparrow\\ RH}}{\bar{\chi}} \not{D}\left(\frac{1+\gamma_5}{2}\right) \underset{\substack{\uparrow\\ LH}}{\psi} = \bar{\chi} \not{D}_L \psi$$

The operator $\not{D}_L$ takes LH spinors to RH spinors, so it is not antihermitian. That is, the zero modes are chiral in an instanton background

$$\not{D} \psi_0 = 0 \quad (\text{instanton})$$
$$\bar{\chi} \overleftarrow{\not{D}} = 0 \quad (\text{anti-instanton})$$

E.g., in a one-instanton background,

$$\psi = a_0 \psi_0 + \sum_{n\neq 0} a_n \psi_n$$
$$\chi = \sum_{n\neq 0} \bar{b}_n \bar{\chi}_n \qquad \int \bar{\chi}_m \psi_n = \delta_{mn}$$

Therefore
$$\int d\bar{\chi}\, d\psi\, e^{\int \bar{\chi} \not{D}_L \psi}\, \psi(x)$$
$$= \int da\, d\bar{b} \prod_{n\neq 0} (1 + \lambda_n \bar{b}_n a_n)(a_0 \psi_0 + \sum_{m\neq 0} a_m \psi_m)$$
$$= \det{}'(\not{D}_L)\, \psi_0(x)$$

Note that we don't have the doubling of the modes that we had in the vector-like theory, since $\gamma_5 \psi = \psi$. So $\det' D\!\!\!/_L$ is the square root of $\det' D\!\!\!/$ in the vector-like theory.

We see that, in the one instanton background, it is $\psi_0(x)$, which is not Lorentz invariant or gauge invariant, which acquires an expectation value. But we have not yet averaged over the orientation of the instanton in space and SU(2). When we do, we get a rotational invariant and an SU(2) invariant. So

$$\langle \psi_0(x) \rangle = 0$$

One instanton doesn't generate an effective operator at all.

Neither can two instantons, since there is no bilinear which is gauge invariant and Lorentz-invariant allowed by Fermi statistics

Four overlapping instantons can give a vev to

$$\psi_i^\alpha \psi_j^\beta \psi_k^\gamma \psi_l^\delta$$

which is totally antisymmetric (because of Fermi statistics) in the 4 indices (i, α) and therefore a Lorentz/gauge singlet. This operator does explicitly break the U(1) Fermion no. which has an anomaly. What is peculiar is that it takes four instantons to do the job.

the properties cited above are odd, but not really paradoxical. Still, one is starting to get suspicious. (Note that the remarks about instantons also apply to the Sp(2n) model.)

In fact, we will argue that the SU(2) theory with one left-handed doublet is ill-defined; it just doesn't exist. It is inconsistent because of a subtlety concerning gauge-invariance in the presence of fermions. One way of expressing the inconsistency is to note that the fermion determinant is not gauge-invariant in this theory. This is a new kind of anomaly which spoils gauge invariance. It is nonperturbative, i.e., unlike the triangle anomaly, it cannot be uncovered by calculation of a Feynman graph.

Actually, the nonabelian triangle anomaly, which spoils gauge invariance unless anomalies cancel, can also be thought of as a failure of the fermion determinant to be gauge-invariant. We'll return to this interpretation of the nonabelian triangle anomaly after explaining the SU(2) anomaly.

Gauge invariance and quantum mechanics together can place severe restrictions on physical theories, if we demand consistency. As a warm up, let's recall a classical example: the quantum mechanics of a charged particle interacting with a magnetic monopole.

Magnetic Monopole:

The equation of motion is

$$m\ddot{\vec{r}} = e\dot{\vec{r}} \times \vec{B}, \quad \vec{B} = \frac{g\hat{r}}{r^2}$$

which is perfectly well-defined and gauge-invariant. But, in order to formulate quantum mechanics, we construct an action. The Lagrangian is

$$L = \tfrac{1}{2} m\dot{\vec{r}}^2 + e\dot{\vec{r}}\cdot\vec{A}, \text{ where } \vec{\nabla}\times\vec{A} = \vec{B}$$

and the action is $S = S_{kin} + S_{int}$,

$$S_{int} = e\int_1^2 dt\, \frac{d\vec{r}}{dt}\cdot\vec{A} = e\int_1^2 d\vec{r}\cdot\vec{A}$$

(depends on path, but not on vel. along path)

S_{int} has a troubling feature; really two, related, troubling features.
First of all, it is not gauge-invariant:

$$\vec{A} \to \vec{A} + \frac{1}{e}\vec{\nabla}\Lambda \Rightarrow S_{int} \to S_{int} + \Lambda_2 - \Lambda_1$$

– it depends on the value of the gauge function at the endpoints. Of course, this is no problem in the classical theory – the equations of motion are still gauge-invariant – but it is potential trouble in the quantum theory

Secondly, it is multivalued. Consider two paths from 1 to 2. The difference

$$(S_{int})_P - (S_{int})_{P'} = e\oint_{P-P'} d\vec{r}\cdot\vec{A} = e\int_{S_{P,P'}} d^2\vec{S}\cdot\vec{B} = e\Phi$$

where Φ is the flux through the patch bounded by the two paths. But now suppose we continuously deform an initial path, allowing it to wind around a sphere enclosing the monopole and return to its starting point. Then the change in S_{int} is determined by the flux through the entire sphere

$$\Delta S_{int} = 4\pi e g$$

So a given path actually has action

$$S_{int} = (S_{int})_0 + 4\pi e g n$$

where n can be any integer.

In quantum mechanics, we assign the phase e^{iS} to every trajectory, and sum over trajectories. For QM to be consistent, it is enough that this phase be well-defined; so we demand

$$4\pi e g = 2\pi n \quad \text{or} \quad \boxed{g = \frac{n}{2e}}$$

QM is consistent only if the magnetic charge of the monopole is an integer multiple of $1/2e$. (Dirac, 1931)

We can also think of this quantization condition as being required by gauge invariance.

Overall phases make no difference in quantum mechanics; only relative phases matter. So it is not necessarily a cause for concern that $S_{1\to 2}$ is not gauge-invariant, as long as the relative phases of two paths beginning at 1 and ending at 2 do not depend on the gauge. We need only worry

in other words, about ambiguities in the phase associated with a closed path. This phase is

$$\exp[ie \oint \vec{A}\cdot d\vec{r}\,]$$


U
L

If we try to define a vector potential everywhere on a closed sphere, it must be singular; it can be nonsingular only if the magnetic flux through the sphere is zero.


U
L

But there is no problem defining a vector potential on a disk. So divide the sphere into an upper and lower part, divided by the path in question. Each part (U and L) is equipped with a nonsingular vector potential ($\vec{A}_U$ and $\vec{A}_L$)

These two vector potentials must define the same physics along the closed path, i.e. each assigns the same phase to the path

$$\exp[ie \oint \vec{A}_U \cdot d\vec{r}\,] = \exp[ie \oint \vec{A}_L \cdot d\vec{r}\,]$$

$$\text{or} \quad \exp[ie \oint (\vec{A}_U - \vec{A}_L)\cdot d\vec{r}\,] = 1 = e^{ie(\Phi_U + \Phi_L)} = e^{i 4\pi e g}$$

This means that

$$\Omega(\vec{r}_f) = \exp ie \int_{\vec{r}_i}^{\vec{r}_f} (\vec{A}_U - \vec{A}_L)\cdot d\vec{r}$$

defines a single-valued gauge transformation on the closed path, and, in fact

$$\vec{A}_U = \vec{A}_L + \underbrace{\frac{1}{ie}(\vec{\nabla}\Omega)\Omega^{-1}}_{\text{along path}}$$

i.e. $\vec{A}_U$ and $\vec{A}_L$ differ by a gauge transformation on the path.

... where n is the winding no. of this

We see now that the Dirac quantization condition is related to the topology of the gauge group. Magnetic charge is quantized because the winding number is an integer, or

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

1st homotopy group ≡ fundamental group.

3-dimension Yang-Mills theory:

Gauge invariance requires quantization of a parameter in 3-dimensional Yang-Mills theory also (i.e. 2+1-dimensional spacetime). Oddly, the quantized parameter is a "mass" for the gauge field.

The most general renormalizable Lagrangian for (2+1)-dimensional Yang-Mills theory is

$$\mathcal{L} = -\frac{1}{2g^2} \operatorname{tr} F_{\mu\nu}F^{\mu\nu} - \frac{\mu}{2g^2} \epsilon^{\mu\nu\lambda} \operatorname{tr}\left(F_{\mu\nu}A_\lambda + \tfrac{2}{3} i g A_\mu A_\nu A_\lambda\right)$$

the (new) second term is gauge-invariant only in a restricted sense.

Cf. our 4-dimensional result (p. 3.50 and p. 3.66)

$$\partial^\mu J_{\mu 5} = \frac{1}{8\pi^2} \partial_\mu \epsilon^{\mu\nu\lambda\sigma} \operatorname{tr}\left(F_{\nu\lambda} A_\sigma + \tfrac{2}{3} i g A_\nu A_\lambda A_\sigma \right) = \partial_\mu K^\mu$$

and we saw that $\int d^3x\, K^0$ was not invariant under a time-independent gauge transformation, but changed by $2p$ under a gauge transformation with winding number p.

Here, under a gauge transformation which goes to a constant at $|x| \to \infty$ in (2+1)-dimensional space-time, we have a change in the action

$$S \rightarrow S + \frac{8\pi^2\mu}{g^2} p$$

this ambiguity will not affect the quantum theory as long as the phase e^{iS} associated with a trajectory is unambiguous. We must require

$$\frac{8\pi^2\mu}{g^2} = 2\pi n \quad \text{or} \quad \boxed{\mu = \frac{g^2}{4\pi} n}$$

the "mass" is quantized.

The quantization condition again has a topological basis. Gauge transformations in (2+1)-dimension spacetime are classified by

$$\pi_3(G) = \mathbb{Z} \qquad \text{(for simple gauge group } G\text{)}$$

as the gauge transformations in time [(0+1)-dim. spacetime] along the charged particle trajectory in the monopole background are classified by $\pi_1(G)$. Quantization is required for the consistency of quantum mechanics with gauge invariance.

The SU(2) Anomaly Again

The SU(2) anomaly is another example of a consistency requirement which arises when we wed quantum mechanics and gauge invariance.

What causes trouble this time is the fermion determinant which arises when we perform the fermionic integration. The determinant is tricky because its sign is not uniquely defined.

More on the SU(2) Anomaly

Consider again, slightly more carefully, the eigenvalue problem discussed on p (3.128)
We have

$$\mathcal{L}_{\substack{ferm\\ Eucl}} = \bar{\chi}\, i\not{D}\left(\frac{1+\gamma_5}{2}\right)\psi = \bar{\chi}\, i\not{D}_L \psi$$

Formally, the fermion integration is

$$\int d\bar{\chi}\, d\psi \; \exp - \int \bar{\chi}\, i\not{D}_L \psi = \det(i\not{D}_L) = (\det i\not{D})^{\frac{1}{2}}$$

But how is $\det(i\not{D}_L)$ to be defined?

Defining it is slightly subtle because the operator $i\not{D}_L$ does not preserve the space on which it acts; it take left-handed fermions to right handed fermions —

$$i\not{D}_L : \mathcal{H}_L \to \mathcal{H}_R$$

There is no problem defining the modulus of the determinant; only the phases of the eigenvalues are ambiguous. We can choose a basis so that

$$i\not{D}_L : \underbrace{|L,a\rangle}_{\text{orthonormal basis}} \to \underbrace{|R,a\rangle}_{\text{orthogonal basis}} ,$$

and define moduli of "eigenvalues" of $i\not{D}_L$ by

$$\lambda_a = \sqrt{\langle R,a | R,a\rangle}$$

and the modulus of the determinant by

$$\left|\det i\not{D}\left(\frac{1+\gamma_5}{2}\right)\right| = \prod_a \lambda_a = (\det i\not{D})^{\frac{1}{2}}$$

this definition gives the square root of the det of $i\not{D}$, in which eigenvalues appear as the pairs $\pm \lambda_a$.

Of course, we need not be concerned about an ambiguity in the overall phase of the effective action functional

$$e^{-W[A]} = \int d\bar{\chi} d\psi\, e^{-\int \bar{\chi} i\not{D}_L \psi} = \text{"} \det i\not{D}_L \text{"}.$$

The overall phase is unobservable. Only relative phase are relevant in quantum mechanics, and were it not for the need for regularization, relative phases would be unambiguous (and gauge invariant). The phase of $i\not{D}_L$ can be defined only if $i\not{D}_L$ has no zero modes, and is invertible, and the operator

$$\not{D}_L^{-1}[A'] \not{D}_L[A]$$

takes H_L to H_L, so the phase of its det is well defined, and it is the relative phase

$$\operatorname{Im}(\ln \det \not{D}_L[A] - \ln \det \not{D}_L[A'])$$

But, in general, the need for regularization can cause relative phases also to be ambiguous. This is how (nonabelian) anomalies arise.

That we need fear an ambiguity in only the phase of the fermion det, and not its modulus, can be understood in another way, which is also enlightening.

If the fermions are in a complex representation R, then, since the Dirac action for these fermions is the complex conjugate of the action for fermions in the conjugate representation $\bar{R}$, we have

$$W_R[A] = W_{\bar{R}}[A]^*$$

therefore $\qquad W_R + W_{\bar{R}} = 2\,\mathrm{Re}\, W_R = \text{real}$

But $W_R + W_{\bar{R}}$ is the effective action for fermions in the representation $R+\bar{R}$. And a theory with fermions in the representation $R+\bar{R}$ can be regularized by the Pauli-Villars method, without breaking gauge invariance. This is because

$$R \times \bar{R} \supset 1,$$

and it is possible to give the fermions gauge-invariant masses.

Therefore $\mathrm{Re}\, W_R$ cannot be afflicted by an anomaly; it is gauge-invariant. Only the imaginary part of W_R can fail to be gauge-invariant. The modulus of e^{-W} is unambiguous, but its phase can be ambiguous.

A similar remark applies to a theory, like the $SU(2)$ theory, with fermions in a pseudo-real representation, i.e. such that

$$(R \times R)_{\text{antisymmetric}} \supset 1.$$

Fermi statistics prevents a Pauli-Villars regulator in this theory from being gauge-invariant, but, if the fermions are in a rep.

$$R+R,$$

then Pauli-Villars regularization is possible therefore

$$e^{-W_{R+R}} = e^{-2W_R} \quad \text{is gauge-invariant}$$

therefore e^{-W_R} ~~is~~ has not an ambiguous phase, but only an ambiguous sign

this is why the SU(2) anomaly is nonperturbative. The perturbative nonabelian anomaly for complex representations is a failure of gauge invariance under an infinitesimal gauge transformation which can be studied in perturbation theory. The anomaly for pseudoreal rep is a failure of gauge invariance under a finite gauge transformation which changes the sign of the effective action; we cannot detect it by considering only infinitesimal changes of the gauge fields.

The topological basis of the SU(2) anomaly is the statement

$$\pi_4(SU(2)) = Z_2$$

Gauge transformations on (compactified) spacetime fall into two classes which cannot be continuously deformed one into the other

(Compare our statements $\pi_1(U(1)) = Z$ about the magnetic monopole, and $\pi_3(G) = Z$ about 3-dim Yang-Mills theory.) The configurations A and A^Ω (where Ω is a gauge transformation with nontrivial Z_2 winding number) are continuously connected by field configurations of finite action, so there is no way to avoid summing over both in the functional integral

$$Z = \int (dA_\mu)(\det i\not{D})^{\frac{1}{2}} \exp[-\tfrac{1}{2g^2}\int \mathrm{tr} F^2]$$

In the absence of fermions, the double counting is harmless, and cancels out in calculations of expectation values.

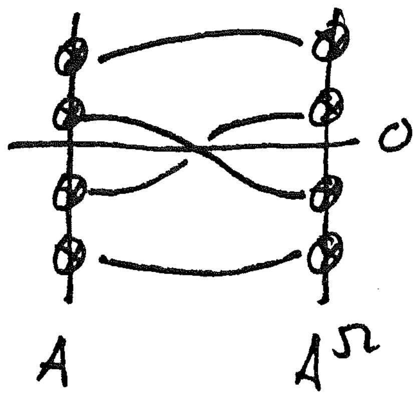
But in the SU(2) model with one left-handed fermion doublet, the double counting causes trouble, because

$$[\det i\not{D}(A)]^{\frac{1}{2}} = -[\det i\not{D}(A^\Omega)]^{\frac{1}{2}}$$

Therefore, a cancellation occurs and $Z=0$; all expectation values are $0/0$ - i.e., are not well-defined.

E.g., suppose we define the $\det^{\frac{1}{2}}$ of $i\not{D}(A)$ to be the product of all positive eigenvalues of $i\not{D}$. Now consider a continuous path in gauge field configuration space from A to A^Ω. Of course, $i\not{D}(A^\Omega)$ has the same eigenvalue spectrum as $i\not{D}(A)$

But what can be shown is that



an odd number of positive and negative eigenvalues change places. Therefore, no matter how we define $[\det i\not{D}(A)]^{1/2}$, $[\det i\not{D}(A^\Omega)]^{\frac{1}{2}}$ has the opposite sign.

This type of eigenvalue flow is required by a "mod 2" version of the Atiyah-Singer index theorem for a five-dimensional Dirac operator. According to this index theorem, the number of zero modes of the 5-dimensional Dirac operator for a gauge field which interpolates (as a function of "time" τ) between A and A^Ω is odd.

The connection between zero modes and eigenvalue flow arises as follows: the 5-dimensional Dirac equation satisfied by a zero mode

$$\not{D}_5 \psi = 0$$

can be written

$$\frac{d}{d\tau}\psi = -\gamma^\tau \not{D}_4 \psi$$

where $\not{D}_4 = [\partial_\mu - iA_\mu(x,\tau)]\gamma^\mu$

is the four-dimensional Dirac operator at time τ

We can choose A_μ to vary slowly with τ, so that the adiabatic approximation applies, and this eqn. has a solution of the form

$$\psi(x,\tau) = f(\tau)\,\phi(x,\tau)$$

where $\quad \gamma^\tau \not{D}_4\, \phi(x,\tau) = \lambda(\tau)\,\phi(x,\tau)$

(i.e. $\phi(\tau)$ is eigenstate of $\not{D}_4(\tau)$)

and f satisfies

$$\frac{d}{d\tau} f(\tau) = -\lambda(\tau) f(\tau) \Rightarrow f(\tau) = f(0)\, e^{-\int_0^\tau d\tau' \lambda(\tau')}$$

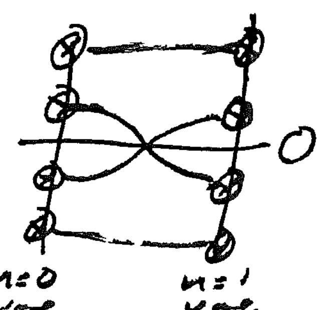
this is a normalizable zero mode only if

$$\lambda(\tau) > 0 \qquad \tau \to +\infty$$
$$\lambda(\tau) < 0 \qquad \tau \to -\infty$$

Hence, every zero mode of $\not{D}_5$ corresponds to a trajectory $\lambda(\tau)$ which passes from negative to positive eigenvalues. The no. of zero modes mod 2 hence determines the change in the sign of the determinant.

Incidentally, the production of chiral fermions due to instanton effects can also be understood as a crossing of zero by eigenvalues of the 3-dimensional Dirac operator, required by zero modes of the 4-dimensional Dirac operator in the presence of the instanton.

An occupied negative energy state becomes a positive energy state, and a fermion pops out of the Dirac sea.

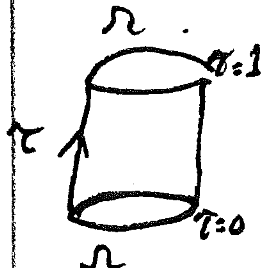


Any theory with $\pi_4(G) \neq 0$ can suffer from this inconsistency. But π_4 is nontrivial only for the $Sp(N)$ groups. $Sp(N)$ with an odd no. of fermions is ill-defined; the same theories to which our remarks on p 3.127 ff apply.

SU(2) Anomaly in the Hamiltonian Formulation

Consider the group $\mathcal{G}$ of time independent gauge transformations (compactified) three dimensional space to the gauge group

$$\mathcal{G} : S^3 \to G$$



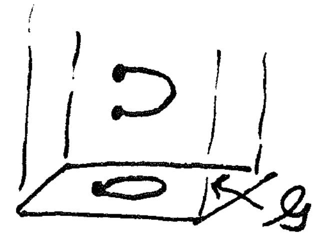
In the Hamiltonian framework, we define physical gauge invariant states by demanding that $\mathcal{G}$ is represented by 1 in the physical subspace of the Hilbert space.

In this framework, the significance of $\pi_4(G) = Z_2$ is that the group $\mathcal{G}$ is not simply connected; there are "loops" in $\mathcal{G}$

$$\Omega(\vec{x},\tau) \quad \text{such that} \quad \Omega(\vec{x},0) = \Omega(\vec{x},1)$$

which cannot be deformed to a point in $\mathcal{G}$.

As is familiar from the representation theory of finite dimensional Lie groups, since $\mathcal{G}$ is not simply connected, exponentiation of a rep. of the Lie algebra of $\mathcal{G}$ (the infinitesimal gauge transformations) need not furnish a single-valued rep of $\mathcal{G}$; but defines a rep of $\tilde{\mathcal{G}}$, the simply-connected covering group of $\mathcal{G}$.



Even if there is no perturbative anomaly, so that it is possible to represent the algebra of gauge transformations locally, there may be a problem globally, if $\pi_1(\mathcal{G}) = Z_2$. In particular, the theory does not exist in the Hamiltonian formulation if there are no physical states which are $\mathcal{G}$ singlets; then the physical subspace is empty. This is surely the case if the identity element in $\mathcal{G}$ is represented by both 1 and -1 in the Hilbert space (center of $\bar{\mathcal{G}}$).

We can demonstrate the inconsistency of the theory by finding a closed path in $\mathcal{G}$ which is represented in Hilbert space by a path which runs from 1 to -1. For if we could solve the Gauss's law constraint

$$Q^a |\psi\rangle = 0$$

then we would have to have $|\psi\rangle$ left invariant by all elements of $\bar{\mathcal{G}}$. But (-1) cannot leave any state invariant. So the Gauss's law constraint has no solutions. The physical subspace is empty.

To find this path in $\mathcal{G}$, we make use of the fact that the <u>conserved</u> fermion number $\tilde{Q}$ is not invariant under a topologically nontrivial gauge transformation

$$\Omega_p^{-1} \hat{Q} \Omega_p = \tilde{Q} + np$$

where Ω_p is a time-independent gauge transformation with winding number p, and n is the no. of chiral doublets.

consider

$$\Omega_\Theta = e^{i\tau_3\theta}\Omega_1^{-1}e^{-i\tau_3\theta}\Omega_1 ,$$

where τ_3 is the diagonal SU(2) generator, for $\theta \in [0, 2\pi]$. This is a closed loop in the space of "small" gauge transformations. For $\theta = 2\pi$, $e^{i2\pi\tau_3}$ is ~ represented in Hilbert space by

$$e^{i2\pi\tau_3} \Rightarrow (-1)^{\tilde{Q}},$$

because only the fermions are in half-odd-integer "spin" reps of SU(2), so

$$e^{i2\pi\tau_3}|\text{state}\rangle = (-1)^{\text{Fermion No}}|\text{state}\rangle.$$

And $\tilde{Q}$ is the conserved charge which has the right commutation relations to represent fermion no.

Since Ω_1 acting on a state acts like a fermion creation operator, we have in the Hilbert space

$$\Omega_\theta \Rightarrow (-1)^{\tilde{Q}}\Omega_1^{-1}(-1)^{\tilde{Q}}\Omega_1^{\cdot 1} = (-1)^n$$

if there are n left-handed fermions which are SU(2) doublets. We conclude that the Gauss's law constraint cannot be solved if n is odd; and the theory is simply not gauge-invariant.

More about the Nonabelian Triangle Anomaly

We now understand that the "nonabelian anomaly" arises when the effective action

$$e^{iW[A]} = \int d\bar\psi\, d\psi\, e^{i\int \bar\psi\, i\gamma^\mu(\partial_\mu - iA_\mu)\psi} \quad \text{(Minkowski space)}$$

fails to be gauge invariant, because gauge-invariance is destroyed by the regularization procedure. (Compare the "abelian anomaly" which arises for a different reason, the noninvariance of the path integral measure under a chiral rotation. Note also that the language "abelian anomaly" and "nonabelian anomaly" is not precise. I am using "nonabelian anomaly" to mean an anomaly which destroys gauge invariance, but such an anomaly can occur in a $U(1)_A$ gauge theory, as well as in a theory with a nonabelian gauge group. Let's derive how the change in the effective action under an infinitesimal gauge transformation is related to the divergence of a gauge current.

Gauge transformation

$$A_\mu \to \Omega A_\mu \Omega^{-1} + \frac{1}{i}\partial_\mu \Omega\, \Omega^{-1}$$

(the gauge coupling g has been absorbed in A)

$$\Omega \sim 1 + i\omega$$

$$\Rightarrow A_\mu \to A_\mu + i[\omega, A_\mu] + \partial_\mu \omega$$

$$\Rightarrow \delta A_\mu = [D_\mu, \omega]$$

Suppose that, to leading order in ω, the change in $W[A]$ under a gauge transformation is...

$$W[A+[D,\omega]] = W[A] - \int \mathrm{tr}\, G\omega$$

then we expand W to obtain the Ward identity

$$W[A] + \int \mathrm{tr} \frac{\delta W[A]}{\delta A_\mu}[D_\mu,\omega] = W[A] - \int \mathrm{tr}\, G\omega$$

or $$-\int \mathrm{tr}\left[D_\mu, \frac{\delta W}{\delta A_\mu}\right]\omega = -\int \mathrm{tr}\, G\omega = \delta_\omega W$$

after an integration by parts. Since $\omega(x)$ is an arbitrary function, we have

$$\left[D_\mu, \frac{\delta W}{\delta A_\mu}\right] = G.$$

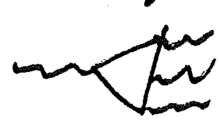
If gauge-invariance were exact, the above would hold with $G=0$; G is the ward-identity anomoly.

Since $\frac{\delta W}{\delta A^a_\mu}$ is the operator $J^{\mu a} = \bar{\psi}\gamma^\mu T^a \psi$

(ψ is a LH two-component fermion) we have the operator equation

$$[D_\mu, J^\mu]^a = G^a.$$

G is the anomalous (covariant) divergence of the gauged current. Back on page 3.75 ff, we considered the form of the term in G which is quadratic in gauge fields. We also noted that there is a higher order term generated by the graph



but we did not compute it. We will now

show that the higher order term can be determined from an algebraic consistency condition which the anomaly must satisfy.

The consistency condition follows directly from the algebraic structure of the infinitesimal gauge transformations:

Under $\Omega = 1 + iu$, we have

$$\delta_u W = \int \mathrm{tr}\, u G = \frac{1}{2}\int u^a G^a$$

W defines a representation of the Lie algebra –

$$\delta_u \delta_v W - \delta_v \delta_u W = \delta_{[u,v]} W$$

– this is the Wess-Zumino consistency condition, which G must satisfy. (Phys. Lett. 37B, 95 (1971).)

To write it out explicitly, the change in G under a gauge transformation is

$$\delta_v G^a(x) = \int dy \, \frac{\delta G^a(x)}{\delta v^b(y)} v^b(y)$$

$$\delta_v \delta_u W = \frac{1}{2}\int dx\,dy\, u^a(x) v^b(y) \frac{\delta G^a(x)}{\delta v^b(y)}$$

$$\delta_u \delta_v W - \delta_v \delta_u W = \frac{1}{2}\int dx\,dy\, u^a(x) v^b(y) \left[\frac{\delta G^b(y)}{\delta u^a(x)} - \frac{\delta G^a(x)}{\delta v^b(y)}\right]$$

$$= \delta_{[u,v]} W = \frac{1}{2}\int dx\, [u,v]^c G^c(x)$$

$$= \frac{1}{2}\int dx\, u^a(x) v^b(x)\, i C^{abc} G^c(x)$$

or $$\boxed{\frac{\delta G^b(y)}{\delta u^a(x)} - \frac{\delta G^a(x)}{\delta v^b(y)} = i\, C^{abc} G^c(x)\, \delta(x-y)}$$

This consistency condition allows us to determine the higher order terms in G once the lowest order term is given.

If the gauge group is simple, the form of the lowest order term in G can be uniquely fixed by demanding exact crossing symmetry, and we computed its normalization on p 3.28. It is

$$[D_\mu, J^\mu]^a = \frac{1}{24\pi^2} \epsilon^{\mu\nu\rho\sigma}\, \mathrm{tr}\left(T^a \partial_\mu A_\nu \partial_\rho A_\sigma\right) + \cdots$$

Exercise 3.8

Use the consistency condition to construct $[D_\mu, J^\mu]^a$, given the first term.