Low energy effective theories and approximate symmetries

We have already encountered the notion of an effective field theory at lower energies obtained by integrating out higher energy processes in the discussion of neutrino masses. This is a general procedure, pioneered by Ken Wilson. The content and interaction strengths of an effective field theory vary according to energy scale. For a given effective low-energy field content, there will be an effective cutoff above which new higher-energy processes begin to take place (e.g. $W_\mu^\pm$ and Z_μ creation above their mass scales $\sim 100\ \mathrm{GeV}/c^2$). Below a given cutoff, interaction strengths vary according to renormalization group flow, which we will come back to later on.

For the strong interactions, there is a natural scale below which an effective field theory description becomes relevant. Owing to asymptotic freedom, the strong interaction coupling decreases as energy scale increases. At energies of the order of the Z_μ boson mass ($\sim 91\ \mathrm{GeV}/c^2$), one finds $\alpha_3(M_Z^2) \equiv \frac{g_3^2}{4\pi} = 0.12$. At lower energies, α_3 increases and actually tends to infinity at energies $\sim \Lambda_{QCD} = 250\ \mathrm{MeV}/c^2$. This doesn't actually mean that couplings diverge. What it does mean is that perturbation theory becomes unreliable for the strong interactions below Λ_{QCD}.

What can one do to describe QCD physics below the Λ_{QCD} scale? Since Λ_{QCD} is about 2% of the electroweak scale of M_W, M_Z (or of the Higgs mass, which we now know to be $\sim 125\ \mathrm{GeV}/c^2$), it is appropriate to integrate out all the electroweak $\mathcal{O}(100\ \mathrm{GeV}/c^2)$ structure. Moreover, quark masses vary significantly according to flavour:

$m_u \sim 2.3\ \mathrm{MeV}/c^2$, $m_d \sim 4.8\ \mathrm{MeV}/c^2$ | $m_s \sim 95\ \mathrm{MeV}/c^2$ "light quarks"

$m_c \sim 1.3\ \mathrm{GeV}/c^2$; $m_b \sim 4.2\ \mathrm{GeV}/c^2$, $m_t \sim 173\ \mathrm{GeV}/c^2$ "heavy quarks"

For energies between m_u, m_d, m_s and Λ_{QCD} an approximation becomes relevant: freeze out all M_W, M_Z, M_H physics and the C, B and T quarks, and treat the "light" u, d and s quarks as <u>massless</u>.

The resulting low(ish) energy theory has a leading-order effective Lagrangian consisting just of the strong interaction gluons and the massless u, d, s quarks

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - i(\bar{u}, \bar{d}, \bar{s}) \begin{pmatrix} \not{D} & & \\ & \not{D} & \\ & & \not{D} \end{pmatrix} \begin{pmatrix} u \\ d \\ s \end{pmatrix}$$

where now the covariant $\not{D} = \gamma^\mu D_\mu = \gamma^\mu(\partial_\mu - \frac{i}{2} g_3 G^a_\mu \lambda^a)$ involves only the SU(3) gauge fields G^a_μ, $a = 1, \dots, 8$.
In this massless-quark limit, an accidental $U_V(3)$ <u>rigid</u> symmetry has clearly appeared, corresponding to flavour rotations of the u, d and s quarks. The $SU_V(3)$ subgroup of this $U_V(3)$ is what Murray Gell-Mann called the <u>eightfold way</u> (alluding to the Noble Eightfold Path of Buddhism). This SU(3) symmetry was independently proposed by Yuval Ne'eman while he was a student of Abdus Salam at Imperial [M. Gell-Mann & Y. Ne'eman, "The Eightfold Way" (Benjamin, 1964)].

Of course, this $U_V(3)$ is broken by the differences in the non-zero quark masses, but it does give a reasonably accurate description of QCD physics. Since $U_V(3) \simeq SU_V(3) \times U_V(1)$, the SU(3) structure and the $U_V(1)$ symmetry can be discussed separately. The $U_V(1)$ subgroup is none other than baryon number symmetry $U_B(1)$. It is the only part of $U_V(3)$ which survives when one includes the rest of the Standard Model.

The reason for the V subscripts is that there is actually more accidental rigid symmetry in $\mathcal{L}_{QCD}$ than just the $U_V(3)$. The Dirac spinor fields u, d and s are <u>reducible</u> under the Spin(3,1) group: they may be separated into L and R chiral projections, so the spinor terms in $\mathcal{L}_{QCD}$ may be separated into $-i(\bar{u}_L \not{D} u_L + \bar{u}_R \not{D} u_R + \bar{d}_L \not{D} d_L + \bar{d}_R \not{D} d_R + \bar{s}_L \not{D} s_L + \bar{s}_R \not{D} s_R)$.

Hence, the full accidental symmetry of $\mathcal{L}_{QCD}$ is actually $U_L(3) \times U_R(3)$. $U_V(3)$ is the diagonal subgroup of this, under which the L and R chiral projections of each light quark are rotated by the same $U_V(3)$ matrix.

How much of this accidental symmetry is actually seen in particle experiments? $U_V(1)$ is baryon number symmetry $U_B(1)$, and is good to a high degree. Ignoring the breakings caused by quark u,d,s masses, the $SU_V(3)$ symmetry organises baryonic bound states of the quarks into $SU_V(3)$ representations. The lightest set of such baryons is the <u>octet</u>, built from three quarks in the 8-dimensional $\begin{smallmatrix}\square\square\\\square\end{smallmatrix}$ representation of $SU_V(3)$. The mixed permutation symmetry $\begin{smallmatrix}\square\square\\\square\end{smallmatrix}$ implies a contraction between the spinor indices of the maximal helicity state: $\epsilon^{ABC} q^{[i}_{A\alpha} q^{j]}_{B\beta} q^{k}_{C\gamma}$ must be antisymmetric in $[\alpha\beta]$ spinor indices, where ϵ^{ABC} contracted with the $SU_c(3)$ strong interaction indices A, B, C produces a colour singlet. So this $SU_V(3)$ approximate multiplet has spin 1/2. A heavier baryon multiplet is the <u>decuplet</u>, in the $\square\square\square$ representation of $SU_V(3)$. The maximal helicity state of this multiplet, formed from $\epsilon^{ABC} q^{(i}_{A\alpha} q^{j}_{B\beta} q^{k)}_{C\gamma}$ must have totally symmetrized $(\alpha\beta\gamma)$ spinor indices, so this multiplet has spin 3/2. $SU_V(3) \times U_B(1)$ agree well with particle data, e.g. the multiplets

i, j: flavour labels 1, 2, 3

Octet: $p^+(uud)$, $n^0(udd)$, $\Lambda^0(uds)$, $\Sigma^+(uus)$, $\Sigma^0(uds)$, $\Sigma^-(dds)$, $\Xi^0(uss)$, $\Xi^-(dss)$

Decuplet: $\Delta^{++}(uuu)$, $\Delta^+(uud)$, $\Delta^0(udd)$, $\Delta^-(ddd)$, $\Sigma^{*+}(uus)$, $\Sigma^{*0}(uds)$, $\Sigma^{*-}(dds)$, $\Xi^{*0}(uss)$, $\Xi^{*-}(dss)$, $\Omega^-(sss)$

What about the remainder of the $U_L(3) \times U_R(3)$ symmetry of $\mathcal{L}_{QCD}$? None of it fits with experimentally observed particle multiplet structure. We shall see that one of the $U_L(3) \times U_R(3)$ generators has an <u>anomaly</u>, and so is removed as a candidate approximate symmetry. Write the infinitesimal action of $U_L(3) \times U_R(3)$ transformations on the u,d,s quarks as

$$\delta \begin{pmatrix} u \\ d \\ s \end{pmatrix} = \left[\frac{i}{2} \omega_L^\theta \lambda_\theta P_L + \frac{i}{2} \omega_R^\theta \lambda_\theta P_R \right] \begin{pmatrix} u \\ d \\ s \end{pmatrix}$$

where λ_θ, $\theta = 0, 1, \dots, 8$ are the Gell-Mann matrices (so SU(3) generators are $\frac{1}{2}\lambda_\theta$, $\theta = 1, \dots, 8$) extended by $\lambda_0 = \sqrt{2/3}\, \mathbb{1}_{3\times 3}$; $\mathrm{tr}(\lambda_\theta \lambda_\varphi) = 2\delta_{\theta\varphi}$.

Then define $\omega_L^\theta = \omega_V^\theta + \omega_A^\theta$, $\omega_R^\theta = \omega_V^\theta - \omega_A^\theta$, so the infinitesimal $U_L(3) \times U_R(3)$ transformation becomes

$$\delta\begin{pmatrix} u \\ d \\ s \end{pmatrix} = \left[\frac{i}{2}\omega_V^\theta \lambda_\theta + \frac{i}{2}\omega_A^\theta \lambda_\theta \gamma_5\right]\begin{pmatrix} u \\ d \\ s \end{pmatrix}.$$

Only the axial ω_A^θ transformations are potentially anomalous. Of them, only one generator is actually anomalous. The anomaly will turn out to involve the $^{(i,j)}$ trace of the axial generators $\lambda_\theta\gamma_5$, and the only such generator with a non-vanishing trace is $\lambda_0 = \sqrt{\frac{2}{3}}\,\mathbb{1}$. Note that while the λ_θ generators generate a closed $U_V(3)$ subalgebra, the $\lambda_\theta\gamma_5$ generators have commutators that go into the $U_V(3)$ generators, since $\gamma_5^2 = \mathbb{1}$. So the $\lambda_\theta\gamma_5$ generators do not in general form a subalgebra.

$[V,V] \to V$
$[V,A] \to A$
$[A,A] \to V$

The exception is $\lambda_0\gamma_5$ and since λ_0 is proportional to the 3×3 unit matrix, $\lambda_0\gamma_5$ commutes with all other generators; it generates an axial $U_A(1)$ subgroup. This is a normal subgroup, since $\lambda_0\gamma_5$ is an ideal, so after it is excluded by the occurence of anomalies, the remaining generators form a group, which is $SU_L(3) \times SU_R(3) \times U_V(1)$.

How much of this chiral accidental symmetry is approximately linearly realized in the experimental particle spectrum? Only the $SU_V(3) \times U_V(1)$ subgroup. The reason is that the $SU_c(3)$ strong interactions give rise to condensation of quark bilinears: $\langle\bar{u}u\rangle = \langle\bar{d}d\rangle = \langle\bar{s}s\rangle = V^3$ (recall that the dimension of Fermi fields is $3/2$, so these condensates are of dimension 3).

$V \sim \Lambda_{QCD}$
equal vevs experimentally suggested

The strong interactions moreover give rise to confinement: the fermionic quark states are not directly seen in the spectrum — only $SU_c(3)$ singlet bound states are observed. This is a classic situation of spontaneous symmetry breaking, where the quark bilinear condensates break the chiral $SU_L(3)\times SU_R(3) \times U_V(1)$ symmetry spontaneously down to $SU_V(3) \times U_V(1)$. The occurrence of such condensates is an inherently nonperturbative phenomenon in QCD.

A complete derivation of the properties of the QCD vacuum and of Colour $SU_c(3)$ confinement is still lacking. (This problem is an extension of the Clay Mathematics Institute Millenium prize offered for showing confinement in pure Yang-Mills theory without quarks. Success at that would be worth one million dollars.) There are a variety of approaches to the problem, including lattice field theory. In the absence of a full ab initio derivation from microscopic QCD theory, however, one can still approach the analysis of the vacuum and spectrum from general mathematical properties of spontaneously broken symmetries.

Consistency with the observed particle spectrum shows that the axial symmetry transformations generated by $\lambda_a \gamma_5$, $a = 1, 2, \ldots 8$ must be spontaneously broken. From Goldstone's Theorem, we know that there must exist corresponding Goldstone modes. Were the $SU_L(3) \times SU_R(3)$ chiral symmetry an exact symmetry, the Goldstone modes arising from its spontaneous breakdown to $SU_V(3)$ would be exactly massless. However, we know that the chiral symmetry is actually broken by a number of effects: quark masses, couplings to $W_\mu^\pm$ and Z_μ and to A_μ. So the erstwhile Goldstone modes are actually only approximately massless. One refers to them as pseudo Goldstone bosons.

Interpretation of the particle spectrum in terms of spontaneously broken $SU_L(3) \times SU_R(3)$ works reasonably well. Experimentally, it is seen that the octet of pseudoscalar mesons is significantly lighter than the next lightest states, which form the octet of vector mesons (including the rho mesons $\sim 770\ \mathrm{MeV}/c^2$). Consistency of this picture of the pseudoscalar mesons as pseudo Goldstone bosons is given support by lattice QCD calculations which show that the dependence of pseudoscalar meson masses on quark masses is in agreement with the predictions of chiral perturbation theory. Moreover, there are two mesons, the η and the η', of which the η is a member of the pseudo Goldstone boson octet, and light, while the η' is an $SU(3)$ singlet and is significantly heavier. This fits naturally with an association of the η' with the anomalous $U_A(1)$ axial generator, which is explicitly broken so not protected by Goldstone's Theorem. This is hard to explain at the quark level.

$m_\pi \sim 135$ MeV/c^2

meson octet: $\pi^\pm$ π^0, 4 kaons, η

$\delta m_{\text{pseudoscalar meson}}$ given by δm_{quark}

Nonlinear Sigma Models

In systems with spontaneous symmetry breaking from a full symmetry group G to a stability subgroup H which is the invariance of the vacuum, the appropriate mathematical description is in terms of nonlinear realizations.
References: Coleman, Wess and Zumino, Phys. Rev. 177 (1969) 2239 and Callan, Coleman, Wess and Zumino, Phys. Rev. 177 (1969) 2247 and C.J. Isham, Nuovo Cimento A59 (1969) 356.

Let the generators of H be denoted by V_i and the remaining generators of G be denoted by Z_ℓ. The V_i and the Z_ℓ together form a complete set of generators of G, which may be taken to be orthonormal with respect to the Cartan-Killing metric $g_{ab} = f_{ad}{}^{c} f_{bc}{}^{d}$ (which is nonsingular for semisimple G).
The Lie algebra of G can be written

$$\left.\begin{array}{ll} [V_i, V_j] = i f_{ij}{}^{k} V_k & \\ [V_i, Z_\ell] = i f_{i\ell}{}^{m} Z_m & Z_\ell \text{ form a representation of } H \\ [Z_\ell, Z_m] = i f_{\ell m}{}^{n} Z_n + i f_{\ell m}{}^{k} V_k & \end{array}\right\}$$

For a compact group G, take the generators to be hermitian, structure constants $f_{ij}{}^{k}$ real.

If $f_{\ell m}{}^{n}$ in the third commutation relation vanish, then G/H is a symmetric space and the algebra of G admits an automorphism $V_i \to V_i$, $Z_\ell \to -Z_\ell$. [This is the case for chiral $SU(N)_L \times SU(N)_R$ groups: the parity operator induces an automorphism which changes the sign of the axial vector generators.]

A general group element $g \in G$ can be represented uniquely in the form $g = e^{i\xi^\ell Z_\ell} h$, where $h \in H$ and the ξ^ℓ parametrize the coset space G/H. Acting on the left with $g_0 \in G$, one can separate the product $g_0 g$ in the same way:

$$g_0 g = e^{i\xi' \cdot Z} h' \qquad \xi \cdot Z = \xi^\ell Z_\ell$$

or

$$g_0 e^{i\xi \cdot Z} = e^{i\xi' \cdot Z} h_1 \quad \text{where} \quad h_1 = h' h^{-1}$$

In general, $\xi^{m'}$ and h_1 are functions of g_0 and ξ^m.
For $g_0 = h_0 \in H$ however, one has

$$(h_0 e^{i\xi \cdot Z} h_0^{-1}) h_0 = e^{i\xi' \cdot Z} h_1$$

Since the Z_ℓ form a representation of H, this implies

$$e^{i\xi'\cdot Z} = h_0 e^{i\xi\cdot Z} h_0^{-1} \quad \text{and} \quad h' = h_0 h, \quad h_1 = h_0$$

So in this case the transformation $\xi^m \to \xi^{m'}$ is linear.
On the other hand, for $g_0 = e^{i\xi_0\cdot Z}$ one has

$$e^{i\xi_0\cdot Z} e^{i\xi\cdot Z} = e^{i\xi'\cdot Z} h_1$$

which is a nonlinear and inhomogenous transformation on the ξ^m.

For infinitesimal transformations, one has, where 1 is the identity,

$$e^{-i\xi\cdot Z}(g_0 - 1)e^{i\xi\cdot Z} - e^{-i\xi\cdot Z}\delta e^{i\xi\cdot Z} = h_1 - 1$$

Consequently, for $g_0 = e^{i\xi_0\cdot Z}$ with ξ_0^m infinitesimal, one has

$$e^{-i\xi\cdot Z}(i\xi_0\cdot Z)e^{i\xi\cdot Z} - e^{-i\xi\cdot Z}\delta e^{i\xi\cdot Z} = h_1 - 1$$

On the RHS of this equation, one has only generators V_i of H. So the coefficients of the Z_ℓ on the LHS must be set equal to zero. Using the algebra of the group G, this allows one to calculate the $\delta\xi^m$.

In the case of a symmetric-space G/H, where there is a $Z_\ell \to -Z_\ell$, $V_i \to V_i$ automorphism, one also has

$$e^{-i\xi_0\cdot Z} e^{-i\xi\cdot Z} = e^{-i\xi'\cdot Z} h_1 \qquad (\text{same } h_1)$$

This allows one to eliminate h_1 and obtain

$$e^{i\xi_0\cdot Z} e^{2i\xi\cdot Z} e^{i\xi_0\cdot Z} = e^{2i\xi'\cdot Z}.$$

From this follows that $\xi^m \to \xi^{m'}$ is a good realization of G.

Above, we have just considered the linear and nonlinear transformations of the coset parameters ξ^m in the abstract. They also give the transformations of nonlinearly transforming fields $\xi^m(x)$ under rigid G transformations generated by V_i or Z_ℓ. In order to construct covariantly transforming differential quantities, one employs the Maurer-Cartan differential form $C = C_\mu dx^\mu$ where

$$C_\mu = e^{-i\xi\cdot Z}\partial_\mu e^{i\xi\cdot Z} = i\,P_\mu\cdot Z + i\,v_\mu\cdot V.$$

Dependence: $P_\mu(\xi,\partial\xi)$, $v_\mu(\xi,\partial\xi)$

The forms $P = P_\mu dx^\mu$ and $v = v_\mu dx^\mu$ transform linearly under H.
For coset generated transformations $g_0 = e^{i\xi_0\cdot Z}$, on the other hand,

$$C_\mu' = h_1 C_\mu (h_1)^{-1} + h_1 \partial_\mu (h_1)^{-1}$$

i.e.

$$P_\mu'\cdot Z = h_1 P_\mu\cdot Z(h_1)^{-1}$$

$$v_\mu'\cdot V = h_1 v_\mu\cdot V(h_1)^{-1} - i\,h_1\partial_\mu(h_1)^{-1}$$

(transforms "locally" like a gauge field for H.)

The g_0 transformation law of P_μ^m is therefore of the same form as for a linear $h \in H$ transformation, except now with the group element $h_1(\xi, g_0)$. Accordingly, H-invariant products of the P_μ^m are automatically invariant under the full group G. For example, one can construct a G-invariant Lagrangian kinetic term $-\frac{1}{2} P_\mu^m P^{\mu n} g_{mn}$, with g_{mn} the Cartan-Killing metric for H. For compact semisimple groups G there is an appropriate normalization such that $g_{mn} = \delta_{mn}$, so then one has simply $-\frac{1}{2} P_\mu^m P^{\mu m}$ as the rigidly G-invariant kinetic term for the $\xi^m(x)$.

The transformation of the H-valued part v_μ of ω_μ is also important for the construction of rigidly G invariant Lagrangians. Let ψ be some additional field carrying a linear representation R of the stability subgroup H. Thus

$$g: \psi \to \psi' = R(h_1)\psi$$

For $h_1 = h_0 \in H$, this is a standard linear realization of the rigid transformation $h_0 \in H$. For a transformation $g_0 = e^{i\xi_0 \cdot Z}$, however, one has $h_1(\xi^m(x), g_0)$ from the general nonlinear realization structure, and although g_0 will be taken to be a rigid (x^μ independent) symmetry transformation, $h_1(\xi^m(x), g_0)$ acquires x^μ dependence from the dependence on x^μ of the nonlinearly transforming "Goldstone" field $\xi^m(x)$.

One can straightforwardly make G invariant Lagrangian terms for ψ simply by hooking up indices in a manifestly H invariant fashion. For nonderivative terms, this poses no problem, but for derivatives the x^μ dependence of $h_1(\xi(x), g_0)$ requires a refined construction. This is where the v_μ^i come in: $i v_\mu'{\cdot}V = i h_1 v_\mu{\cdot}V (h_1)^{-1} + h_1 \partial_\mu (h_1^{-1})$ is a transformation of gauge-field type, even for x^μ independent g_0. Recall that $v_\mu(\xi, \partial\xi)$ is not an independent gauge field in its own right, but is built from $\xi^m(x)$ and $\partial_\mu \xi^m(x)$ in the Cartan-Maurer form.

Use $v_\mu(\xi,\partial\xi)$ to construct a <u>covariant derivative</u> for the $R(h_1)$ representation transformation of ψ. For any representation R of H, one has $R(h\tilde{h}) = R(h)R(\tilde{h})$ for all $h,\tilde{h}\in H$. So, defining $D_\mu\psi = \partial_\mu\psi + iR(v_\mu\cdot V)\psi$ one finds
$\partial_\mu\psi \to \partial_\mu(R(h_1)\psi) = R(h_1)\partial_\mu\psi + R(\partial_\mu h_1)\psi$ and
$$iR(v_\mu\cdot V)\psi \to iR(h_1 v_\mu\cdot V(h_1^{-1}))R(h_1)\psi + R(h_1\partial_\mu(h_1^{-1}))R(h_1)\psi$$
$$= iR(h_1)R(v_\mu\cdot V)\psi - R(h_1)R(h_1^{-1})R(\partial_\mu h_1)R(h_1^{-1})R(h_1)\psi$$
$$= iR(h_1)R(v_\mu\cdot V)\psi - R(\partial_\mu h_1)\psi$$
So, putting things together $D_\mu\psi \to R(h_1)D_\mu\psi$, transforming <u>linearly</u> in the R representation of H for the element $h_1(\xi, g_0)$.

In this manner, one also can construct G invariant Lagrangians simply by hooking up the R representation indices to make invariants. Note that <u>any</u> H invariant term constructed in this way using P_μ^l and v_μ^i will yield a rigidly G-invariant term by this construction.

The construction of the covariant derivatives $\nabla_\mu\xi^l$ and $D_\mu\psi$ are made more explicit using the formula
$$w_\mu = e^{-i\xi\cdot Z}\partial_\mu e^{i\xi\cdot Z} = \left[(1-e^{-i\Delta_{\xi\cdot Z}})/\Delta_{\xi\cdot Z}\right]\partial_\mu\xi\cdot Z$$
where the operator $\Delta_{\xi\cdot Z}$ is defined by $\Delta_{\xi\cdot Z}X = [\xi\cdot Z, X]$ and functions of $\Delta_{\xi\cdot Z}$ are given by their power-series expansions. Then it follows that $(\nabla_\mu\xi^l)Z_l = P_\mu^l Z_l = (\partial_\mu\xi^l)Z_l + \cdots$ where the $\cdots$ terms are nonlinear. For systems such as the $SU_L(N)\times SU_R(N)/SU_V(N)$ coset, the commutator of two Z_l generators gives a V_i generator and it is then useful to separate the odd from the even multiple commutators. Then one obtains $v_\mu\cdot V = -i\left[(1-\cos\Delta_{\xi\cdot Z})/\Delta_{\xi\cdot Z}\right]\partial_\mu\xi\cdot Z$
and $P_\mu\cdot Z = \left[\sin\Delta_{\xi\cdot Z}/\Delta_{\xi\cdot Z}\right]\partial_\mu\xi\cdot Z$.

G/H Symmetric space

Using P_μ^l and v_μ^i, one can construct G-invariant Lagrangian terms $-\frac{1}{2}P_\mu^l P^{\mu l}$ and have consequently G-covariant field equations $(D_\mu P^\mu)^l = 0$, i.e. $\partial_\mu P^{\mu l} + v_\mu^i f_{ilm}P^{\mu m} = 0$, which can also be written $D_\mu P^\mu\cdot Z = \partial_\mu P^\mu\cdot Z + i[v_\mu\cdot V, P^\mu\cdot Z] = 0$, using the representation generators $T_{ilm} = -i f_{ilm}$ for the H rep carried by P_μ^l.

Chiral Symmetry related to nonlinear sigma models

Now let's see how $G = SU(N)_L \times SU(N)_R$ chiral symmetry as derived from N underlying quark species fits in with the standard mathematics of nonlinear realizations, based on exponentiation of G/H coset and H stability subgroup generators, with the Goldstone modes as coset parameters. For chiral symmetry, the $SU(N)_{L/R}$ factors act on the $q_L = P_L q$ and $q_R = P_R q$ projections of the quarks, respectively, and the H stability subgroup is the diagonal "vector" subgroup $SU(N)_V$. An arbitrary element g of G may be written

$P_\pm = P_{L/R} = \frac{1}{2}(\mathbb{1} \pm \gamma_5)$

$g = L(\alpha_L)R(\alpha_R)$ where $L(\alpha_L) = e^{i\alpha_L \cdot T_L}$, $R(\alpha_R) = e^{i\alpha_R \cdot T_R}$ where $T_{Li} = T_i P_L$, $T_{Ri} = T_i P_R$ and the T_i are the usual SU(N) generators (e.g. $\frac{1}{2}\sigma_i$ for SU(2) or $\frac{1}{2}\lambda_i$ for SU(3)). Since $P_L P_R = 0$, one may trivially use the BCH theorem to write

$$g = \exp\left(\tfrac{i}{2}(\alpha_L^j + \alpha_R^j)\mathcal{T}_j + \tfrac{i}{2}(\alpha_L^\ell - \alpha_R^\ell)\gamma_5 \mathcal{T}_\ell\right) \text{ where } \mathcal{T}_i = T_i \otimes \mathbb{1}_{4\times4}.$$

The H stability subgroup generators $\mathcal{T}_i$ have $\alpha_L^i = \alpha_R^i$, while the G/H cosets may be parametrized by $e^{i\xi^\ell A_\ell}$ with $A_\ell = \gamma_5 \mathcal{T}_\ell$. In a nonlinear field realization, the ξ^ℓ become Goldstone fields $\xi^\ell(x)$.

Accordingly, for a nonlinear realization of $G = SU(N)_L \times SU(N)_R$ with stability subgroup $H = SU(N)_V$, one writes

$$g(x) = e^{i\xi^\ell(x)A_\ell}\, h = U(\xi^\ell(x))\, h,$$

polarising into G/H and H factors. In terms of our general nonlinear realisation notation, the H subgroup generators are $V_i = \mathcal{T}_i$. For the V_i and broken A_ℓ generators, we have the algebra

$$[V_i, V_j] = i f_{ij}{}^k V_k, \quad [V_i, A_\ell] = i f_{i\ell}{}^m A_m \quad \text{so } A_\ell \text{ form an adjoint rep. of H}$$
$$[A_\ell, A_m] = i f_{\ell m}{}^k V_k$$

where all the $f_{ij}{}^k$, $f_{i\ell}{}^m$ and $f_{\ell m}{}^k$ are the usual SU(N) generators. This algebraic structure makes G/H a symmetric space and there is also a <u>parity automorphism</u> $V_i \to V_i$, $A_\ell \to -A_\ell$. Using the Weyl representation for the Spin(3,1) γ-matrices, one has $\gamma_5 = \mathrm{diag}(\mathbb{1}, -\mathbb{1})$ where $\mathbb{1}$ is a 2x2 unit matrix.

We have seen in the discussion of Pin groups that the parity operation on a spinor field is $\psi(x) \to \psi_p(x) = \eta_p \Lambda_p \psi(x_p)$ where $x_p^\mu = (x^0, -x^i)$, η_p is a parity phase and $\Lambda_p = i\gamma^0$. We will mainly be concerned with an effective theory of Goldstone bosons in which the η_p phases cancel out, but the Λ_p operator plays nonetheless a rôle. In the Weyl representation, one has simply $\Lambda_p = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$. Acting on γ_5 with a Λ_p similarity transformation, one has $\Lambda_p \gamma_5 \Lambda_p^{-1} = \text{diag}(-\mathbb{1}, \mathbb{1})$, ie. this transformation sends $\gamma_5 \xrightarrow{P} -\gamma_5$. Accordingly, for the $SU_L(N) \times SU_R(N)$ generators, one has $V_i \to \Lambda_p V_i \Lambda_p^{-1} = V_i$ for $V_i = \mathcal{T}_i = T_i \otimes \mathbb{1}_{4\times4}$ but for $A_\ell = \gamma_5 \mathcal{T}_\ell$ one has $A_\ell \to \Lambda_p A_\ell \Lambda_p^{-1} = -A_\ell$, thus implementing the parity automorphism.

For $SU(3)_L \times SU(3)_R$ u is a 12x12 matrix: (3x4)x(3x4)

For the coset representative field $u = e^{i\xi^\ell(x) A_\ell} = e^{i\xi^\ell(x)\mathcal{T}_\ell \gamma_5}$ one has in Weyl representation $u = \begin{pmatrix} w\mathbb{1} & 0 \\ 0 & w^\dagger \mathbb{1} \end{pmatrix}$ where $w = e^{i\xi\cdot T}$. The polarised structure of the nonlinear $u(\xi)$ realisation is determined by requiring preservation of the diagonal $u = \text{diag}(w\mathbb{1}, w^\dagger\mathbb{1})$ structure of u: $u \to u' = \text{diag}(e^{i\xi'\cdot T}\mathbb{1}, e^{-i\xi'\cdot T}\mathbb{1})$. The effect of a parity automorphism on the coset representative $u(\xi)$ is to interchange w and $w^\dagger$: $u \xrightarrow{P} u_p = \Lambda_p u \Lambda_p^{-1} = \text{diag}(w^\dagger \mathbb{1}, w\mathbb{1})$. A general $G = SU_L(N) \times SU_R(N)$ transformation by $g_0 \in G$ acting on $u(\xi)$ is given by $g_0 u(\xi) = u(\xi') h_1$ where $h_1 \in H$. Since the A_ℓ form an adjoint representation of H, for $g_0 = h_0 \in H$ one has $u(\xi') = e^{i\xi'\cdot A} = h_0 e^{i\xi\cdot A} h_0^{-1}$ and $h_1 = h_0$, so the Goldstone fields $\xi^\ell(x)$ transform linearly under H. For G/H coset transformations $g_0 = e^{i\xi_0 \cdot A}$, one has instead a nonlinear transformation $e^{i\xi_0\cdot A} u(\xi) = u(\xi') h_1(g_0, u(\xi))$.

We have seen that in order to construct Lagrangian terms that are invariant under the nonlinear G transformations, one needs to use the Maurer-Cartan differential forms to construct H invariants where the $v_\mu^i(\xi, \partial\xi)$ forms (recall $u^{-1}\partial_\mu u = i p_\mu \cdot A + i v_\mu \cdot V$) transform like gauge fields for H,

$V_\mu^{\,i} \cdot V = h_1 V_\mu \cdot V (h_1)^{-1} - i\, h_1 \partial_\mu (h_1)^{-1}$ while the $P_\mu^{\,l}$ forms transform as $P_\mu' \cdot A = h_1 P_\mu \cdot A (h_1)^{-1}$, i.e. using a linear representation of H but transforming "locally" with $h_1(g_0, u(\xi(x)))$. The $V_\mu^{\,i}$ forms are needed to make "covariant derivatives" for the "local" $h_1(g, u(\xi(x)))$ transformations.

In order to make nonlinear G invariants involving other fields, one needs to have them transform with $h_1(g_0, u(\xi(x)))$ under some linear representation R of the stability subgroup H, $g: \psi \to \psi' = R(h_1)\psi$. To obtain such fields, one may start with a set of additional fields transforming linearly under the full G group. Call these other fields "quarks", to give them a name. For $G = SU_L(N) \times SU_R(N)$, Dirac spinor quarks q will transform as $q \xrightarrow{G} q' = L q_L + R q_R$, separately on their P_L and P_R projections. In order to mesh such linear G representations with fields transforming under h_1 or the Cartan-Maurer forms, one needs variants of the q fields that transform just under H, using the corresponding $h_1(g_0, u(\xi))$ element. Such modified variants $\tilde q_L$, $\tilde q_R$ are defined by

e.g. q in fundamental of G; otherwise $q' = R(L) q_L + R(R) q_R$

$L = \begin{pmatrix} g_L & 0 \\ 0 & 1\!\!1 \end{pmatrix}$ $R = \begin{pmatrix} 1\!\!1 & 0 \\ 0 & g_R \end{pmatrix}$

$$q_L = u \tilde q_L \,, \qquad q_R = u^\dagger \tilde q_R \quad \left(\text{for } q^L_R \text{ in representation } R,\ q^L_R = {R(u) \atop R(u^\dagger)} \tilde q^L_R \right)$$

or $\tilde q_L = u^\dagger q_L$, $\tilde q_R = u q_R$

Note that, providing one implements the parity automorphism on the $\tilde q^L_R$ fields by $\Lambda_p \tilde q_L = \tilde q_R$, $\Lambda_p \tilde q_R = \tilde q_L$ (exchanging L and R projections just the same as for q_L, q_R), the above definitions of $\tilde q_L$ and $\tilde q_R$ are parity covariant, since one also has $\Lambda_p u \Lambda_p^{-1} = u^\dagger$.

Consider now making a pure $SU_L(N)$ transformation $L = \begin{pmatrix} g_* & 0 \\ 0 & 1\!\!1 \end{pmatrix}$: $\begin{pmatrix} g_* w 1\!\!1 & 0 \\ 0 & w^\dagger 1\!\!1 \end{pmatrix} = \begin{pmatrix} w' 1\!\!1 & 0 \\ 0 & w'^\dagger 1\!\!1 \end{pmatrix} h_1(g_*, u)$, i.e. $Lu = u' h_1$, then make a parity transformation : $\Lambda_p L \Lambda_p^{-1} = \begin{pmatrix} 1\!\!1 & 0 \\ 0 & g_* \end{pmatrix}$, an R transf, so $\begin{pmatrix} w^\dagger 1\!\!1 & 0 \\ 0 & g_* w 1\!\!1 \end{pmatrix} = \begin{pmatrix} w'^\dagger 1\!\!1 & 0 \\ 0 & w' 1\!\!1 \end{pmatrix} h_1(g_*, u)$, which is of the form $R u^\dagger = u^{\dagger\prime} h_1$. Combine then the rules $Lu = u' h_1$; $R u^\dagger = u^{\dagger\prime} h_1$ with the defining relations for $\tilde q^L_R$:

$$q' = L u \tilde q_L + R u^\dagger \tilde q_R = u' h_1 \tilde q_L + u^{\dagger\prime} h_1 \tilde q_R \text{ to find}$$

$\tilde q_L' = h_1 \tilde q_L$ and $\tilde q_R' = h_1 \tilde q_R$.

Accordingly, the $\tilde{q}_L$ and $\tilde{q}_R$ modified "quarks" transform just with the "local" $h_1 \in H$ subgroup transformation, as needed for constructing nonlinear realization invariants.

Note that by hermitian conjugating $R u^+ = u^{+\prime} h_1$ one obtains $u R^+ = (h_1)^+ u'$. Heretofore, the L, R labels denoted left & right chiral projections on spinor fields. However, one may also reinterpret the $u R^+ = (h_1)^+ u'$ relation in terms of a right-sided transformation of the coset representative. Similarly, hermitian conjugation of $L u = u' h_1$ gives $u^+ L^+ = (h_1)^+ u^{+\prime}$. Summarising, for $g_0 = L$ respectively $g_0 = R$, one has

$R^{-1} = R^+$
$(h_1)^{-1} = h^+$
for $SU(N)$

$$u' = L u h_1^+ \quad \text{resp.} \quad u' = h_1 u R^+$$
$$u^{+\prime} = h_1 u^+ L^+ \quad \text{resp.} \quad u^{+\prime} = R u^+ h^+ \quad .$$

Now consider the detail of which $h_1(g_0, u) \in SU(N)_V$ transformation is induced by the L, R transformations with a given $g_* \in SU(N)$. For $L = \begin{pmatrix} g_* & 0 \\ 0 & \mathbb{1} \end{pmatrix}$ one has $\Lambda_P L \Lambda_P^{-1} = R = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & g_* \end{pmatrix}$ as the parity conjugate. The L transformation of u is

$$u' = \begin{pmatrix} g_* & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} w \mathbb{1} & 0 \\ 0 & w^+ \mathbb{1} \end{pmatrix} h_1^+ = \begin{pmatrix} w' \mathbb{1} & 0 \\ 0 & w'^+ \mathbb{1} \end{pmatrix} .$$

Preserving the diagonal G/H representative form of u' requires $(g_* w h_1^+)^+ = w^+ h_1^+$ i.e. $h_1 w^+ g_*^+ = w^+ h_1^+$. This nonlinear condition implicitly determines $h_1(L, u)$ in terms of g_* and u. Now consider the parity conjugate $R = L_P = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & g_* \end{pmatrix}$ transformation:

$$u' = h_1 \begin{pmatrix} w \mathbb{1} & 0 \\ 0 & w^+ \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & g_*^+ \end{pmatrix} = \begin{pmatrix} w' \mathbb{1} & 0 \\ 0 & w'^+ \mathbb{1} \end{pmatrix} .$$

Preserving the diagonal G/H representative form of this u' requires $(h_1 w)^+ = h_1 w^+ g_*^+$, i.e. $w^+ h_1^+ = h_1 w^+ g_*^+$, which is the <u>same</u> condition determining h_1 in terms of g_* and u. Thus, the $L = \mathrm{diag}(g_*, \mathbb{1})$ and $R = L_P = \mathrm{diag}(\mathbb{1}, g_*)$ transformations induce the same $h_1 \in H = SU(N)_V$ stability subgroup element.

The identity of the $h_1(g_*, u)$ stability subgroup elements obtained from left-side and right-side versions of a given $g_0 \in G$ transformation extends from the pure L or R

transformations considered separately to combined $g_0 = X = LR$ transformations. For $X = \mathrm{diag}(\ell \mathbb{1}, r\mathbb{1})$, one finds that $h_1(g_0, u(x))$ is determined by the nonlinear relation $h_1 u^+ \ell^+ = r u^+ h_1^+$ equivalently for left-sided and right-sided versions of the transformation:

$$u' = X u h_1^+ \longleftrightarrow u' = h_1 u X_p^+$$
$$u^{+\prime} = h_1 u^+ X^+ \longleftrightarrow u^{+\prime} = X_p u^+ h_1^+$$

where $X_p = \Lambda_p X \Lambda_p^{-1} = \mathrm{diag}(r\mathbb{1}, \ell\mathbb{1})$ is the parity conjugate of g_0.

The lowest-lying hadrons are described in the chiral effective field theory by composites of the modified $\tilde{q}_L$ and $\tilde{q}_R$ fields. Since these transform only under the nonlinear $h_1 \in SU(3)_V$, there are fewer representation types that can be built from them as composites than there would be under a linear realisation of the full $G = SU(3)_L \times SU(3)_R$. For example, bilinears in $\tilde{q}_L$ and $\tilde{q}_R$ form the chiral effective theory compositions of the mesons. These can have spin zero or spin one. Even-parity spin zero composites are formed from $i(\bar{\tilde{q}}_L \tilde{q}_R + \bar{\tilde{q}}_R \tilde{q}_L)$, which decomposes under $SU(3)_V$ into an $\underline{8} \oplus \underline{1}$ rep., or "nonet". Such scalar mesons exist in the particle spectrum, but their experimental status is somewhat unclear owing to large decay widths and the existence of multiple decay channels. The pseudoscalar mesons, on the other hand, are represented by the G/H coset representative u itself, transforming linearly under $SU(3)_V$ but nonlinearly under the G/H broken symmetries. Interpreted as pseudo Goldstone bosons, they are relatively light, with masses occasioned by quark masses in the underlying QCD theory, as given in first approximation by the Gell-Mann-Oakes-Renner relation.

no $L \leftrightarrow R$ distinction in h_1 transformations

Note that the analysis of allowable $\tilde{q}$ composites goes exactly the same as solely under consideration of $SU(3)_V$, $\mathrm{Spin}(3,1)$ and $SU(3)_{\mathrm{colour}}$ (which requires overall singlets for confinement). Building such $\tilde{q}$ composites indicates in addition how the pion multiplet of pseudo Goldstone bosons enters into the effective theory construction.

Now we can consider how the conventional formulation of the chiral effective field theory in terms of an $SU(N)$ group element

$\hat{u} \in SU(N)$ $\hat{u}$ transforming as $\hat{u} \to g_L \hat{u} g_R^\dagger$ relates to the general mathematical formalism for nonlinear sigma models. We do this for $\mathcal{U} = \begin{pmatrix} \hat{u}\mathbb{1} & 0 \\ 0 & \hat{u}^\dagger \mathbb{1} \end{pmatrix}$ by letting $\mathcal{U} = u^2$, i.e. $\hat{u} = \omega^2$, where the two u factors transform separately by left-side and right-side G transformations: for $X \in G = SU(N)_L \times SU(N)_R$ one has $\mathcal{U}' = u'^2 = (X u h_1^\dagger)(h_1 u X_p^\dagger) = X u^2 X_p^\dagger = X \mathcal{U} X_p^\dagger$ in which the $h_1(X,u)$ cancels out. For general G transformations,

so $\hat{u} \to l \hat{u} r^\dagger$

$X = \begin{pmatrix} l\mathbb{1} & 0 \\ 0 & r\mathbb{1} \end{pmatrix}$ and $X_p = \begin{pmatrix} r\mathbb{1} & 0 \\ 0 & l\mathbb{1} \end{pmatrix}$ so $\mathcal{U}' = \begin{pmatrix} l\mathbb{1} & 0 \\ 0 & r\mathbb{1} \end{pmatrix} \mathcal{U} \begin{pmatrix} r^\dagger \mathbb{1} & 0 \\ 0 & l^\dagger \mathbb{1} \end{pmatrix}$.

Elements of the $H = SU(N)_V$ stability subgroup have $l = r$, while the G/H coset representatives have $l = g_* = r^\dagger$, i.e. structure $X = \begin{pmatrix} g_* \mathbb{1} & 0 \\ 0 & g_*^\dagger \mathbb{1} \end{pmatrix}$ for $g_* \in SU(N)$. Accordingly, under G/H transformations, one has $\mathcal{U}' = \begin{pmatrix} g_*\mathbb{1} & 0 \\ 0 & g_*^\dagger \mathbb{1} \end{pmatrix} \mathcal{U} \begin{pmatrix} g_*\mathbb{1} & 0 \\ 0 & g_*^\dagger \mathbb{1} \end{pmatrix} = X \mathcal{U} X$.

This is exactly of the form

$$e^{2i\xi' \cdot A} = e^{i\xi_0 \cdot A} e^{2i\xi \cdot A} e^{i\xi_0 \cdot A}$$

that one obtains for G/H nonlinear transformations when there is an $A_\ell \to -A_\ell$, $V_i \to V_i$ automorphism.

The Maurer-Cartan form for this system is, for $u = \begin{pmatrix} \omega\mathbb{1} & 0 \\ 0 & \omega^\dagger \mathbb{1} \end{pmatrix}$, given by

$$C = u^{-1} du = \begin{pmatrix} \omega^\dagger d\omega\, \mathbb{1} & 0 \\ 0 & \omega\, d\omega^\dagger \mathbb{1} \end{pmatrix}.$$

Break this into G/H and H generators:

note: $d(\omega\omega^\dagger) = 0$ i.e. $d\omega\omega^\dagger + \omega d\omega^\dagger = 0$

$$P = P^\ell A_\ell = \frac{1}{2} \begin{pmatrix} (\omega^\dagger d\omega - \omega d\omega^\dagger)\mathbb{1} & 0 \\ 0 & (\omega d\omega^\dagger - \omega^\dagger d\omega)\mathbb{1} \end{pmatrix}$$

$$v = v^i V_i = \frac{1}{2} \begin{pmatrix} (\omega^\dagger d\omega^\dagger + \omega d\omega^\dagger)\mathbb{1} & 0 \\ 0 & (\omega d\omega^\dagger + \omega^\dagger d\omega)\mathbb{1} \end{pmatrix}$$

Then the Goldstone field (pion multiplet) kinetic term is

$\overline{\mathrm{tr}}$: trace on $Spin(3,1)$ and $SU(N)$ indices

tr: trace just on $SU(N)$ indices

$$\mathcal{L}_{kin} = \overline{\mathrm{tr}}(P_\mu P^\mu) = \frac{4}{4} \mathrm{tr}(\omega^\dagger \partial_\mu \omega - \omega \partial_\mu \omega^\dagger)(\omega^\dagger \partial^\mu \omega - \omega \partial^\mu \omega^\dagger)$$

$$= -2\, \mathrm{tr}(\partial_\mu \omega^\dagger \partial^\mu \omega + \omega^\dagger (\partial_\mu \omega) \omega (\partial^\mu \omega^\dagger))$$

$$= -\mathrm{tr}(\partial_\mu (\omega^{\dagger 2}) \partial^\mu (\omega^2)) = -\mathrm{tr}(\partial_\mu \hat{u} \partial^\mu \hat{u}^\dagger)$$

$$= -\frac{1}{4} \overline{\mathrm{tr}}(\partial_\mu \mathcal{U} \partial^\mu \mathcal{U}^\dagger)$$

Hence the normalised kinetic term for the $SU(2)_L \times SU(2)_R$ pion multiplet is

$$\mathcal{L}_{kin}^\pi = -\frac{F_\pi^2}{4} \mathrm{tr}\left(\partial_\mu (e^{2i\pi\cdot\tau/F_\pi}) \partial^\mu (e^{-2i\pi\cdot\tau/F_\pi})\right)$$

$$= -\frac{F_\pi^2}{16} \overline{\mathrm{tr}}(\partial_\mu \mathcal{U} \partial^\mu \mathcal{U}^\dagger) = +\frac{F_\pi^2}{4} \overline{\mathrm{tr}}(P_\mu P^\mu).$$

Chiral Effective field Theory

Now consider how to construct an effective theory Lagrangian. A conventional way to write the nonlinear sigma model for $SU_L(3)\times SU_R(3)/SU_V(3)$ is in terms of the $SU(3)$ matrix

π^a transform as an adjoint 8 under $SU(3)$

$$\overset{(3)}{U}(x) = \exp\left(2i\frac{\pi^a T^a}{F_\pi}\right) = \exp\left[\frac{\sqrt{2}\,i}{F_\pi}\begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0+\frac{1}{\sqrt{6}}\eta^0 & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0+\frac{1}{\sqrt{6}}\eta^0 & K^0 \\ \bar{K}^- & \bar{K}^0 & -\sqrt{2/3}\,\eta^0 \end{pmatrix}\right]$$

where the $SU(3)$ generators are $T^a = \frac{1}{2}\lambda^a$, with λ^a the eight Gell-Mann matrices. F_π is a constant with dimensions of mass/energy, known as the pion decay constant. Fitting to data gives $F_\pi = 92\ \mathrm{MeV}/c^2$.

The $SU_L(3)\times SU_R(3)/SU_V(3)$ chiral sigma model gives an account of the interactions of pions, kaons and the η^0 that works well phenomenologically after the large strange quark mass is incorporated via perturbation theory. To simplify the story, however, let us now exclude the third, strange, flavour and focus just on the $SU_L(2)\times SU_R(2)/SU_V(2)$ subsystem for the pions:

$\pi^\pm = \frac{1}{2}(\pi^1 \mp i\pi^2)$

$$U(x) = \exp\left(2i\frac{\pi^a \tau^a}{F_\pi}\right) = \exp\left[\frac{i}{F_\pi}\begin{pmatrix} \pi^0 & \sqrt{2}\pi^+ \\ \sqrt{2}\pi^- & -\pi^0 \end{pmatrix}\right] \in SU(2)$$

where $\tau^a = \frac{1}{2}\sigma^a$ are the $SU(2)$ generators. Under $SU_L(2)\times SU_R(2)$, $U(x)$ transforms as $U \to g_L U g_R^\dagger$. The unbroken $SU_V(2)$ has $g_L = g_R$. The simplest invariant under $SU_L(2)\times SU_R(2)$ that one can write is

$$\mathcal{L}_{chiral} = -\frac{F_\pi^2}{4}\,\mathrm{tr}\left[(D^\mu U)(D_\mu U)^\dagger\right]$$

where the electromagnetic coupling to A_μ has been restored in

D_μ gauges τ^3 in U

$D_\mu\phi = (\partial_\mu - iq_\phi A_\mu)\phi$ for a scalar field with electromagnetic charge ϕ. This explicitly breaks the $SU_L(2)\times SU_R(2)$ rigid chiral symmetry, but is included for physical completeness. Setting A_μ to zero restores the chiral symmetry. Expanding $\mathcal{L}_{chiral}$ out to quadratic order yields normal scalar kinetic terms and electromagnetic couplings, followed by specific interactions

$$\mathcal{L}_{chiral} = -\frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 - D_\mu\pi^+D^\mu\pi^- + \frac{1}{3F_\pi^2}\left((\pi^0)^2\partial_\mu\pi^+\partial^\mu\pi^- + \cdots\right) - \frac{1}{18F_\pi^4}\left((\pi^+\pi^-)^2\partial_\mu\pi^0\partial^\mu\pi^0 + \cdots\right) + \cdots$$

In this non-renormalizable effective theory, terms with more derivatives are also expected, but these are constrained by the $SU_L(2) \times SU_R(2)$ symmetry. The next possible terms have four derivatives:

$$\mathcal{L}_4 = L_1 \left(\mathrm{tr}\left[(D_\mu U)(D^\mu U)^\dagger\right]\right)^2 + L_2\, \mathrm{tr}\left[D_\mu U (D_\nu U)^\dagger D^\mu U (D^\nu U)^\dagger\right] + L_3\, \mathrm{tr}\left[D_\mu U (D^\mu U)^\dagger D_\nu U (D^\nu U)^\dagger\right]$$

with the coefficients determined from scattering experiments

$$L_1 = 0.65 \qquad L_2 = 1.89 \qquad L_3 = -3.06$$

Since these terms have extra derivatives, their contributions at low energies will be suppressed at low energies by powers of $\frac{E}{F_\pi}$ compared to the predictions of the leading two-derivative term, for which the $-\frac{F_\pi^2}{4}$ coefficient gives canonically normalized $-\frac{1}{2}\partial_\mu \pi^0 \partial^\mu \pi^0 - D_\mu \pi^+ D^\mu \pi^-$ kinetic terms.

E: energy

Now let us see what can be said about the effect of quark mass terms in the underlying QCD theory. Quark masses explicitly break the chiral $SU_L(2) \times SU_R(2)$ symmetry, with the consequence that the pions are not massless Goldstone bosons, but instead are light pseudo Goldstone bosons. Write the quark mass terms as $\mathcal{L}_m = -i\bar{q} M q$, $q = \binom{u}{d}$ with $M = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix}$. Now we employ a trick to restore the chiral symmetry that is broken by $\mathcal{L}_m$: assign a transformation property to M: $M \to g_R M g_L^\dagger$. This makes $\mathcal{L}_m$ invariant, formally. When constants are treated as fields in this way, they are called <u>spurions</u>.

Although the low-energy effective theory of pions does not contain quarks, its details can depend on the mass matrix M. One can use the spurious $SU_L(2) \times SU_R(2)$ to constrain the effect of the quark mass matrix M on the nonlinear chiral effective Lagrangian. The leading $SU_L(2) \times SU_R(2)$ term that can be added to the pion nonlinear sigma model is $\mathcal{L}_M = \frac{V^3}{2} \mathrm{tr}(MU + M^\dagger U^\dagger)$. Expanding this, one obtains $\mathcal{L}_M = V^3(m_u + m_d) - \frac{V^3}{2F_\pi^2}(m_u+m_d)(\pi_0^2 + \pi_1^2 + \pi_2^2) + \mathcal{O}(\pi^3)$

The coefficient $\frac{V^3}{2}$ is fixed by requiring the vacuum energy arising from $\mathcal{L}_M$ to match that arising from $\mathcal{L}_m$. For $\langle \bar{u}u \rangle = \langle \bar{d}d \rangle = V^3$, one has $\mathcal{L}_m = V^3(m_u + m_d)$, which matches the constant term in the expansion of $\mathcal{L}_M$. Given that assignment, one obtains the pion masses from the quadratic term in pions: $m_\pi^2 = \frac{V^3}{F_\pi^2}(m_u + m_d)$.

This is the <u>Gell-Mann-Oakes-Renner</u> relation: the square of the pion mass scales linearly with the quark masses. For $V \sim \Lambda_{QCD} \sim 250\ \mathrm{MeV}/c^2$, $F_\pi = 92\ \mathrm{MeV}/c^2$ and $m_\pi = 140\ \mathrm{MeV}/c^2$

m_u, m_d: "current quark masses"

one obtains $m_u + m_d \sim 10.6\ \mathrm{MeV}/c^2$. This sort of value agrees with results from lattice QCD. So, although based just on leading approximations, the constraints of spontaneously broken chiral symmetry give rather good results in comparison with experiment and more elaborate phenomenological calculations.

Another relation that one can obtain from the effective chiral symmetry theory concerns interactions between pions and nucleons. Historically, interactions between pions π^k and nucleons $N = \binom{p}{n}$ were described in terms of a Yukawa interaction $\mathcal{L}_{NN\pi} = -g_{NN\pi}(\bar{N}\gamma_5\tau_k N)\pi^k$, with $g_{NN\pi}$ found to be close to 14. (Note: γ_5 coupling since π^k are pseudoscalars)

On the other hand, consider the most general $SU_L(2) \times SU_R(2)$ and parity invariant coupling between pions and nucleons

$N \to g_R N$

$$\mathcal{L}_{\pi NN} = -i\bar{N}(\partial\!\!\!/ + m)N - \frac{ig}{2}\bar{N}\gamma^\mu\gamma_5 U^\dagger\partial_\mu U N$$
$$= -i\bar{N}(\partial\!\!\!/ + m)N + \frac{g}{2F_\pi}\bar{N}\gamma^\mu\gamma_5\tau_k N\partial_\mu\pi^k + \frac{1}{4F_\pi^2}\epsilon_{k\ell m}\bar{N}\gamma^\mu\tau^k N\pi^\ell\partial_\mu\pi^m + \cdots$$

Integrating ∂_μ by parts and using the nucleon Dirac equation $(\partial\!\!\!/ + m)N = 0$ yields an interaction $-\frac{mg}{F_\pi}\bar{N}\gamma_5\tau_k N\pi^k + \cdots$ leading to the identification $g_{NN\pi} = \frac{gm}{F_\pi}$.

This is the <u>Goldberger-Treiman relation.</u>

The value of the pion-nucleon coupling constant g is obtained from consideration of neutron decay. Obtain the Noether currents $j_V^{\mu i}$ and $j_A^{\mu i}$ from the $SU_L(2)\times SU_R(2)$ transformations $g_L = \mathbb{1} + i(\delta\alpha + \delta\beta)\cdot\tau$ $\quad g_R = \mathbb{1} + i(\delta\alpha - \delta\beta)\cdot\tau$ and let $\delta\alpha^i$ and $\delta\beta^i$ become space-time dependent to calculate the currents. In the original QCD action, one has $\delta I_{QCD} = -\int d^4x \left(\partial_\mu \delta\alpha_i\, j_V^{\mu i} + \partial_\mu \delta\beta_i\, j_A^{\mu i}\right)$
while in the effective low-energy theory, one has $U' = g_L U g_R^{-1}$,
i.e. $\delta U = i\left(\delta\alpha\cdot[\tau, u] + \delta\beta\cdot\{\tau, u\}\right)$
so from $\delta I_{pion} = -\int d^4x \left(\partial_\mu\delta\alpha\cdot J_V^\mu + \partial_\mu\delta\beta\cdot J_A^\mu\right)$
one finds

$$J^\mu_{Vk} = -\epsilon_{k\ell m}\pi_\ell \partial^\mu \pi_m + \cdots \qquad \left[\delta_\alpha \pi^i = -\epsilon^{ijk}\delta\alpha^j\pi^k + \cdots\right]$$

$$J^\mu_{Ak} = F_\pi \partial^\mu \pi_k + \cdots \qquad \left[\delta_\beta \pi^i = \tfrac{F_\pi}{2}\delta\beta^i + \cdots\right]$$

When coupling to nucleons is included, these become

$$J^\mu_{Vk} = -\epsilon_{k\ell m}\pi_\ell \partial^\mu \pi_m + \frac{i}{2}\bar{N}\gamma^\mu \tau_k N$$

$$J^\mu_{Ak} = F_\pi \partial^\mu \pi_k + \frac{ig}{2}\bar{N}\gamma^\mu\gamma_5 \tau_k N + \cdots$$

Nucleon matrix elements are

$\sigma,\sigma' = 1,2$ pm

$$\langle N(p,\sigma)|J^\mu_{Vk}|N'(p',\sigma')\rangle = \frac{i}{2}e^{iq\cdot x}\bar{N}\gamma^\mu F_1(q^2)\tau_k N + \cdots$$

$$\langle N(p,\sigma)|J^\mu_{Ak}|N'(p',\sigma')\rangle = \frac{i}{2}e^{iq\cdot x}\bar{N}\gamma^\mu\gamma_5 G_1(q^2)\tau_k N + \cdots$$

with $F_1(0) = 1$ and $G_1(0) = g$, $q^\mu = p^\mu - p^{\mu\prime}$
Then from the neutron decay rate, one finds $g = 1.269$

Anomalies

We saw that the classical QCD Lagrangian with just the three u,d,s quarks had a $U_L(3)\times U_R(3)$ symmetry, or, keeping just the u,d quarks, a $U_L(2)\times U_R(2)$ symmetry. Spontaneous symmetry breaking down to $SU_V(2)\times U_V(1)$ (with $U_V(1) = U_B(1)$, Baryon number symmetry) gives a good approximate description of linearly transforming baryon and meson multiplets other than the pions. The pions (and their kaon and η^0 partners when the s quark is included) are interpreted as Goldstone bosons, or pseudo Goldstone bosons when quark masses are turned on. As mentioned earlier however, the count of such apparent Goldstone modes is wrong for the $U_L(2)\times U_R(2)$ symmetry: in addition to the π^a, $a=1,2,3$ for the broken $\frac{1}{2}\sigma^a\gamma_5$ generators, there would be one more, for the $\mathbb{1}\gamma_5$ generator, corresponding to $U_A(1)$ chiral transformations $q_i \to e^{i\beta\gamma_5}q_i$: $\bar{q}_i \to (e^{i\beta\gamma_5}q_i)^+\gamma^0 = q_i^+ e^{-i\beta\gamma_5}\gamma^0 = \bar{q}_i e^{i\beta\gamma_5}$ for each u, d quark type. Although there is a candidate pseudoscalar meson, the η', which would seem to fit into the 3-flavour $U_L(3)\times U_R(3)$ picture, it is not light ($m_{\eta'} = 958\ \mathrm{MeV}/c^2$ as compared to $m_\pi \sim 135\ \mathrm{MeV}/c^2$).

What is happening is the phenomenon of anomalies: although the axial $U_A(1)$ is a good symmetry at the classical level, it is broken at the quantum level. Derive the Noether current for $U_A(1)$ by the standard procedure of letting the parameter $\beta \to \beta(x)$ become spacetime dependent. Then

ignore SU(3) coupling here

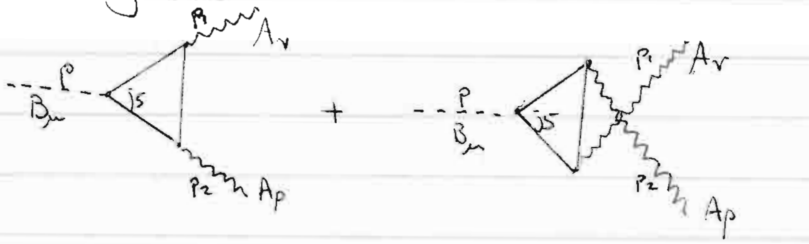
$\delta(-i\bar{q}_i\gamma^\mu\partial_\mu q_i) = -i\bar{q}_i e^{i\beta\gamma_5}\gamma^\mu\partial_\mu e^{i\beta\gamma_5}q_i = \bar{q}_i\gamma^\mu\gamma_5 q_i\,\partial_\mu\beta = j_5^\mu\partial_\mu\beta$, so the $U_A(1)$ axial current is $j_5^\mu = \sum_i \bar{q}_i\gamma^\mu\gamma_5 q_i$, summed over all quark types q_i. If one includes $-im_i\bar{q}_iq_i$ mass terms for the quarks, use of the Dirac equation $(\not{\partial}+m_i)q_i=0$ and $\partial^\mu\bar{q}_i\gamma_\mu - m_i\bar{q}_i = 0$ gives the broken conservation equation $\partial_\mu j_5^\mu = \partial^\mu(\bar{q}_i\gamma^\mu\gamma_5 q_i) = 2m_i\bar{q}_i\gamma_5 q_i$, so for $m_i=0$ the j_5^μ is classically conserved.

To see what happens to this axial current at the quantum level, restrict the discussion to just one species of fermion, $q_i \to q$, and couple it to an electromagnetic field A_μ, and also to a fictitious B_μ gauge field that couples to the J_μ^5 current:

$$-\tfrac{1}{4}F^{(A)}_{\mu\nu}F^{\mu\nu}(A) - i\bar{q}(\not{\partial} - ie\not{A} + m)q - B_\mu\bar{q}\gamma^\mu\gamma_5 q\,.$$

No kinetic term is included for B_μ — its rôle is purely to facilitate discussion of quantum matrix elements involving insertions of j_5^μ.

There are a variety of ways to evaluate the effects of j^μ_5 insertions. The most straightforward way proceeds with Feynman diagrams



B_μ lines couple to j^μ_5

We are particularly interested in the case where $B_\mu \to \partial_\mu \beta$, i.e. where one evaluates matrix elements $\langle A | \partial_\mu j^\mu_5 | A \rangle$ taken between external electromagnetic fields A_μ. As we have seen, $\partial_\mu j^\mu_5 = 2m\bar{q}\gamma_5 q$ for massive fermions q.

The triangle diagrams for the insertion of j^μ_5 look superficially linearly divergent $\sim \int \frac{d^4k}{k^3}$, but in fact converge since external momenta $p_1^\mu p_2^\nu$ come out, leaving an ultraviolet convergent result, at worst $\int \frac{d^4k}{k^5}$. One finds $\langle A | \bar{\Psi}\gamma^5\Psi | A \rangle = \frac{e^2}{32\pi^2 m} \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu}(A) F_{\alpha\beta}(A)$. This suggests that

$$\langle A | \partial_\mu j^\mu_5 | A \rangle = \frac{e^2}{16\pi^2} \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu}(A) F_{\alpha\beta}(A) = \frac{e^2}{8\pi^2} F_{\mu\nu}(A) {}^*F^{\mu\nu}(A)$$

where ${}^*F^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}$. Note that the mass m of the fermion field going around the triangle loop has cancelled out.

The above "anomalous" result was very puzzling when it was recognized in the post-war 1940s (eg. Steinberger, 1949). For zero fermion mass, the axial current is classically conserved, but the non-vanishing and m independent result gives non-conservation at the quantum level. Doing the calculation directly for $m=0$ massless fermions produced a host of other puzzles, not least how to obtain an electromagnetic gauge invariant result. Schwinger solved that problem in 1951. The modern understanding of anomalies was achieved through the work of Adler, Bell and Jackiw in 1969. The most elegant derivation of anomalies was given by Fujikawa in 1979-1981 in terms of non-invariance of the quantum field theory path integral measure under chiral γ_5 transformations.

The physical implications of axial anomalies for rigid symmetries are varied. Recalling that the pseudoscalar π^0 pion also couples to fermions q through $\pi^0 \bar{q}\gamma_5 q$, the triangle diagrams give the dominant contribution to $\pi^0 \to 2\gamma$ (2 gamma ray) decay. The non-conservation of the axial $U_A(1)$ current in the low-energy effective pion theory explains why the η' is not protected from a large mass in the same way as the pions (and, to a lesser extent, the kaons and η^0). So rigid axial anomalies can produce physically important results.

Anomalies can also occur for chirally coupled fermions in gauge theories, as well, and there they are definitely not welcome. Anomalies in Yang-Mills gauge currents ruin renormalizability and can produce strange fractional degrees of freedom. So a key aim in constructing a gauge theory such as the Standard Model is to ensure anomaly freedom. For left-handed spinor fields ψ^m transforming under some symmetry as $\delta\psi^m = i\epsilon^a (T_a)_m{}^n P_L \psi_n$ where $(T^a)_m{}^n$ are hermitian symmetry generators, a theory will be anomaly free if the anomaly coefficients

A_{abc} totally (abc) symmetric

$\{A,B\} = AB+BA$

$$A_{abc} = \mathrm{tr}(T_a \{T_b, T_c\})$$ vanish for all a, b and c,

where the trace is taken over all types of fermions. So, for example, one needs to sum over every quark of every flavour and every lepton in each generation. Fields such as the e_{Rm} in the Standard Model which are defined to be right-handed are included in the calculation of A_{abc} by listing their left-handed charge congigates with appropriately congugated T_a reps. (like the left-handed e_{Rmc}). When the anomaly coefficients don't vanish, the anomaly is

$$\partial_\mu J^\mu_a = \frac{A_{abc}}{64\pi^2} g^2 \epsilon^{\mu\nu\alpha\beta} F^b_{\mu\nu} F^c_{\alpha\beta}.$$

A consequence of the anomaly coefficient structure is that anomalies cancel for <u>real</u> or <u>pseudoreal</u> fermion representations T_a. A pseudoreal representation is one for which the generators (iT_a) are real up to a similarity transformation, i.e. $T_a^* = -S T_a S^{-1}$ for some invertible matrix S. To see how this works, note that for hermitian generators $T_a^T = T_a^*$, so

$$A_{abc} = \mathrm{tr}\left(T_a \{T_b, T_c\}\right) = \mathrm{tr}\left[\left(T_a\{T_b,T_c\}\right)^T\right]$$
$$= \mathrm{tr}\left(\{T_c^T, T_b^T\} T_a^T\right) = \mathrm{tr}\left(\{T_c^*, T_b^*\} T_a^*\right)$$
$$= -\mathrm{tr}\left(S\{T_c, T_b\} T_a S^{-1}\right) = -\mathrm{tr}\left(\{T_c, T_b\} T_a\right) = -A_{abc} = 0$$

since the A_{abc} are totally symmetric.

As an application of this result, consider a case where fermion number is a good conserved quantity and where left- and right-handed fermions transform in the same t_a of the group in question. (For this purpose, the right-handed fermions are considered in their right-handed forms, not yet charge-conjugated.) Then, assembling the total representation T_a of left-handed fermions (right-handed now charge conjugated), one may write T_a in block-diagonal form:
$$T_a = \begin{pmatrix} t_a & 0 \\ 0 & -t_a^* \end{pmatrix}$$

since charge conjugation sends $e^{i\epsilon^a t_a}$ to $e^{-i\epsilon^a t_a^*}$. This total representation T_a is consequently pseudoreal, since $T_a^* = -S T_a S^{-1}$ with $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Consequently, any symmetry which is left-right symmetric in this way is anomaly-free, for example in theories with just Dirac fermions that are given an overall symmetry representation t_a.

An important application of the existence of anomalies in the case of a rigid chiral symmetry is to π^0 decay. Recall that $U = \exp\left(\frac{i}{F_\pi}\begin{pmatrix} \pi^0 & \sqrt{2}\pi^+ \\ \sqrt{2}\pi^- & -\pi^0 \end{pmatrix}\right)$ and $U \to g_L U g_R^+$, $g_L \in SU(2)_L$, $g_R \in SU(2)_R$ with the axial (spontaneously broken) symmetries having $g_R = g_L^{+}$.

Let $g_L = \exp\left(i\frac{\omega}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right) = g_R^+$ so under this transformation $\exp\left(\frac{i}{F_\pi}\pi^0\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right) \to \exp\left(i\frac{\omega}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right)\exp\left(\frac{i}{F_\pi}\pi^0\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right)\exp\left(i\frac{\omega}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right)$

i.e. $\pi^0 \to \pi^0 + \omega F_\pi$.

Recalling the electromagnetic couplings $D_\mu = (\partial_\mu - iqA_\mu)$ for charged fields of charge q in the chiral Lagrangian $\mathcal{L}_{chiral} = -\frac{F_\pi^2}{4}\mathrm{tr}\left[(D^\mu U)(D_\mu U)^\dagger\right]$, one notes that the chiral $SU_L(2)\times SU_R(2)$ symmetry is actually broken by such couplings (it is also broken by small corrections $\sim(m_u+m_d)$ as one saw from the Gell-Mann-Oakes-Renner relation). The only part of the $SU_L(2)\times SU_R(2)$ axial symmetry that survives as a symmetry of the combined strong plus electromagnetic theory is the part that commutes with the $\binom{u}{d}$ quark electric-charge matrix $Q_{Em} = \begin{pmatrix}\frac{2}{3}&0\\0&-\frac{1}{3}\end{pmatrix}$. The surviving axial generator is $T_{3A} = \begin{pmatrix}1/2&0\\0&-1/2\end{pmatrix}\gamma_5$ (i.e. $\tau^3\gamma_5$).

broken: $\mathcal{O}(1\%)$ corrections for $\pi^\pm$

Note rôle of γ_5: RH fermions transf. opposite to LH under T_{3A}. But then $(RH)_c$ transforms the same as LH, doubling result.

Experimentally, the decay of the neutral pion π^0 into photons is well-described by an interaction term

$$\mathcal{L}_{\pi^0AA} = \frac{e^2}{32\pi^2F_\pi}\pi^0\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}(A)F^{\rho\sigma}(A).$$

The problem with such an interaction term, however, is that it does not respect the $U_{3A}(1)$ transformation generated by the supposedly surviving T_{3A}, under which $\pi^0 \to \pi^0 + \omega F_\pi$. Other interactions like $\mathcal{L}' = \frac{e^2}{32\pi^2F_\pi}\left(\frac{\Box\pi^0}{\Lambda^2}\right)\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}$ would address that problem, but for $\Lambda \sim 1\,\mathrm{GeV}/c^2$, they give decay rates that are too small for any value of Λ large enough to justify use of the effective pion theory.

What is happening, however, is an anomaly in $U_{3A}(1)$, specifically the anomaly arising from $A_{3A\,Q_{Em}Q_{Em}} = A$. From $A_{abc} = \mathrm{tr}(T_a,\{T_b,T_c\})$, one has $A = \frac{2N_c}{2}\left((2/3)^2 - (-1/3)^2\right)\times 2$ where the final factor of 2 comes from inclusion of the charge-conjugated right-handed u_{Rc} and d_{Rc} quarks.

The N_c factor comes from the number of colours from the $SU_c(3)$ strong interactions. In fact, this is a key point. The anomaly coefficient is $A = 2N_c/3$, so the violation of the anomalous $U_{3A}(1)$ symmetry in the Standard Model is $\delta \mathcal{L}_{SM} = \omega \partial_\mu J^\mu_{3A} = \frac{\omega A}{64\pi^2} e^2 \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma}$.

For $N_c = 3$, i.e. 3 colours, this gives $\delta \mathcal{L}_{SM} = \frac{e^2}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} \omega$ which precisely reproduces the $U_{3A}(1)$ violation coming from the pion-photon interaction $\mathcal{L}_{\pi^0 AA} = \frac{e^2}{32\pi^2 F_\pi} \pi^0 \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma}$ under $\delta\pi^0 = \omega F_\pi$.

Now consider the dangerous potential anomalies in gauge currents. For consistency of the Standard Model, it is essential that such anomalies cancel. One needs to check the anomaly coefficients $A(3,3,3)$, $A(3,3,2)$, $A(3,3,1)$, $A(3,X,Y)$ $(X,Y = 2 \text{ or } 1)$, $A(2,2,2)$, $A(2,2,1)$, $A(1,1,1)$ where $A(3,3,3)$ denotes all the anomalies involving three $SU_c(3)$ generators, etc. $A(3,3,3)$ is straightforward, since the $SU_c(3)$ representations are all left-right symmetric, ie. the left-handed Q_{LmA} and right-handed u_{RmA} and d_{RmA} all carry fundamental $\underline{3}$ representations of $SU_c(3)$. $A(3,3,2)$ vanish because they are all proportional to the trace of the $SU_L(2)$ generators $\frac{1}{2}\sigma^a$, which are all traceless. $A(3,3,1)$ are a bit more involved. The $SU_c(3)$ generators are $T_\theta = \frac{1}{2}\lambda_\theta$, where the λ_θ are the Gell-Mann matrices, which satisfy $\{\lambda_\alpha, \lambda_\beta\} = \frac{4}{3}\delta_{\alpha\beta} + 2 d_{\alpha\beta\gamma}\lambda_\gamma$. The $SU_c(3)$ generators λ_γ are traceless, so the $d_{\alpha\beta\gamma}\lambda_\gamma$ part gives zero, while the trace over colours of the other part gives $4\delta_{\alpha\beta}$. Consequently, the $A(3,3,1)$ coefficient is proportional to the trace over all left-handed coloured fields of the $U_Y(1)$ hypercharge

$$\sum_{\text{quarks}} Y = 3(2y_{QL} + y_{u_R^c} + y_{d_R^c}) = 3\left[2(1/6) + (-2/3) + (1/3)\right] = 0$$

where the overall factor of 3 is the number of generations and the 2 on y_{QL} comes from the two $SU_L(2)$ flavours.

$A(3,X,Y)$ with $(X,Y = 2 \text{ or } 1)$ vanish because $\operatorname{tr}\lambda_\theta = 0$ for $SU_c(3)$.

For $A(2,2,2)$ recall that the only nontrivial $SU_L(2)$ irreps in the Standard Model are doublets, generated by $T_a = \frac{1}{2}\sigma_a$. The Pauli matrices all satisfy $\sigma_a^* = -\sigma_2 \sigma_a \sigma_2$, ie. this representation is pseudoreal, so $A(2,2,2) = 0$. For $A(2,2,1)$, note that the Pauli matrices satisfy $\{\frac{1}{2}\sigma_a, \frac{1}{2}\sigma_b\} = \frac{1}{2}\delta_{ab}$, and the factor of $\frac{1}{2}$ is cancelled by summing over doublets. So

L_L: left-handed lepton doublet

$$A(2,2,1) = \sum_{\text{doublets}} Y = 3(3y_{Q_L} + y_{L_L}) = 3(3(\tfrac{1}{6}) + (-\tfrac{1}{2})) = 0$$

where the factor of 3 on y_{Q_L} comes from colours. $A(2,1,1)$ involves the trace of a single Pauli matrix, which vanishes.

Finally, $A(1,1,1)$ is proportional to the sum over all left-handed fermions of the cube of the hypercharge

$$A(1,1,1) = 2\sum_{\text{all fermions}} Y^3 = 6(2y_{L_L}^3 + Y_{e_{RC}}^3 + 6y_{Q_L}^3 + 3y_{u_{RC}}^3 + 3y_{d_{RC}}^3)$$
$$= 6(2(-\tfrac{1}{2})^3 + (+1)^3 + 6(\tfrac{1}{6})^3 + 3(-\tfrac{2}{3})^3 + 3(\tfrac{1}{3})^3)$$
$$= 6(-\tfrac{1}{4} + 1 + \tfrac{1}{36} - \tfrac{8}{9} + \tfrac{1}{9}) = 0.$$

This rather non-trivial anomaly cancellation hinges upon a number of physical issues. The equality of proton and electron charge absolute values depends on having $Y_{e_{RC}} = 1$, while the neutrality of the neutron requires $Y_{d_{RC}} = \frac{1}{3}$. Moreover, if $Y_{e_{RC}}$ were different from 1, this would give rise to a shift in the identification of the electric charge generator, and this shift would then disturb the electric charge cancellation for the neutrino. So gauge anomaly cancellation is quite essential for the consistency and physical correctness of the Standard Model.

The QCD Θ parameter

In discussing the adjustments to the generic form of the Standard Model action needed to make the quark mass matrix real, positive and diagonal, we have used chiral transformations that are exactly of the form of the $SU_L(N) \times SU_R(N)$ effective low-energy theory's rigid approximate symmetries. So we need to be concerned about the effects of anomalies on these mass-matrix-arranging transformations.

The impact of such fermionic field redefinitions is most clearly seen from Fujikawa's approach to anomalies. For a rigid chiral transformation of a fermionic field $\psi \to e^{i\omega\gamma_5}\psi$, the classical action may be invariant but the path integral measure is not:

$$\int \mathcal{D}\psi\, \mathcal{D}\bar{\psi} \longrightarrow \int \mathcal{D}\psi\, \mathcal{D}\bar{\psi} \exp\left(-i\omega \frac{g^2}{32\pi^2} \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu}^a F_{\alpha\beta}^a\right)$$

where the $F_{\mu\nu}^a$ are Yang-Mills field strengths for any gauge symmetries under which the fermionic field ψ is charged and g is the corresponding charge. Note how the coefficient of ω agrees with the formula $A_{abc} = \mathrm{tr}(T_a\{T_b T_c\})$, e.g. for colour $SU_c(3)$. This is a mixed anomaly involving a rigid axial $U_A(1)$ with generator $T = 1\cdot\gamma_5$ and two $SU_c(3)$ generators $T_b = \frac{1}{2}\lambda_b$, $T_c = \frac{1}{2}\lambda_c$. Using again $\{\lambda_b, \lambda_c\} = \frac{4}{3}\delta_{bc} + 2 d_{bce}\lambda_e$ and noting that $\mathrm{tr}(\lambda_e) = 0$, one has $A_{Tbc} = \mathrm{tr}(T)\frac{\delta_{bc}}{3}\cdot 3 = \delta_{bc}\mathrm{tr}(T)$ (remembering the 3 colours of ψ). Then for $\mathrm{tr}(T)$ note that $\psi_L \to e^{i\omega}\psi_L$, $\psi_R \to e^{-i\omega}\psi_R$ so $(\psi_R)_c \to e^{i\omega}(\psi_R)_c$ so $\mathrm{tr}(T) = 2$ and one gets $\partial_\mu J^\mu_T = \frac{2}{64\pi^2} g^2 \epsilon^{\mu\nu\alpha\beta} F^b_{\mu\nu} F^b_{\alpha\beta}$. Recall that for a spacetime dependent parameter $w(x)$ in the transformation $\psi \to e^{i\omega\gamma_5}\psi$, one has $\delta\mathcal{L} = j^\mu_5 \partial_\mu w$, which one can rewrite as $-\partial_\mu j^\mu_5 w$. So, subsequently taking $w = \frac{\theta}{2}$ spacetime independent, one finds agreement with the anomalous variation of the path integral integrand $e^{i\int d^4x \mathcal{L}} \to e^{i\int d^4x(\mathcal{L} - \partial_\mu j^\mu_5 \frac{\theta}{2})}$.

For chiral transformations rotating between multiple generations, $\psi^i_L \to L^{ij}\psi^j_L$, $\psi^i_R \to R^{ij}\psi^j_R$, the angle θ is given by $\det(R^\dagger L) = r e^{i\theta}$ where $r \in \mathbb{R}$, so $\theta = \arg\det(R^\dagger L)$. From this, one sees that for a non-chiral transformation for which $L^{ij} = R^{ij}$, one has $\theta = 0$, corresponding to the absence of an anomaly.

above single ψ trans: $L = e^{i\omega}$ $R = e^{-i\omega}$

The inclusion of a term $\mathcal{L}_\theta = \frac{-\theta g^2}{64\pi^2}\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu}^a F_{\rho\sigma}^a$ in the Lagrangian at the classical level would have no effect on physical phenomena, because $\frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu}^a F_{\rho\sigma}^a = \partial_\mu K^\mu$ is a total derivative, where $K^\mu = 2\epsilon^{\mu\nu\rho\sigma}(A_\nu^a \partial_\rho A_\sigma^a + \frac{1}{3} f^{abc} A_\nu^a A_\rho^b A_\sigma^c)$ is the Chern-Simons current. This total derivative structure clearly is what is required by the $-\frac{\theta}{2}\partial_\mu j^\mu_5$ change in the Lagrangian.

$\tilde F^{a\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F^a_{\rho\sigma}$

Θ angle terms $\mathcal{L}_\Theta = \frac{-\Theta g^2}{32\pi^2} F^a_{\mu\nu} \tilde F^{a\mu\nu}$ are invariant under charge conjugation C because they are quadratic in $F^a_{\mu\nu}$, but they violate P and CP, and consequently violate T. The key test for non-invariance under P is the presence of a single $\epsilon^{\mu\nu\rho\sigma}$ tensor, which has the effect that in each term in the sum $F^a_{\mu\nu} \tilde F^{a\mu\nu}$ there is one sign flip, either because $\partial_i \to -\partial_i$ or because A^a_μ flips sign (recall that under C, gauge fields have sign flips or not according to the symmetry/antisymmetry of T^a, while under P $A^a_i(x) \to -A^a_i(x^0, -x^i)$)

under C, $F^a_{\mu\nu}$ transforms like A^a_μ

Once one has admitted the possibility of CP non-invariant total derivative terms in the SM Lagrangian, there are three possible Θ angle terms:

$$\mathcal{L}_{theta} = \frac{-g_3^2\Theta_3}{32\pi^2} G^\alpha_{\mu\nu} \tilde G^{\alpha\mu\nu} - \frac{g_2^2\Theta_2}{32\pi^2} W^i_{\mu\nu} \tilde W^{i\mu\nu} - \frac{g_1^2\Theta_1}{32\pi^2} B_{\mu\nu}\tilde B^{\mu\nu}.$$

However, not all of these have observable consequences, at least in the original Standard Model without neutrino masses. Precisely because there are anomalous variations of the fermionic path integral measure under rigid chiral transformations, there is ambiguity in where the effects of theta-angle terms will be encountered. Revisit the transformations made in order to diagonalize the u and d quark mass matrices $D^{(u)}, D^{(d)}$. We wrote the original Yukawa coupling matrices

$Y^{(u)}_{mn} \to k_{mn}$

$Y^{d}_{mn} \to h_{mn}$

$Y^{(u)} = U^{u_L\dagger} D^{(u)} U^{u_R}$, $Y^{(d)} = U^{d_L\dagger} D^{(d)} U^{d_R}$, but we could just as well have written $Y^{(u)} = U^{u_L\dagger} D^{(u)} U^{u_L} U'^{u_R}$, $Y^{(d)} = U^{d_L\dagger} D^{(d)} U^{d_L} U'^{d_R}$, so the quark mass matrices can be diagonalized by first making U'^{u_R} and U'^{d_R} chiral rotations on the RH quarks only, which does change Θ_3 and Θ_1 since the RH quarks couple to $SU_c(3)$ and $U_Y(1)$, followed by non-chiral rotations ($\tilde u_R = U'^{u_R} u_R$; $\tilde d_R = U'^{d_R} d_R$; $\tilde u_L = u_L$, $\tilde d_L = d_L$ then $\tilde{\tilde u}_R = U^{u_L} \tilde u_R$, $\tilde{\tilde u}_L = U^{u_L} \tilde u_L$; $\tilde{\tilde d}_R = U^{d_L} \tilde d_R$, $\tilde{\tilde d}_L = U^{d_L} \tilde d_L$), the non-chiral rotations changing none of the Θ angles.

non-chiral transfs: no anomaly

The phase induced by the U'^{uR} and U'^{dR} rotations that changes Θ_3 is $\arg\det(U'^{uR}U'^{dR}) = -\arg(\det(D^{(u)}D^{(d)})/\det(Y^{(u)}Y^{(d)}))$ $= \arg\det(Y^{(u)}Y^{(d)})$ since $D^{(u)}$ and $D^{(d)}$ are diagonal, real matrices. Suppose, on the other hand, one wants instead to eliminate Θ_3. That would be achieved by an inverse set of transformations $U'^{uR\dagger}$, $U'^{dR\dagger}$, but these would then restore a phase $-\arg\det(U'^{uR\dagger}U'^{dR\dagger}) = \arg\det(U'^{uR}U'^{dR})$ into $\arg\det(Y^{(u)}Y^{(d)})$ for the Yukawa terms, thus making the quark mass matrix complex. Moving a phase back and forth in this way is just a matter of basis choice, and cannot affect physical quantities (eg. S-matrix), which are invariant under such changes of variable. The basis-independent quantity is $\bar{\Theta}_3 = \Theta_3 + \arg\det(Y^{(u)}Y^{(d)})$.

For the $SU_L(2)$ and $U_Y(1)$ Θ angles, the situation is different. Since the RH quarks do not couple to $SU_L(2)$, Θ_2 is unaffected by the U'^{uR}, U'^{dR} transformations, but it can be removed by transformations of the LH quarks. Consequently, one may make LH quark transformations to remove Θ_2, thus moving its effect into the Yukawa couplings, but then make compensating transformations on the RH quarks to remove those effects without disturbing the now zero Θ_2. Similarly, for Θ_1, one can make LH neutrino transformations to set it to zero and since there are no Yukawa couplings to the LH neutrinos after symmetry breaking, one doesn't have to worry about Θ_1 moving into the Yukawa terms. So only the strong interaction $\bar{\Theta}_3$ is of physical relevance in the minimal Standard Model. (Matters can be different after inclusion of neutrino masses, but the resulting physical effects would be very small.)

Now consider how a total derivative term such as $-\frac{g_3^2\Theta_3}{32\pi^2}G^a_{\mu\nu}\tilde{G}^{a\mu\nu}$ can have physical effects. Quantum mechanical tunneling processes are calculated using solutions i,j contractions on same line because this is now Euclidean

to Euclideanized field theory, obtained after a Wick rotation $t = -i\tau$ of the time coordinate. After this Wick rotation, the $e^{i\int d^4x\, \mathcal{L}}$ path-integral phase becomes e^{-S_E}, where S_E is the Euclidean action. For Yang-Mills theories, $S_E = \frac{1}{4}\int d^dx\, F^a_{ij}F^a_{ij}$ (d-dimensional), $i,j = 1 \dots d$. Solutions to the Euclidean field equations derived from S_E extremality can contribute to tunneling amplitudes if S_E is finite when evaluated at such solutions. Finite Euclidean action S_E requires $F^a_{ij} = 0$ as one goes to infinity ($|x| \to \infty$), but F^a_{ij} can be nonvanishing provided the Euclidean space is $d = 4$ dimensional.

The restriction to $d = 4$ dimensions comes about from a scaling argument. In order for $S_E[A]$ to be extremal, the field equations $\nabla_i F^a_{ij} = 0$ must be satisfied. Moreover, these are preserved by transforming $A^a_i(x) \to A^{aR}_i(x) = R^{-1} A^a_i(x/R)$ (under which $F^a_{ij}(A) \to R^{-2} F^a_{ij}(A(y))\big|_{y = x/R} = F^{aR}_{ij}(x)$ and $\nabla_i = \frac{\partial}{\partial x_i} - ig A_i \cdot T \to R^{-1}\nabla_i$. In other words, the Yang-Mills equations are scale covariant. But $S_E[A^R] = R^{d-4} S_E[A]$ so for $d \neq 4$, there can be no non-trivial extrema of $S_E[A]$ unless $S[A] = 0$. But if $S_E[A] = 0$, then $F^a_{ij} = 0$ everywhere. By a gauge transformation, one can then make $A_i = 0$ everywhere as well.

For $d = 4$, however, one can find "instanton" solutions with F^a_{ij} non-vanishing except at infinity. In order to have finite $S_E[A]$, $A^a_i(x)$ can vanish as slowly as $1/|x|$ provided the field approaches a pure gauge configuration as $|x| \to \infty$: $it^a A^a_i(x) \to g^{-1}(\hat{x})\, \partial_i g(\hat{x})$ where $g(\hat{x})$ is a direction-dependent element of the gauge group G. Moreover, $A^a_i(x)$ is unaffected if one replaces $g(\hat{x})$ with $g_0 g(\hat{x})$ for any fixed $g_0 \in G$. Then by choosing $g_0 = g^{-1}(\hat{x}_1)$ for some specific direction $\hat{x}_1$, one arranges that $g(\hat{x}_1) = \mathbb{1}$ (the identity element of G). Each gauge field with finite $S_E[A]$ thus defines a mapping from the Euclidean unit sphere $|\hat{x}| = 1$ to the group manifold of G, with $\hat{x}_1$ mapped to the identity element of G.

The set of classes of such topologically distinct mappings $S_{d-1} \to G$ with one point of S_{d-1} mapped onto a fixed element of G is $\pi_{d-1}(G)$, the $(d-1)$th homotopy group of the group manifold. For $d=4$, one is therefore interested in $\pi_3(G)$; this homotopy group is non-trivial for any semisimple Lie group G.

Using the fact that $0 \le \int (F^a_{ij} \mp \frac{1}{2}\epsilon^E_{ijkl} F^a_{kl})(F^a_{ij} \mp \frac{1}{2}\epsilon_{ijmn} F^a_{mn}) d^4x$ where ϵ^E_{ijkl} is the Euclidean ϵ-tensor, $\epsilon^E_{1234} = 1$, and noting that for $\tilde{F}^a_{ij} = \frac{1}{2}\epsilon_{ijkl} F^a_{kl}$ one has $\tilde{\tilde{F}}^a_{ij} = F_{ij}$, it follows that

$$0 \le 2\int F^a_{ij} F^a_{ij} \mp 2\int F^a_{ij} \tilde{F}^a_{ij} \quad \text{so} \quad S_E \ge \frac{1}{2}\left|\int F^a_{ij} \tilde{F}^a_{ij}\right| .$$ Bogomolnyi inequality

The lower bound is thus reached if and only if F^a_{ij} is either self-dual or anti-self-dual: $F^a_{ij} = \pm \frac{1}{2}\epsilon_{ijkl} F^a_{kl}$. Note that such self-dual or anti-self-dual field strengths automatically satisfy the $\nabla_i F^a_{ij} = 0$ field equation because $\nabla_i \tilde{F}^a_{ij} \equiv 0$.

The integrated quantity $\nu = \frac{g^2}{32\pi^2}\int d^4x\, F^a_{ij} \tilde{F}^a_{ij}$ is a topological invariant known as the winding number of the Euclidean field configuration, the integral of the Chern-Pontryagin density. The winding number so expressed is an integer $\nu \in \mathbb{Z}$. Often, one sees it given in terms of differently scaled field strengths $\hat{F}^a_{ij} = gF^a_{ij} = \partial_i \hat{A}^a_j - \partial_j \hat{A}^a_i + f^{abc}\hat{A}^b_i \hat{A}^c_j$ with $\hat{A}_i = gA_i$, in terms of which $\nu = \frac{1}{32\pi^2}\int d^4x\, \hat{F}^a_{ij} \hat{\tilde{F}}^a_{ij}$.

Solutions with $F^a_{ij} = \pm \tilde{F}^a_{ij}$, which extremize the Bogomolnyi inequality $S_E \ge \frac{1}{2}|F^a_{ij}\tilde{F}^a_{ij}|$ were found by Belavin, Polyakov, Schwarz and Tyupkin (Phys. Lett. 59B (1975), 89):

$$iA_i(x) = \left(\frac{r^2}{r^2 + R^2}\right) g_1^{-1}(\hat{x})\, \partial_i\, g_1(\hat{x}) \qquad r = \sqrt{x^i x^i} \quad i = 1, \dots 4$$

where R is an arbitrary scale factor and $g_1(\hat{x})$ is an element of an $SU(2)$ subgroup of the gauge group G with

$$g_1(\hat{x}) = \left(\frac{x^4 + 2i\, x \cdot \tau}{r}\right), \quad \tau^a = \frac{1}{2}\sigma^a,\ a = 1,2,3 \ SU(2) \text{ generators}.$$

This is the original instanton solution, and has winding number $\nu = 1$. It is centered at $x^i = 0$ in the Euclidean $\mathbb{R}^4$, hence the "instanton" name. The Euclidean action for this:

solution is $S_E[A] = \frac{8\pi^2}{g^2}$, corresponding to $\int d^4x\, F^a_{ij}\tilde{F}^a_{ij} = 32\pi^2\nu$ with $\nu = 1$.

Having found an instanton solution with winding number $\nu = 1$, it is then clear that there exist solutions with any $\nu \in \mathbb{Z}$. For instance, solutions with $\nu = N$ can be approximately constructed by superposing N solutions with $\nu = 1$ taken with centers far apart, so the nonlinearities of the YM field equations can be ignored. A solution with $\nu = -1$ can be obtained by replacing $g_1(\hat{x})$ with $g_1^{-1}(\hat{x})$; this then leads to $\nu = -N$ solutions by superposition.

Asymptotically as $|x| \to \infty$, the instanton solutions approach pure gauge configurations (e.g. $iA_i \to g_1^{-1}(\hat{x})\partial_i g_1(\hat{x})$ for the $\nu = 1$ BPST solution), and consequently $F^a_{ij} \to 0$ as $|x| \to \infty$.

temporal gauge: A_i becomes x^4 independent as $|x^4| \to \infty$

Looking back to the Minkowski space theory, in temporal gauge with $A^a_4 = 0$, one can think of an instanton as giving a transition between an $F^a_{ij} = 0$ solution as $t \to -\infty$ and a different $F^a_{ij} = 0$ solution as $t \to +\infty$; in other words a "vacuum to vacuum" tunnelling. In the quantum theory, one needs to sum over all histories in the path integral and this will thus necessarily include the effects of all $\nu \in \mathbb{Z}$ instantons. The weighting of these various contributions can be shown on general field theoretic grounds (especially cluster decomposition) to be of the form $e^{i\theta\nu}$, which is thus equivalent to the inclusion of $\mathcal{L}_\theta = -\frac{\theta g^2}{32\pi^2} F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$ since upon returning to Minkowski space one has $(d^4x)_E = i(d^4x)_M$, $\tilde{F}^a_{k4} = -i\tilde{F}^a_{k0}$ $(k = 1,2,3)$ and $\epsilon^{1230} = -1$ so $\nu = -\frac{g^2}{32\pi^2}\int d^4x\, F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$ in the Minkowski space theory.

For a winding number ν instanton, the Euclidean action is $S_E[A] = \frac{8\pi^2|\nu|}{g^2}$, so the path integral is weighted by

restoring $\hbar$

$\exp[-\frac{1}{\hbar}S_E] = \exp[-\frac{8\pi^2|\nu|}{g^2\hbar}]$. As $g \to 0$, this exponential vanishes strongly enough to kill any powers of g that occur in perturbative corrections: $g^{-n}\exp\left[\frac{-8\pi|\nu|}{g^2 \hbar}\right]$ and all its derivatives with respect to g vanish at $g=0$. Such contributions in instanton backgrounds are accordingly highly non-perturbative: they are never encountered in any finite order in perturbation theory.

Despite the non-perturbative nature of the instanton contributions, one can obtain information about their effects using the chiral nonlinear sigma model effective theory abstracted from QCD. We had $U(x) = \exp(i\pi^a\tau^a/F_\pi)$ where $\tau^a = \frac{1}{2}\sigma^a$, and π^a are the pion fields. Including the mass term, the chiral effective theory Lagrangian we took to be

$$\mathcal{L}_{\text{chiral, m}} = -\frac{F_\pi^2}{4}\operatorname{tr}\left[(D_\mu U)(D^\mu U)^\dagger\right] + \frac{V^3}{2}\operatorname{tr}[MU + M^\dagger U^\dagger]$$

where $V^3 = \langle\bar{u}u\rangle = \langle\bar{d}d\rangle$ and M is the QCD quark mass matrix. We looked earlier at expanding the mass term and comparing the vacuum energy arising from it to that arising from the QCD level quark mass terms in the presence of the $\langle\bar{u}u\rangle = \langle\bar{d}d\rangle$ condensates. This led to the Gell-Mann-Oakes-Renner relation $F_\pi^2 m_\pi^2 = V^3(m_u+m_d)$. Now consider the impact of the invariant $\bar{\Theta}_3$ parameter. One may make chiral transformations on the u_R and d_R quarks to remove the $\bar{\Theta}_3$ coefficient from the $F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$ term in the QCD Lagrangian, at the price of having its effects appear in the Yukawa couplings, i.e. in $\arg\det(y^{(u)}y^{(d)})$. After electroweak symmetry breaking, this leads to complex quark masses: one has

$\arg\det(y^{(u)}y^{(d)}) = \arg\det M$

$$M = \begin{pmatrix} m_u & 0 \\ 0 & m_d\end{pmatrix} e^{\frac{i}{2}\bar{\Theta}_3}$$ for the u,d quark mass matrix.

This has the effect of making the vacuum energy depend on $\bar{\Theta}_3$:

$E(\bar{\Theta}_3) = V^3(m_u+m_d)\cos(\frac{1}{2}\bar{\Theta}_3) = F_\pi^2 m_\pi^2 \cos(\frac{1}{2}\bar{\Theta}_3)$ using the Gell-Mann-Oakes-Renner relation. This indicates that the parameter $\bar{\Theta}_3$ controls the energies of distinct vacuum states, called Θ-vacua.

for N flavours, $\frac{1}{2}\bar{\Theta}_3 \to \frac{1}{N}\bar{\Theta}_3$

Another consequence of the QCD $\bar{\Theta}_3$ parameter is that the neutron acquires an electric dipole moment proportional to $\bar{\Theta}_3$. Then from the experimental bound on the neutron EDM, one obtains a bound $\bar{\Theta} < 5\cdot 10^{-10}$. Why so small?

The Strong CP problem concerns the extreme smallness of $\bar{\Theta}_3$ despite the relatively large amount of CP violation arising in the weak interactions of the Standard Model.

One conceivable solution would be that one of the quarks is massless. Then one could eliminate Θ_3 in a fashion similar to the elimination of Θ_1 by a chiral rotation without effect on the quark masses. The chief candidate for masslessness might be the u quark. There are several problems with this suggestion, however. One is that there is no symmetry protecting $m_u = 0$ alone, so m_u would just have to be tuned to be small instead of tuning $\bar{\Theta}_3$ to be small. So proposing that $m_u = 0$ just moves the fine-tuning problem around. Another problem is that it conflicts with lattice gauge theory calculations, which give a non-zero value for m_u with a large statistical significance.

eg. setting to zero a Yukawa coef.

The most favorably discussed idea for solving the Strong CP problem is the introduction of a hypothetical particle field called the axion $a(x)$. This field is assumed to be a (pseudo) Goldstone boson for an approximate axial symmetry involving at least one of the quarks. If this symmetry were exact, the axion would have a shift symmetry $a(x) \rightarrow a(x) + \epsilon$, for constant parameter ϵ. Such a symmetry is called a Peccei-Quinn symmetry after its original authors. The idea has subsequently been developed by Weinberg and Wilczek. Suppose that the axion has an effective Lagrangian

$$\mathcal{L}_{axion} = -\frac{1}{2}\partial_\mu a \partial^\mu a + \frac{(a/f_a - \bar{\Theta}_3 g_3^2)}{32\pi^2} G^{\theta}_{\mu\nu}\tilde{G}^{\theta\mu\nu}$$

dimension 5 coupling similar to π^0 coupling to $F_{\mu\nu}\tilde{F}^{\mu\nu}$

f_a = axion decay constant

where the axion shift symmetry is anomalous. Let the constant vacuum value of $a(x)$ be $\bar{a}$. Then $\left(\frac{\bar{a}}{g_3^2 f_a} - \bar{\Theta}_3\right)$ can be eliminated by making an axial rotation of the quarks. This produces a phase $\exp\left(\frac{i}{2}\left(\bar{\Theta}_3 - \frac{\bar{a}}{g_3^2 f_a}\right)\right)M$ in the mass matrix $M = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix}$.

Consequently, the energy of the vacuum is

$$E(\bar{\Theta}, \bar{a}) = V^3 (m_u + m_d) \cos\left(\frac{1}{2}\left(\bar{\Theta}_3 - \frac{\bar{a}}{g_3^2 f_a}\right)\right)$$

$$= F_\pi^2 m_\pi^2 \cos\left(\frac{1}{2}\left(\bar{\Theta}_3 - \frac{\bar{a}}{g_3^2 f_a}\right)\right)$$ using Gell-Mann-Oakes-Renner

So the minimum of vacuum energy occurs for $\bar{a} = g_3^2 f_a \bar{\Theta}_3$, at which value the phase of the quark mass matrix vanishes, and the strong CP violating effects are cancelled out.

Proposing a new particle species in this way to cause the CP violating $\bar{\Theta}_3$ angle to relax and cancel out the Strong-interaction CP violation is not cost-free, however. As with every extension of a successful theory, one has to ensure that new physical effects arising from the extension are not in conflict with observation. The $(\text{mass})^2$ of the axion may be obtained from $\frac{\partial^2}{\partial \bar{a}^2} E(\bar{\Theta}, \bar{a})\Big|_{\bar{a} = g_3^2 f_a \bar{\Theta}_3} = \frac{F_\pi^2 m_\pi^2}{4 g_3^4 f_a^2} = m_{ax}^2$

So $m_{ax} = \frac{F_\pi m_\pi}{2 g_3^2 f_a}$, inversely proportional to f_a.

If $f_a \sim$ TeV, this yields a mass $m_{ax} \sim$ keV, while if $f_a \sim 10^{16}$ GeV, it yields a mass $m_{ax} \sim 10^{-9}$ eV.

Astrophysical bounds, particularly from red giants, require $f_a > 10^{10}$ GeV, while cosmological bounds require $f_a < 10^{12}$ GeV (too many axions would overclose the universe). Consequently, the axion should be very weakly coupled with a mass in the range $10^{-4}\,\text{eV} < m_{ax} < 10^{-2}\,\text{eV}$.

The possibility that the axion couples not only to the Strong interaction $G^a_{\mu\nu}\tilde{G}^{a\,\mu\nu}$ but also to $F_{\mu\nu}\tilde{F}^{\mu\nu}$ in QED means that in a strong electromagnetic field an axion could convert into a photon. This is the basis for a search by the ADMX experiment at the University of Washington, predicated on the supposition that our galaxy has a <u>dark matter</u> halo composed largely of axions.