Lecture 15: THE STANDARD MODEL OF FUNDAMENTAL INTERACTIONS

Having discussed in chapter 14 the structure and the formal properties of non-Abelian gauge theories, we now apply the formalism to the description of some general properties of the Standard Model (SM) of Weak, Electromagnetic and Strong Interactions.

The weak–electromagnetic theory with $SU(2) \times U(1)$ spontaneously broken gauge symmetry and three quark and lepton generations, thanks to the smallness of the coupling constant at low energy, has been tested quite systematically and when radiative corrections are taken into account, provides up to now a precise description of all collider experiments.

With the discovery in 2012 of the Higgs particle at the new Large Hadron Collider at CERN with a mass of 125 Gev, all particles of the initial Standard Model have now been identified and all parameters determined.

Important remains a precise experimental verification of the couplings of the Higgs particle to other particles since, in the SM, these couplings determine the masses of all fundamental particles.

A disagreement would be a sign of new physics. Moreover, the mass spectrum of the fundamental particles remains one of the mysterious SM features.

In recent years a series of experiments have demonstrated the existence of oscillations between the three neutrino species $\nu_\mathrm{e}, \nu_\mu, \nu_\tau$. This implies neutrino masses and thus a modification of the SM.

We assume here a minimal modification in the form of Dirac neutrinos. With such an assumption, consistent up to now with experimental data, we exclude Majorana-type mass terms and thus lepton-number violating neutrinoless β decay.

The lepton and quark sectors then have a similar structure. In particular, the lepton sector now involves a mixing matrix in analogy the quark sector (three angles have been determined, the fourth CP violating angle is still unknown)

The situation for the Strong Interaction sector, is somewhat different. It is described by Quantum Chromodynamics (QCD), a theory of quarks interacting through gluons associated to gauge fields corresponding to a $SU(3)$ colour gauge group.

The QCD RG β-function, calculated at leading order in perturbation theory, exhibits the property of large momentum asymptotic freedom (see section 15.4.2). Thus, a number of high-energy properties, like the so-called inclusive cross-sections, can be predicted.

However, low energy properties like $SU(3)$ colour confinement, cannot be derived from perturbation theory, the effective interaction becoming too large. Evidence for the validity of the quark confinement scheme thus relies on increasingly precise, non-perturbative, numerical investigations of lattice gauge theories (see chapter 16).

In some cases, when gauge fields are coupled to axial currents, the WT identities that are necessary to prove the consistency of gauge theories are not satisfied beyond the tree approximation. They are spoiled by anomalies. Therefore, the second part of the chapter is devoted to the discussion of this important problem. Results are illustrated by some physics consequences like, the cancellation of anomalies in the Standard Model that requires three-colour quarks, the π_0 decay and the solution to the $U(1)$ problem.

A general remark is that, though the Standard Model depends on many parameters, its structure is tightly constrained and satisfies a number of non-trivial conditions.

15.1 Weak–electromagnetic interactions: gauge and Higgs fields

We first describe the Standard Model of weak and electromagnetic interactions (suitably modified to accommodate Dirac neutrinos), which provides a physics application of the non-Abelian Higgs mechanism.

We mainly restrict the presentation to one generation and two flavours, eventually indicating how it generalizes to three generations (with essential new features like the CKM and PMNS mixing matrices and CP violating angles).

The gauge group of this sector of the Standard Model is $SU(2) \times U(1)$. The action for the Higgs field sector is related to the action (14.44) considered in section 14.5.3 by gauging a $U(1)$ subgroup of the remaining $SU(2)$ non-gauge symmetry. The pattern of symmetry breaking is similar.

A $U(1)$ gauge symmetry remains unbroken, whose generator is a linear combination of the original $U(1)$ generator and one of the $SU(2)$ generators.

It is associated with electromagnetic interactions.

Since the gauge group is a product of two groups, the model depends on two independent gauge couplings and, therefore, weak and electromagnetic interactions are combined rather than completely unified.

Note that since in the SM, the left-handed and right-handed components of the fermion fields are treated differently, the breaking of parity symmetry is explicit.

Detailed phenomenological applications can be found in the literature.

Notation. We use the conventions of section 14.5.3, in particular, for the $SU(2)$ transformations and the τ Pauli matrices.

15.1.1 Scalar and gauge fields

The scalar action. The scalar field is a two-component complex field, a $SU(2)$ doublet $\boldsymbol{\varphi}$. In the absence of gauge fields, the scalar action has the form

$$\mathcal{S}_{\text{scal.}} = \int \mathrm{d}^4 x \left[\sum_{\mu} \left| \partial_{\mu} \boldsymbol{\varphi}(x) \right|^2 + \mathcal{U}\left(|\boldsymbol{\varphi}(x)|^2 \right) \right] \tag{15.1}$$

with

$$\mathcal{U}(\rho) = u\rho + \tfrac{1}{6}\lambda\rho^2 .$$

In this form the action has an obvious $U(2) \sim SU(2) \times U(1)$ symmetry:

$$\boldsymbol{\varphi}(x) \mapsto \mathbf{U}\boldsymbol{\varphi}(x)\,, \;\; \mathbf{U} \in U(2).$$

Actually, rewriting the action (15.1) in terms of the four real components of $\boldsymbol{\varphi}$, one verifies that action has a larger $O(4)$ symmetry, but only the $SU(2) \times U(1)$ group is gauged.

The $\boldsymbol{\varphi}$-gauge invariant action. One introduces a three-component real gauge vector $\mathbf{A}_\mu$ for the $SU(2)$ component and a real gauge field B_μ for the Abelian $U(1)$ component.

The gauge invariant scalar field action can be written as

$$\mathcal{S}_{\text{scal.}} = \int \mathrm{d}^4x \left[\sum_\mu |\mathbf{D}_\mu \boldsymbol{\varphi}(x)|^2 + \mathcal{U}\left(|\boldsymbol{\varphi}(x)|^2\right) \right] \tag{15.2}$$

with

$$\mathbf{D}_\mu = \mathbf{1}\left(\partial_\mu + \tfrac{1}{2} i Y_\varphi g' B_\mu\right) + \tfrac{1}{2} i g \mathbf{A}_\mu \cdot \boldsymbol{\tau} \, .$$

We normalize g' in such a way that $\boldsymbol{\varphi}$ has a charge $Y_\varphi = 1$ with respect to the $U(1)$ group.

We normalize the curvature tensors by

$$[\mathbf{D}_\mu, \mathbf{D}_\nu] = \tfrac{1}{2} i g \mathbf{F}_{\mu\nu} + \tfrac{1}{2} i g' B_{\mu\nu}$$

with (in $SO(3)$ notation)

$$\mathbf{F}_{\mu\nu} = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu - g \mathbf{A}_\mu \times \mathbf{A}_\nu \,, \quad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \,.$$

The $SU(2) \times U(1)$ gauge field action, in this normalization, is the sum

$$\mathcal{S}(\mathbf{A}, \mathbf{B}) = \tfrac{1}{4} \int \mathrm{d}^4 x \sum_{\mu,\nu} \left[\mathbf{F}^2_{\mu\nu}(x) + B^2_{\mu\nu}(x) \right] .$$

15.1.2 The Higgs mechanism: classical approximation.

We assume that the $\boldsymbol{\varphi}$ potential is such that in the classical approximation the field $\boldsymbol{\varphi}$ has a non-zero expectation value (*i.e.*, $u < 0$), which we choose of the form,

$$\langle \boldsymbol{\varphi}(x) \rangle = \frac{v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; v > 0 \,.$$

With this form, the expectation value of $\boldsymbol{\varphi}$ is invariant under the transformations,

$$\boldsymbol{\varphi}(x) \longmapsto \mathrm{e}^{i\theta(1+\tau_3)/2}\,\boldsymbol{\varphi}(x)\,.$$

They correspond to a $U(1)$ subgroup whose generator is the sum of the generator of the initial $U(1)$ group and the τ_3 generator of $SU(2)$. Therefore, as anticipated, $SU(2)\times U(1)$ is broken down to $U(1)$.

Replacing $\boldsymbol{\varphi}$ by its expectation value in action (15.2), we read off the mass terms for the gauge fields in the classical approximation:

$$\tfrac{1}{8}v^2\int \mathrm{d}^4x \sum_{\mu}\left[\left(g'B_\mu - gA^{(3)}_\mu\right)^2 + g^2\left|A^{(1)}_\mu + iA^{(2)}_\mu\right|^2\right].$$

First, it follows that the linear combination $g'B_\mu - gA^{(3)}_\mu$ is massive while the orthogonal combination remains massless and thus represents the photon field. One defines the weak angle θ_W by

$$g'/g = \tan\theta_\mathrm{W}\,. \tag{15.3}$$

The photon field A_μ and the massive neutral field Z_μ then correspond to

$$A_\mu = \cos\theta_{\rm W} B_\mu + \sin\theta_{\rm W} A_\mu^{(3)}, \quad Z_\mu = -\sin\theta_{\rm W} B_\mu + \cos\theta_{\rm W} A_\mu^{(3)}. \tag{15.4}$$

Conversely,

$$B_\mu = \cos\theta_{\rm W} A_\mu - \sin\theta_{\rm W} Z_\mu\,, \quad A_\mu^{(3)} = \sin\theta_{\rm W} A_\mu + \cos\theta_{\rm W} Z_\mu\,. \tag{15.5}$$

The components $A_\mu^{(1,2)}$ are coupled to A_μ and correspond to charged vector fields, which are usually written in complex notation as

$$W_\mu^\pm = (A_\mu^{(1)} \pm i A_\mu^{(2)})/\sqrt{2}\,. \tag{15.6}$$

From the coupling of the charged vector bosons with the photon, one derives the relation between electric charge e and coupling constants g and g':

$$e = gg'/\sqrt{g^2 + g'^2} = g\sin\theta_{\rm W} = g'\cos\theta_{\rm W}\,. \tag{15.7}$$

The Z mass is then

$$m_Z = \tfrac{1}{2}v\sqrt{g^2 + g'^2} = \frac{gv}{2\cos\theta_\mathrm{W}} = \frac{ev}{\sin 2\theta_\mathrm{W}}\,.$$

The two charged boson vectors $W_\mu^\pm$ have the common mass,

$$m_W = \tfrac{1}{2}gv = \frac{ev}{2\sin\theta_\mathrm{W}} = m_Z\cos\theta_\mathrm{W}\,.$$

Finally, the mass of the remaining massive scalar particle, the Higgs boson, is (equation (14.46))

$$m_H = v\sqrt{\lambda/3}\,.$$

In 2014 the experimental values are

$$m_W = 80.385 \pm 0.015\,\mathrm{GeV}\,,\ \ m_Z = 91.1876 \pm 0.0021\,\mathrm{GeV}\,,$$
$$m_H = 125.7 \pm 0.4\,\mathrm{GeV}\,.$$

They lead to a weak angle such that $\sin^2\theta_\mathrm{W} = 0.223$, instead of the experimental value $\sin^2\theta_\mathrm{W} = 0.231$. The agreement improves when radiative corrections are taken into account and between different experimental determinations, like the one coming from neutral currents.

From the value of the fine structure constant $\alpha^{-1} = 137.036$ one also infers $v = 250.7$ GeV (see also equation (15.15)).

Finally, from the experimental mass of the Higgs particle, one infers $\lambda \approx 0.75$, which is rather small.

15.2 Leptons: minimal SM extension with Dirac neutrinos

We consider here only the electron and the corresponding neutrino ν_e, which we assume to be a Dirac fermion, since the gauge couplings of the leptons of the two other generations (μ, ν_μ), (τ, ν_τ) have exactly the same structure.

The left-handed parts of the neutrino and electron are combined into a left-handed doublet L of $SU(2)$:

$$L = \begin{pmatrix} (\nu_e)_{\mathrm{L}} \\ e_{\mathrm{L}} \end{pmatrix} \equiv \tfrac{1}{2}(1-\gamma_5)\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}.$$

The right-handed parts of the neutrino $R_1 = \frac{1}{2}(1+\gamma_5)\nu_e$ (an addition to the SM) and electron $R_2 = \frac{1}{2}(1+\gamma_5)e^-$ form $SU(2)$ singlets.

Note that, in contradistinction to our usual euclidean notation, we use the Minkowskian conventions,

$$\bar{L} = (\bar{\nu}_e\,, e^+)(1+\gamma_5)/2\,,\ \bar{R}_1 = \bar{\nu}_e(1-\gamma_5)/2\,,\ \bar{R}_2 = e^+(1-\gamma_5)/2\,.$$

15.2.1 Lepton–gauge action

The action for the two leptons coupled to gauge fields can then be written as

$$\mathcal{S}_{\nu,e} = -\int \mathrm{d}^4 x \left[\sum_i \bar{R}_i \left(\not\partial + \tfrac{1}{2} i Y_{i,\mathrm{R}} g' \not B \right) R_i + \bar{L} \left(\not\partial + \tfrac{1}{2} i Y_{\mathrm{L}} g' \not B + \tfrac{1}{2} i g \not\mathbf{A} \cdot \boldsymbol{\tau} \right) L \right]. \tag{15.8}$$

Using relations (15.5, 15.7) to replace B_μ and $A_\mu^{(3)}$ by their components on A_μ , one obtains the coupling terms to the photon field,

$$\tfrac{1}{2} i e \sum_i Y_{i,\mathrm{R}} \bar{R}_i \not A R_i + \tfrac{1}{2} i e \bar{L} (Y_{\mathrm{L}} + \tau_3) \not A L \,.$$

Denoting by T_3 the eigenvalue of τ_3 the generator of $SU(2)$, Y the $U(1)$ charge and Q the electric charge, one obtains for each fermion the relation

$$Q = \tfrac{1}{2}(T_3 + Y). \tag{15.9}$$

Explicitly,

$$Q_1 = \tfrac{1}{2} Y_{1\mathrm{R}}, \quad Q_2 = \tfrac{1}{2} Y_{2\mathrm{R}}, \quad Q_1 = \tfrac{1}{2}(Y_{\mathrm{L}} + 1), \quad Q_2 = \tfrac{1}{2}(Y_{\mathrm{L}} - 1). \quad (15.10)$$

First, one notes the consistency condition

$$Q_1 - Q_2 = 1 \,, \quad (15.11)$$

consequence of the $SU(2)$ doublet assignment.

Then, all $U(1)$ charges are related to Q_1 by

$$Y_{1\mathrm{R}} = 2Q_1 \,, \; Y_{2\mathrm{R}} = 2(Q_1 - 1) \,, \; Y_{\mathrm{L}} = 2Q_1 - 1 \,. \quad (15.12)$$

Identifying Q_1 with the neutrino charge, $Q_1 = 0$, one finds $Q_2 = -1$ and the $U(1)$ charges,

$$Y_{1\mathrm{R}} = 0 \,, \quad Y_{2\mathrm{R}} = -2 \,, \quad Y_{\mathrm{L}} = -1 \,.$$

The value $Y_{1\mathrm{R}} = 0$ is consistent with the observation that ν_{R} does not participate in the weak and electromagnetic interactions.

15.2.2 Lepton masses

The $SU(2)$ symmetry forbids fermion mass terms. On the other hand, the coupling terms between the scalar field and the leptons,

$$\mathcal{S}_1 = iG_\nu \int \mathrm{d}^4x \left[\bar{R}_1(x)\big(\boldsymbol{\varphi}(x)\tau_2 L(x)\big) + \big(\bar{L}(x)\tau_2\boldsymbol{\varphi}^*(x)\big)R_1(x)\right], \quad (15.13a)$$

$$\mathcal{S}_2 = G_e \int \mathrm{d}^4x \left[\bar{R}_2(x)\big(\boldsymbol{\varphi}^*(x)\cdot L(x)\big) + \big(\bar{L}(x)\cdot\boldsymbol{\varphi}(x)\big)R_2(x)\right], \quad (15.13b)$$

are allowed by the $SU(2)\times U(1)$ symmetry (If $\mathbf{U}$ belongs to $SU(2)$, $\mathbf{U}^* = \tau_2\mathbf{U}\tau_2$), independently of the value of Q_1 provided $Y_\varphi = 1$.

Replacing the $\boldsymbol{\varphi}$-field by its expectation value, we see that the spontaneous breaking of the $SU(2)\times U(1)$ symmetry generates the masses

$$m_\nu = v|G_\nu| \quad m_e = v|G_e|\,.$$

Therefore, the fermion masses are calculable, but, in the absence of a new dynamic principle, in terms of arbitrary parameters, the Yukawa coupling constants G_e, G_ν.

In the SM, the Yukawa couplings of leptons are proportional to their mass. In particular, the perturbative approximation becomes worse for heavier leptons.

15.2.3 The Fermi constant

The coupling constant G_{F} (the Fermi constant), characteristic of the strength of weak interactions, is defined in terms of an effective low energy current–current and thus four-fermion interaction:

$$\frac{G_{\mathrm{F}}}{\sqrt{2}} \int \mathrm{d}^4 x \sum_{\mu} J_{\mu}(x) J_{\mu}^{\dagger}(x). \tag{15.14}$$

The contribution to the charged current J_{μ} coming from the electron and the neutrino has the form:

$$J_{\mu}(x) = \bar{e}(x)(1 - \gamma_5)\gamma_{\mu}\nu_{\mathrm{L}}(x) = 2\bar{e}_{\mathrm{L}}(x)\gamma_{\mu}\nu_{\mathrm{L}}(x).$$

The relation between G_{F} and the coupling constants g and g' is obtained by taking the large W-mass limit of the electron–neutrino scattering amplitude in the classical approximation.

The result can be obtained by integrating over the vector fields $A_\mu^{(1,2)}$, taking only into account the mass term and neglecting the kinetic part. The corresponding part of the action is

$$\tfrac{1}{8}g^2v^2\sum_\mu\left[\left(A_\mu^{(1)}\right)^2+\left(A_\mu^{(2)}\right)^2\right]+\tfrac{1}{2}ig\sum_\mu\bar{L}\gamma_\mu\left(A_\mu^{(1)}\tau_1+A_\mu^{(2)}\tau_2\right)L\,.$$

Completing squares one immediately obtains the result of the integration,

$$\frac{1}{2v^2}\sum_\mu\left[\left(\bar{L}\gamma_\mu\tau_1L\right)^2+\left(\bar{L}\gamma_\mu\tau_2L\right)^2\right]=\frac{2}{v^2}\sum_\mu\bar{\nu}_{\mathrm{L}}\gamma_\mu e_{\mathrm{L}}\bar{e}_{\mathrm{L}}\gamma_\mu\nu_{\mathrm{L}}\,.$$

Comparing with the definition (15.14), one concludes,

$$G_{\mathrm{F}}/\sqrt{2}=1/2v^2\,. \tag{15.15}$$

Inserting the experimental value $G_{\mathrm{F}} = 1.166\ldots \times 10^{-5}\,\mathrm{GeV}^{-2}$ one finds the value $v = 246.22$ GeV.

The effective low energy Fermi model of low energy charged weak interactions determines all parameters of the Standard Model (with massless neutrinos) but two, for example, the weak angle θ_{W} and the Higgs field self-coupling λ, which have to be inferred from additional experimental results. The direct measurements of the W and Z masses, for example, determine the parameter θ_{W}. At leading order the W masses can be rewritten as

$$M_W^2 = \frac{e^2}{4\sqrt{2}G_{\mathrm{F}}\sin^2\theta_{\mathrm{W}}},$$

that is,

$$M_W = \left(\frac{\pi\alpha}{\sqrt{2}G_{\mathrm{F}}}\right)^{1/2}\frac{1}{\sin\theta_W} = \frac{37.28}{\sin\theta_{\mathrm{W}}}\,\mathrm{GeV} \ \Rightarrow\ \sin^2\theta_W = 0.215\,.$$

The various estimates give an idea of the precision of leading order estimates.

The coupling of the charged vector bosons $W^{\pm}$ to (e^{-}, ν_e) is obtained by introducing the definition (15.6) into equation (15.8):

$$\begin{aligned}\tfrac{1}{2} g \bar{L} \gamma_\mu \left(\tau^1 A_\mu^{(1)} + \tau^2 A_\mu^{(2)}\right) L &= \frac{g}{2\sqrt{2}} \left[\left(\bar{\nu}_{\mathrm{L}} \gamma_\mu e_{\mathrm{L}} + \bar{e}_{\mathrm{L}} \gamma_\mu \nu_{\mathrm{L}}\right) \left(W_\mu^{+} + W_\mu^{-}\right) \right. \\ &\qquad \left. + \left(\bar{\nu}_{\mathrm{L}} \gamma_\mu e_{\mathrm{L}} - \bar{e}_{\mathrm{L}} \gamma_\mu \nu_{\mathrm{L}}\right) \left(W_\mu^{+} - W_\mu^{-}\right) \right] \\ &= \frac{g}{\sqrt{2}} \left(\bar{\nu}_{\mathrm{L}} \gamma_\mu e_{\mathrm{L}} W_\mu^{+} + \bar{e}_{\mathrm{L}} \gamma_\mu \nu_{\mathrm{L}} W_\mu^{-}\right).\end{aligned}$$

Using the definitions (15.3, 15.5), one also obtains the couplings of fermions to the neutral vector fields A_μ and Z_μ,

$$\begin{aligned}&\frac{e Z_\mu}{\sin 2\theta_{\mathrm{W}}} \left(2 \sin^2 \theta_{\mathrm{W}} \bar{e}_{\mathrm{R}} \gamma_\mu e_{\mathrm{R}} - \cos 2\theta_{\mathrm{W}} \bar{e}_{\mathrm{L}} \gamma_\mu e_{\mathrm{L}} + \bar{\nu}_{\mathrm{L}} \gamma_\mu \nu_{\mathrm{L}}\right) \\ &\quad - e A_\mu \left(\bar{e}_{\mathrm{R}} \gamma_\mu e_{\mathrm{R}} + \bar{e}_{\mathrm{L}} \gamma_\mu e_{\mathrm{L}}\right).\end{aligned}$$

Note that from the study of the Z decays, one infers that the number of generations with 'light' (*i.e.*, with a mass below $m_Z/2 \approx 45$ GeV) active neutrinos is exactly three.

Beyond the Standard Model: neutrino oscillations. The experimental discovery of neutrino oscillations implies that neutrinos have masses. One infers

$$m^2(\nu_\mu) - m^2(\nu_e) \approx 7.5 \times 10^{-5}\text{eV}^2\,,\ |m^2(\nu_\tau) - m^2(\nu_\mu)| \approx 2.5 \times 10^{-3}\text{eV}^2,$$

and the mass eigenstates differ from the linear combinations of neutrinos appearing in the weak interactions. Including cosmological observations, one finds a bound on the sum of masses of about 0.2 eV. As we have shown, this necessitates a slight modification of the Standard Model. If neutrinos are Dirac fermions, this implies a mixing (PMNS) matrix like in the quark sector, as described in section 15.3, with three measured CP-conserving angles

$$\sin^2(2\theta_{12}) = 0.846 \pm 0.021\,,\ \sin^2(2\theta_{23}) = 1 + 0 - 0.017\,,$$
$$\sin^2(2\theta_{13}) = 9.3 \pm 0.8 \times 10^{-2},$$

and one unknown CP violating angle.

A major issue is the very small neutrino masses. This problem could have a solution in the framework of Grand Unified Theories, in a scheme with Majorana mass terms for neutrinos. This would imply some form of lepton number violation (see section 11.8.3).

15.3 Quarks and weak–electromagnetic interactions

With the addition of the right-handed neutrino, the structure of the lepton and quark sectors become quite similar. We first consider again only one generation with two flavours, corresponding to the quarks **u** and **d**, the coupling of other generations being analogous. Each quark has a colour quantum number and forms a $SU(3)$ triplet (see next section). The left components of the quarks belong to a $SU(2)$ doublet $\mathbf{Q}_{\mathrm{L}}$.

$$\mathbf{Q}_{\mathrm{L}} = \{\mathbf{u}_{\mathrm{L}}\ ,\ \mathbf{d}_{\mathrm{L}}\}\ .$$

All right-handed components $\mathbf{Q}_{1,2\mathrm{R}}$ form $SU(2)$ singlets. The gauge invariant quark action can be written as

$$\mathcal{S}_{\mathrm{Q}} = -\int \mathrm{d}^4x \Big[\sum_i \bar{\mathbf{Q}}_{i\mathrm{R}} \left(\partial\!\!\!/ + \tfrac{1}{2} i g' Y_{i\mathrm{R}} B\!\!\!/\right) \mathbf{Q}_{i\mathrm{R}} + \bar{\mathbf{Q}}_{\mathrm{L}} \left(\partial\!\!\!/ + \tfrac{1}{2} i g' Y_{\mathrm{L}} B\!\!\!/ + \tfrac{1}{2} i g \mathbf{A}\!\!\!/ \cdot \boldsymbol{\tau}\right) \mathbf{Q}_{\mathrm{L}}\Big]. \qquad (15.16)$$

We now use the relations (15.11, 15.12),

$$Q_1 - Q_2 = 1\,,\ Y_{1\mathrm{R}} = 2Q_1\,,\ Y_{2\mathrm{R}} = 2(Q_1 - 1)\,,\ Y_{\mathrm{L}} = 2Q_1 - 1\,.$$

The proton is a **uud** state and has charge $+1$. Thus, $2Q_1+Q_2 = 3Q_1-1 = 1$ and the quark **u** has charge $2/3$. Then, the quark **d** has charge $Q_2 = Q_1 - 1 = -1/3$. The neutron is a **udd** state and one verifies that it has indeed charge 0, a result consistent with the $SU(2)$ doublet assignment. Finally,

$$Y_{\mathrm{L}} = 1/3\,,\ Y_{1\mathrm{R}} = 4/3\,,\ Y_{2\mathrm{R}} = -2/3\,.$$

We verify in section 15.7 that the $SU(3)$ triplet structure of quarks leads to the cancellation of the possible anomaly due to the chiral coupling of gauge fields to fermions in each generation and, therefore, ensures the consistency of the gauge theory of weak and electromagnetic interactions.

15.3.1 Couplings to the Higgs field and quark masses

As for leptons direct quark mass terms are forbidden by the $SU(2)$ symmetry. The quark masses are produced by the coupling to the Higgs scalar field and the spontaneous symmetry breaking. The $SU(2)\times U(1)$ invariant Higgs–quark coupling terms can be easily inferred from the expressions (15.13),

$$\mathcal{S}_{\text{qmass}} = \int \mathrm{d}^4x \,\Big\{iG_{q1}\left[\bar{\mathbf{Q}}_{1\mathrm{R}}(x)\big(\boldsymbol{\varphi}(x)\tau_2\mathbf{Q}_{\mathrm{L}}(x)\big) + \big(\bar{\mathbf{Q}}_{\mathrm{L}}(x)\tau_2\boldsymbol{\varphi}^*(x)\big)\mathbf{Q}_{1\mathrm{R}}(x)\right] \\ + G_{q2}\left[\bar{\mathbf{Q}}_{2\mathrm{R}}(x)\big(\boldsymbol{\varphi}^*(x)\cdot\mathbf{Q}_{\mathrm{L}}(x)\big) + \big(\bar{\mathbf{Q}}_{\mathrm{L}}(x)\cdot\boldsymbol{\varphi}(x)\big)\mathbf{Q}_{2\mathrm{R}}(x)\right]\Big\}, \tag{15.17}$$

which can provide masses for the two quarks. This is at least the situation for one generation. However, six quarks belonging to three generations have been discovered (see table 15.1).

Therefore, in the interactions (15.17) the spinors which appear on the right and the left then need not be, and are not the same.

When one replaces the scalar field φ by its expectation value, one obtains in general a non-diagonal mass matrix of the form

$$\sum_{\alpha,\beta}\left[\bar{\mathbf{Q}}^{\alpha}_{1\mathrm{R}} M_{\alpha\beta}\mathbf{Q}^{\beta}_{1\mathrm{L}} + \bar{\mathbf{Q}}^{\alpha}_{1\mathrm{L}} M^{\dagger}_{\alpha\beta}\mathbf{Q}^{\beta}_{1\mathrm{R}}\right],$$

for the quarks of charge $2/3$, and a similar one for the charge $-1/3$ quarks. Performing independent unitary transformations $\mathbf{U}_{\mathrm{R,L}}$ on the right and left quark components, it is possible to replace the matrix $\mathbf{M}$ by a real diagonal matrix $\mathcal{M}$:

$$\mathbf{U}^{\dagger}_{\mathrm{R}}\mathbf{M}\mathbf{U}_{\mathrm{L}} = \mathcal{M}\,,$$

In this representation the quarks are mass eigenstates. However, the weak interactions no longer have the simple form (15.16) because the unitary transformations on the quark components $\mathbf{Q}_{1\mathrm{L}}$ and $\mathbf{Q}_{2\mathrm{L}}$ are in general different. It is customary to put the blame onto the charge $-1/3$ quarks.

The mismatch is expressed in terms of a 3×3 unitary matrix (because three generations are observed), the Cabibbo–Kobayashi–Maskawa matrix (CKM) which relates the quark mass eigenstates **d**, **s** and **b** to the quarks appearing in the weak interactions:

$$\left[\mathbf{Q}_2^{\alpha}\right]_{\text{weak int.}} = \sum_{\beta} U_{\alpha\beta} \left[\mathbf{Q}_2^{\beta}\right]_{\text{mass eigenst.}} .$$

With only two generations (**d** and **s**), it is possible to cast the matrix into the form

$$\mathbf{U}_{\mathrm{C}} = \begin{pmatrix} \cos\theta_{\mathrm{C}} & \sin\theta_{\mathrm{C}} \\ -\sin\theta_{\mathrm{C}} & \cos\theta_{\mathrm{C}} \end{pmatrix},$$

in which θ_{C} is the Cabibbo angle, after unobservable changes of the relative phases between the quarks.

Table 15.1

Quarks and leptons. The three generations (2014).

Charge 2/3 quarks	Charge $-1/3$ quarks	Charge -1 leptons	Neutrinos
u, 1.7 to 3.0 MeV	**d**, 4.5 to 5.3 MeV	e, 0.511 MeV	*cf.* 15.2
c, 1.275 $\pm$ 0.025 GeV	**s**, 95 $\pm$5 MeV	μ, 105.6 MeV	*cf.* 15.2
t, 173.2 $\pm$ 1.2 GeV	**b**, 4.18 $\pm$0.03 GeV	τ, 1.777 GeV	*cf.* 15.2

In the presence of the third **b** quark, the 3×3 CKM matrix can be parametrized in terms of three rotation angles and one CP violating phase responsible for the observed direct CP violation in neutral kaon and B_0 meson decay. However, the resulting CP violation seems too small to explain the disappearance of antimatter.

Beyond the Standard Model: grand unification. Theoretical speculations based on the search for a unifying simple group including $U(1) \times SU(2) \times SU(3)$ as a subgroup, have mainly focused on $SU(5)$ (the larger $SO(10)$ has also been discussed). This group deals nicely with fermions, has 12 additional super-heavy gauge bosons but necessitates a large collection of Higgs fields.

Running the three independent couplings to higher energies an apparent unification is then observed at energies of order 10^{15} GeV. The non-observation of the predicted proton decay has shifted the focus to the minimal supersymmetric extension of $SU(5)$ where the problem with proton decay is less severe and the apparent unification of the running coupling constants more precise. However, no supersymmetric particles have been found yet at LHC (2015).

15.3.2 Elementary scalar fields: a few problems

The Standard Model has (Higgs) scalar fields as an essential ingredient. This is the source of several difficulties. The Higgs field is responsible for the masses of all fundamental particles, but in the Standard Model these masses are all given in terms of arbitrary parameters, like for example the Yukawa couplings that determine the fermions masses. This could perhaps be expected from an effective low energy theory. Yet puzzling is the diversity of these couplings. If the couplings were 'natural', that is, of order unity, all fermion masses would be in the few 100 GeV range, like the W and Z masses or the Higgs expectation value. In this sense only the top quark (**t**) mass, which is about 173 GeV, is natural. By contrast, only in the quark sector the masses span about five order of magnitudes, something that even a possible RG running of the fermion–Higgs couplings cannot easily explain. Taking into account the lepton sector makes the problem even worse.

Another, perhaps even more fundamental, problem is related to the scalar field mass renormalization. Generically the scalar bare and physical masses are expected to be of the order of the momentum ‘cut-off’ which gives the scale of some new physics. It is only by fine tuning the scalar bare mass that one can render the renormalized mass much smaller than the cut-off. In the statistical physics interpretation of the ϕ^4 theory the divergence of the correlation length (the inverse physical mass in the particle language) is obtained by adjusting the temperature, and thus the bare mass, close to a critical value where a second-order phase transition occurs. However, in particle physics all parameters are given and it is somewhat unnatural for the scalar bare mass to lie accidentally close to such a critical value. To get a rough idea about the severity of the problem we can use perturbation theory, since the ϕ^4 coupling λ is of order 1.

Neglecting other interactions, the one-loop mass counter-term with a momentum cut-off Λ is

$$\delta m_0^2 = \frac{\lambda}{16\pi^4} \int \frac{\mathrm{d}^4 p}{p^2(1 + p^2/\Lambda^2)^2} = \frac{\lambda}{16\pi^2}\Lambda^2.$$

The Higgs mass m_{H} at leading order is given by

$$m_{\mathrm{H}}^2 = \tfrac{1}{3}\lambda v^2.$$

Therefore,

$$f \equiv \delta m_0^2 / m_{\mathrm{H}}^2 = \frac{3}{16\pi^2}(\Lambda/v)^2.$$

Another important contribution comes a top quark loop, which is of order $(\Lambda/v)^2/\pi^2$. Considering the unexplained range of fermion masses, it is difficult to decide how much fine tuning is acceptable.

If one demands that the correction should be smaller than the leading term, then one obtains $\Lambda < 700\,\mathrm{GeV}$, which is almost excluded by the LHC.

If no new physics beyond the Standard Model is found at the LHC, the fine tuning factor will about 10. On the other hand, an absence of any new physics below Planck's scale (10^{19} GeV) is difficult to believe.

Fine tuning solutions. At the scale of new physics the scalar field problem must be cured. Three types of schemes have been proposed so far:

(i) The Higgs boson is a bound state of a new type of fermions. This requires a specific model, hopefully not involving new scalars again. Models in which the forces are again due to gauge interactions have been proposed and fall under the name of technicolour (see next section). Such models have problems generating fermion masses and no widely accepted model has been proposed.

(ii) The Higgs boson remains associated with a fundamental field, but the mass renormalization problem is solved with the help of supersymmetry. Since fermions, due to chiral symmetry, can be naturally massless, the idea is to use supersymmetry to relate them to scalars. In such models the scalar mass renormalization grows only logarithmically with the cut-off (it would be absent in the absence of supersymmetry breaking) and thus the problem is much less severe even if the cut-off is of the order of the Planck mass. The main difficulty with this approach is that none of the superpartners of existing particles have yet been found. Moreover, the mechanism of spontaneous supersymmetry breaking is not fully understood.

(iii) The Higgs particle is related to additional dimensions of space, for instance, is a component of a vector field. Then, it must transform under the corresponding adjoint representation.

15.4 Quantum ChromoDynamics: renormalization group

We now concentrate on Quantum Chromodynamics (QCD) a gauge theory of quarks and vector particles called gluons, which produces the observed hadrons and their Strong Interactions, neglecting completely the weak and electromagnetic interactions that we have described in the preceding sections. Quantum Chromodynamics, as it stands today, consists in a set of quarks characterized by a flavour quantum number, relevant for Weak Interactions, which are also triplets of a gauged symmetry, the $SU(3)$ colour, realized in the symmetric phase. Their interactions are mediated by the corresponding gauge fields, associated to gluons,

$$\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{Q}}, \mathbf{Q}) = -\int \mathrm{d}^4 x \left[\frac{1}{4g^2}\sum_{\mu,\nu} \operatorname{tr} \mathbf{F}_{\mu\nu}^2 + \sum_{\text{flavours}} \bar{\mathbf{Q}}_f \left(\not{\mathbf{D}} + m_f\right) \mathbf{Q}_f\right].$$

The most important physical arguments in favour of such a model are

(i) Quarks behave almost like free particles at short distances, as indicated by deep inelastic scattering experiments or the spectrum of bound states of heavy quarks. We calculate below the RG β-function and show that a pure non-Abelian gauge theory is asymptotically free (AF) at large momentum in four dimensions (like the non-linear σ-model in two dimensions). This property survives the inclusion of a limited number of fermions and, furthermore, the property is specific to non-Abelian gauge theories.

(ii) No free quarks and gluons have ever been observed at large distance (but they manifest themselves indirectly in the jet physics). This is consistent with the simplest picture in which the β-function (which, due to AF, is negative at small coupling) remains negative for all couplings in such a way that the effective coupling constant grows without bounds at large distances. Numerical simulations strongly support this conjecture, called the colour confinement hypothesis (see section 16).

15.4.1 RG equations in the covariant gauge

We first discuss the gauge dependence of RG equations and functions of pure gauge theories in the covariant gauge (14.34), that is, the dependence on the parameter ξ. A short discussion of the Abelian case can be found in section 13.10. We call Z_A the gauge field renormalization constant and Z_g the renormalization constant of the coupling constant

$$\alpha = g^2/4\pi \,.$$

α in Strong Interactions is in general denoted α_s to distinguish it from its QED analogue. In this section no confusion is possible. It can be shown that for such gauges, as in the Abelian case, the gauge fixing term is not renormalized. Therefore,

$$\xi_0 = \xi Z_A/Z_g \,.$$

In terms of the renormalization scale μ the RG equation for the gauge field n-point function reads:

$$\left[\mu\frac{\partial}{\partial\mu}+\beta(\alpha,\xi)\frac{\partial}{\partial\alpha}+\delta(\alpha,\xi)\xi\frac{\partial}{\partial\xi}-\frac{n}{2}\eta_A(\alpha,\xi)\right]\tilde{\Gamma}^{(n)}(\mu,\alpha,\xi)=0\,,$$

where,

$$\delta(\alpha,\xi)\equiv\xi^{-1}\,\mu\frac{\partial}{\partial\mu}\bigg|_{\alpha_0\,,\,\xi_0\,\text{fixed}}\xi=\beta(\alpha,\xi)-\eta_A(\alpha,\xi).$$

One proves that the bare correlation functions of gauge invariant operators are gauge independent. This in particular implies that they are independent of ξ_0. The same property applies to the renormalization constants needed to render these correlation functions finite. It is thus possible to construct renormalized correlation functions which are also ξ_0 independent.

We denote by Γ such a correlation function. It satisfies

$$\left.\frac{\partial}{\partial \xi}\right|_{\alpha_0\,,\,\text{cut-off fixed}} \Gamma = \left(\frac{\partial}{\partial \xi} + \rho(\alpha,\xi)\frac{\partial}{\partial \alpha}\right)\Gamma(\mu,\alpha,\xi) = 0 \tag{15.18}$$

with

$$\rho(\alpha,\xi) = \left.\frac{\partial \alpha}{\partial \xi}\right|_{\alpha_0\,,\,\text{cut-off fixed}} . \tag{15.19}$$

Γ also satisfies a RG equation which we assume to be homogeneous:

$$\left[\mu\frac{\partial}{\partial \mu} + \beta(\alpha,\xi)\frac{\partial}{\partial \alpha} + \delta(\alpha,\xi)\xi\frac{\partial}{\partial \xi} - \eta_\Gamma(\alpha,\xi)\right]\Gamma(\mu,\alpha,\xi) = 0\,.$$

Using equation (15.18) to eliminate $\partial/\partial\xi$, we obtain the new RG equation for Γ,

$$\left[\mu\frac{\partial}{\partial \mu} + \tilde\beta(\alpha,\xi)\frac{\partial}{\partial \alpha} - \eta_\Gamma(\alpha,\xi)\right]\Gamma(\mu,\alpha,\xi) = 0 \text{ with } \tilde\beta = \beta - \xi\delta\rho\,. \tag{15.20}$$

Expressing then the compatibility condition between the two linear equations (15.18) and (15.20), one obtain two equations,

$$\left(\frac{\partial}{\partial\xi}+\rho(\alpha,\xi)\frac{\partial}{\partial\alpha}\right)\eta_\Gamma(\alpha,\xi)=0\,,$$
$$\left(\frac{\partial}{\partial\xi}+\rho(\alpha,\xi)\frac{\partial}{\partial\alpha}\right)\tilde\beta=\frac{\partial\rho}{\partial\alpha}\tilde\beta\,.$$

The first equation expresses that, as expected, the multiplicative renormalization of Γ is independent of ξ_0. The second equation shows that the zeros of $\tilde\beta$ are gauge independent. Differentiating the equation with respect to α one also finds that the slope of β at its zeros is gauge independent. Finally, one verifies that in a minimal subtraction scheme the function ρ vanishes. In dimensional regularization the relation between α_0 and α takes the form

$$\alpha_0=\mu^\varepsilon\alpha Z_g=\mu^\varepsilon\alpha\left(1+\frac{Z_g^1(\alpha,\xi)}{\varepsilon}+\frac{Z_g^2(\alpha,\xi)}{\varepsilon^2}+\cdots\right).$$

The important point is that the term without pole in ε in the expansion of Z_g is ξ independent. Using the definition (15.19) of ρ, we then find

$$0 = \rho\left(1 + \frac{\partial Z_g^1}{\partial \alpha}\frac{1}{\varepsilon} + \cdots\right) + \alpha\left(\frac{\partial Z_g^1}{\partial \xi}\frac{1}{\varepsilon} + O\left(\frac{1}{\varepsilon^2}\right)\right). \tag{15.21}$$

Therefore, the expansion of ρ for ε small has only singular contributions. Since ρ is finite, all singular contributions must cancel and thus ρ vanishes identically. It follows that in the minimal subtraction scheme the β-function and η_Γ are independent of ξ.

15.4.2 The RG β-function at one-loop order

We now calculate the RG β-function at leading order in a gauge theory corresponding to a simple group G in particular to verify asymptotic freedom, because no simple explanation has yet been proposed which allows one to understand the sign without explicit calculation.

The calculation can be done by various methods, for example, we could use the background field method. Here, instead, we calculate directly the β-function from the renormalization of the gauge coupling constant as defined by the fermion-gauge field vertex. We thus need the divergent parts of the gauge field and fermion two-point functions, and the fermion gauge field three-point function. We work in the Feynman gauge and use dimensional regularization. The normalizations of vertices and propagators are those given in section 14.4.

The gauge field two-point function. Four diagrams contribute to the two-point function, corresponding to the gauge field loops, the Faddeev–Popov ghost loop and the fermion loops (see figure 15.1).

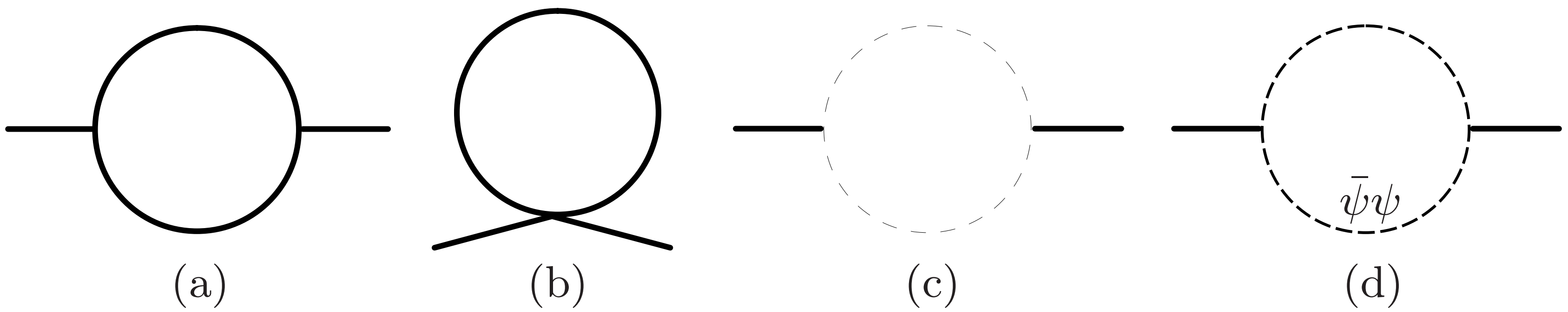

Fig. 15.1 – The gauge field two-point function at one-loop (dotted lines represent ghosts).

The diagram (b) corresponding to the self-contraction of the gauge four-point vertex vanishes in dimensional regularization. The fermion loop contribution (d) has already been calculated in section 13.11 up to a simple geometric factor.

Diagram (a) is given by

$$(\mathrm{a}) = \tfrac{1}{2} f_{acd} f_{bcd} \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{N_{\mu\nu}(k,q)}{q^2 (k+q)^2}$$

with

$$N_{\mu\nu}(k,q) = \delta_{\mu\nu}(5k^2 + 2k \cdot q + 2q^2) + k_\mu k_\nu (d-6) + (q^\mu k_\nu + q^\nu k_\mu)(2d-3) \\ + 2q_\mu q_\nu (2d-3).$$

To calculate the diagrams, we project the integrand over $\delta_{\mu\nu}$ and $k_\mu k_\nu$ and use repeatedly the identity

$$2k \cdot q = (k+q)^2 - k^2 - q^2.$$

We set

$$\sum_{c,d} f_{acd} f_{bcd} = C(G) \delta_{ab} \,.$$

A short calculation yields the divergent part:

$$(\mathrm{a})_{\mathrm{div}} = \delta_{ab} \frac{C(G)}{12} \left(19k^2 \delta_{\mu\nu} - 22 k_\mu k_\nu\right) \frac{g^2}{8\pi^2 \varepsilon}.$$

Diagram (c) is given by

$$(\mathrm{c}) = -\sum_{c,d} f_{acd} f_{bcd} \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{q_\mu (k+q)_\nu}{q^2 (k+q)^2}.$$

The divergent part is

$$(\mathrm{c})_{\mathrm{div}} = \delta_{ab} \left(k^2 \delta_{\mu\nu} + 2 k_\mu k_\nu\right) \frac{1}{12} C(G) \frac{g^2}{8\pi^2 \varepsilon}.$$

Note that both divergent contributions are not separately transverse. By adding them we get the divergent part of the two-point function in the absence of fermions, which now is transverse as expected:

$$\left[\tilde{\Gamma}^{(2)ab}_{\mu\nu}(k)\right]_{\mathrm{div}} = \delta_{ab} \left(k^2 \delta_{\mu\nu} - k_\mu k_\nu\right) \frac{5}{3} C(G) \frac{g^2}{8\pi^2 \varepsilon}.$$

Denoting by Z_A and Z_g the renormalization constants of the gauge field and the coupling constant g^2, we obtain the relation at one-loop order:

$$\frac{Z_A}{Z_g} = 1 + \frac{5}{3} C(G) \frac{g^2}{8\pi^2 \varepsilon}.$$

Adapting the result (13.64), we infer the additional fermion contribution. We finally obtain

$$\frac{Z_A}{Z_g} = 1 + \left(\frac{5}{3} C(G) - \frac{4}{3} T(R) \right) \frac{g^2}{8\pi^2 \varepsilon}, \tag{15.22}$$

where the fermions belong to the representation R and $T(R)$ is the trace of the square of the generators in this representation:

$$\operatorname{tr} \mathbf{t}^a \mathbf{t}^b = -\delta_{ab} T(R).$$

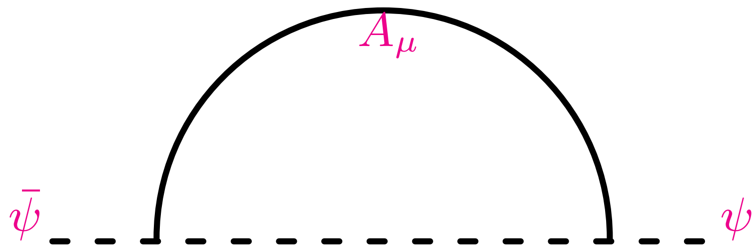

Fig. 15.2 – One-loop contribution to the fermion two-point function.

The fermion two-point function. One diagram contributes to the fermion two-point function $\Gamma^{(2)}$ (see figure 15.2), which differs from its QED counterpart only by a geometric factor:

$$\tilde{\Gamma}^{(2)}_{1\,\text{loop}}(k) = g^2 \int \frac{\mathrm{d}^d q}{(2\pi)^d} t^a \gamma_\mu \frac{1}{i\not{q} + m} t^a \gamma_\mu \frac{1}{(k-q)^2}$$

(for simplicity we have given the same mass m to all fermions since this does not affect the result). Since we need only the field renormalization we can project the integrand over $\not{k}$. The following identity is useful:

$$\sum_\nu \gamma_\nu \gamma_\mu \gamma_\nu = (2-d)\gamma_\mu\,.$$

Calculating the divergent part of the integral, we obtain the fermion field renormalization Z_{F} at one-loop order:

$$Z_{\mathrm{F}} = 1 - C(R)\frac{g^2}{8\pi^2\varepsilon} \quad \text{with} \quad \sum_a \mathbf{t}^a\mathbf{t}^a = -C(R)\mathbf{1}\,. \tag{15.23}$$

The gauge field fermion vertex. Two diagrams contribute at one-loop order (see figure 15.3), the first has a QED counterpart, the second being specific to a non-Abelian theory.

Fig. 15.3 – The gauge field fermion vertex at one-loop.

$$(\mathrm{a}) = g^2 \int \frac{\mathrm{d}^d q}{(2\pi)^d} \sum_{\nu,b} \mathbf{t}^b \gamma_\nu \frac{1}{i\not{q} + m} \mathbf{t}^a \gamma_\mu \frac{1}{i\not{q} + i\not{k} + m} \mathbf{t}^b \gamma_\nu \frac{1}{(p_1 - q)^2}.$$

To calculate the divergent part of the integral we multiply by γ_μ and take the trace. We also use the identity

$$\sum_b \mathbf{t}^b \mathbf{t}^a \mathbf{t}^b = \left(\tfrac{1}{2} C(G) - C(R)\right) \mathbf{t}^a.$$

We then find

$$(\mathrm{a})_{\text{div.}} = \left(C(R) - \tfrac{1}{2} C(G)\right) \mathbf{t}^a \gamma_\mu \frac{g^2}{8\pi^2 \varepsilon},$$

$$(\mathrm{b}) = i \sum_{b,c} f_{abc} g^2 \int \frac{\mathrm{d}^d q}{(2\pi)^d} \mathbf{t}^b \sum_{\nu,\rho} \gamma_\nu \frac{1}{i\not{p}_1 - i\not{q} + m} \mathbf{t}^c \gamma_\rho \frac{V_{\mu\nu\rho}(k, q, -k-q)}{q^2 (k+q)^2}$$

with (equation (14.40)):

$$V_{\mu\nu\rho}(k,q,r) = (r-q)_\mu \delta_{\nu\rho} + (k-r)_\nu \delta_{\rho\mu} + (q-k)_\rho \delta_{\mu\nu} \,.$$

The divergent part is

$$(\mathrm{a})_{\mathrm{div.}} = \frac{3}{2} C(G) t^a \gamma_\mu \frac{g^2}{8\pi^2 \varepsilon}.$$

It follows that

$$Z_{\mathrm{F}} Z_A^{1/2} = 1 - \left(C(R) + C(G)\right) \frac{g^2}{8\pi^2\varepsilon},$$

and, therefore, (equation (15.23)):

$$Z_A = 1 - 2C(G) \frac{g^2}{8\pi^2\varepsilon}.$$

Finally, using the result (15.22) we obtain

$$Z_g = 1 - \left(\tfrac{11}{3} C(G) - \tfrac{4}{3} T(R)\right) \frac{g^2}{8\pi^2\varepsilon}.$$

The β-function at one-loop order follows:

$$\beta(g^2) = -\varepsilon\left[\frac{\mathrm{d}\ln\left(g^2 Z_g\right)}{\mathrm{d}g^2}\right]^{-1} = -\left[\tfrac{11}{3}C(G) - \tfrac{4}{3}T(R)\right]\frac{g^4}{8\pi^2} + O\left(g^6\right). \tag{15.24}$$

In the case of the $SU(N)$ group with N_{F} fermions in the fundamental representation the values of $C(G)$ and $T(R)$ are $C(G) = N$, $T(R) = \frac{1}{2}N_{\mathrm{F}}$ and, therefore,

$$\beta(g^2) = -\left(\frac{11N}{3} - \frac{2N_{\mathrm{F}}}{3}\right)\frac{g^4}{8\pi^2} + O\left(g^6\right). \tag{15.25}$$

The theory is asymptotically free, that is, the β-function is negative for small coupling for

$$N_{\mathrm{F}} < 11N/2\,,$$

which, in the case of $SU(3)$, means at most 16 flavours. If this condition is met, $g = 0$ is a UV fixed point.

15.5 The Abelian anomaly

We have pointed out in sections 13.6 and 14.4 that none of the standard regularization methods can deal in a straightforward way with one-loop diagrams in the case of chiral gauge fields. We now show that indeed gauge theories with massless fermions and chiral symmetry can be found where the axial current is not conserved. The divergence of the axial current, when it does not vanish, is called an anomaly. This leads in particular to obstructions to the construction of gauge theories where the gauge field couples differently to the two fermion chiral components. Several examples are physically important like the theory of weak electromagnetic interactions, the electromagnetic decay of the π_0 meson, or the $U(1)$ problem.

We first discuss the Abelian axial current, in four dimensions (the generalization to all even dimensions is straightforward), and then the general non-Abelian case.

The only possible source of anomalies are one-loop fermion diagrams in gauge theories when chiral properties are involved. This reduces the problem to the discussion of fermions in the background of gauge fields, or equivalently to the properties of the determinant of the gauge covariant Dirac operator.

15.5.1 Abelian axial current and Abelian vector gauge field

We first consider the QED-like fermion action $\mathcal{S}(\bar{\psi},\psi)$ for massless Dirac fermions $\psi,\bar{\psi}$ in the background of an Abelian gauge field A_μ:

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}^4x\,\bar{\psi}(x)\not{D}\psi(x),\quad \not{D} \equiv \not{\partial} + ie\not{A}\,, \tag{15.26}$$

and the corresponding field integral

$$\mathcal{Z}(A_\mu) = \int \left[\mathrm{d}\psi\mathrm{d}\bar{\psi}\right]\exp\left[-\mathcal{S}(\psi,\bar{\psi})\right] = \det\not{D}\,.$$

In what follows we denote by $\langle \bullet \rangle$ expectation values with respect to the measure.

We can find regularizations which preserve gauge invariance, and since the fermions are massless, chiral symmetry. Therefore, we would naively expect the corresponding axial current to be conserved. However, the proof of current conservation involves space-dependent chiral transformations and, therefore, steps that cannot be regularized without breaking one of the symmetries.

The coefficient of $\partial_\mu \theta(x)$ in the variation of the action under a space-dependent chiral transformation,

$$\psi_\theta(x) = \mathrm{e}^{i\theta(x)\gamma_5}\,\psi(x), \quad \bar{\psi}_\theta(x) = \bar{\psi}(x)\,\mathrm{e}^{i\theta(x)\gamma_5}, \tag{15.27}$$

yields the axial current $J^5_\mu(x)$. For the action (15.26) one finds,

$$\delta\mathcal{S} = \int \mathrm{d}^4x \sum_\mu \partial_\mu \theta(x) J^5_\mu(x) \quad \text{with} \quad J^5_\mu(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x).$$

After the transformation (15.27), $\mathcal{Z}(A_\mu)$ becomes

$$\mathcal{Z}(A_\mu,\theta) = \det\left[e^{i\gamma_5\theta(x)}\,\not{D}\,e^{i\gamma_5\theta(x)}\right].$$

Since $e^{i\gamma_5\theta}$ has a determinant which is unity, one might naively conclude that $\mathcal{Z}(A_\mu,\theta)=\mathcal{Z}(A_\mu)$ and, therefore, that the current $J^5_\mu(x)$ is conserved. This is a conclusion we now check by an explicit calculation of the expectation value of $\sum_\mu \partial_\mu J^5_\mu(x)$ in the case of the action (15.26).

Regularizations.

(i) For any regularization which is consistent with the hermiticity of γ_5

$$|\mathcal{Z}(A_\mu,\theta)|^2 = \det(\not{D}\not{D}^\dagger).$$

Therefore, an anomaly can appear only in the imaginary part of $\ln\mathcal{Z}$.

(ii) If the regularization is gauge invariant, $\mathcal{Z}(A_\mu, \theta)$ is also gauge invariant. Therefore, a possible anomaly will also be gauge invariant. One regularization scheme that has the required property is based on regulator fields. But as the discussion of section 13.6.1 has shown, at least one regulator field must be an unpaired massive boson with spin, dividing the fermion determinant by a factor $\det(\not{D} + \Lambda)$. If this boson has a chiral charge global chiral symmetry is broken by the mass Λ; if it has no chiral charge global chiral symmetry is preserved, and the determinant is independent of θ for $\theta(x)$ constant, but then the ratio of determinants is not invariant under local chiral transformations.

General form of the possible anomaly. The operator $\sum_\mu \partial_\mu J^5_\mu(x)$ has dimension 4 and since a possible anomaly is a large momentum or short distance effect, $\langle \sum_\mu \partial_\mu J^5_\mu(x) \rangle$ can only be a local function of A_μ of dimension 4. In addition parity implies that it is proportional to the completely antisymmetric tensor $\epsilon_{\mu\nu\rho\sigma}$. This determines $\langle \sum_\mu \partial_\mu J^5_\mu(x) \rangle$ up to a multiplicative constant,

$$\sum_\lambda \left\langle \partial_\lambda J^5_\lambda(x) \right\rangle \propto e^2 \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} \partial_\mu A_\nu(x) \partial_\rho A_\sigma(x) \propto e^2 \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \,,$$

$F_{\mu\nu}$ being the electromagnetic tensor. The possible anomaly is automatically gauge invariant. It is also a total derivative since

$$\sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = 4 \sum_{\mu,\nu,\rho,\sigma} \partial_\mu \left(\epsilon_{\mu\nu\rho\sigma} A_\nu \partial_\rho A_\sigma \right). \tag{15.28}$$

To find the multiplicative factor, which is the only regularization dependent feature, it is sufficient to calculate the coefficient of term quadratic in A in the expansion of $\langle \sum_\lambda \partial_\lambda J^5_\lambda(x) \rangle$ in powers of A.

We define the three-point function

$$\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{\delta}{\delta A_\mu(p_1)}\frac{\delta}{\delta A_\nu(p_2)}\left\langle J^5_\lambda(k)\right\rangle\Big|_{A=0}, \tag{15.29}$$
$$= \frac{\delta}{\delta A_\mu(p_1)}\frac{\delta}{\delta A_\nu(p_2)} i\,\mathrm{tr}\left[\gamma_5\gamma_\lambda \not{D}^{-1}(k)\right]\Big|_{A=0}.$$

$\Gamma^{(3)}$ is the sum of the two Feynman diagrams of figure 15.4.

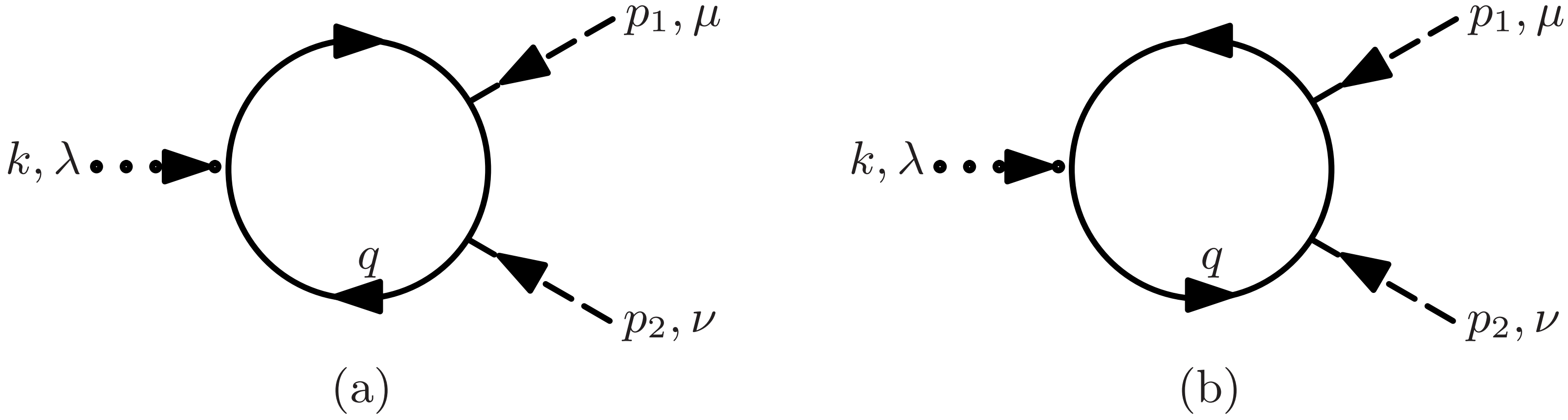

Fig. 15.4 – Anomalous diagrams.

The contribution of diagram (a) is

$$(\mathrm{a}) \mapsto \frac{e^2}{(2\pi)^4} \operatorname{tr}\left[\int \mathrm{d}^4 q\, \gamma_5 \gamma_\lambda (\not{q} + \not{k})^{-1} \gamma_\mu (\not{q} - \not{p}_2)^{-1} \gamma_\nu \not{q}^{-1}\right],$$

and the contribution of diagram (b) is obtained by exchanging $p_1, \gamma_\mu \leftrightarrow p_2, \gamma_\nu$.

Power counting tells us that the function $\Gamma^{(3)}$ may have a linear divergence which, due to the presence of the γ_5 factor, must be proportional to $\epsilon_{\lambda\mu\nu\rho}$, symmetric in the exchange $p_1, \gamma_\mu \leftrightarrow p_2, \gamma_\nu$, and thus proportional to

$$\sum_\rho \epsilon_{\lambda\mu\nu\rho}(p_1 - p_2)_\rho \,. \tag{15.30}$$

On the other hand, by commuting γ_5 we notice that $\Gamma^{(3)}$ is formally a symmetric function of the three sets of external arguments. A divergence breaks the symmetry between external arguments.

Therefore, a symmetric regularization of the kind we shall adopt leads to a finite result. The result is not ambiguous because a possible ambiguity again is proportional to (15.30).

In the same way if the regularization is consistent with vector gauge invariance the WT identity

$$\sum_{\mu} p_{1\mu} \Gamma^{(3)}_{\lambda\mu\nu}(k; p_1, p_2) = 0\,,$$

is satisfied. Applied to the divergent part it yields

$$\sum_{\mu,\rho} -p_{1\mu} p_{2\rho} \epsilon_{\lambda\mu\nu\rho} = 0\,,$$

a condition that is not satisfied. Therefore, the sum of the two diagrams is finite. Different regularizations may still differ by finite quantities of the form (15.30) but again all regularizations consistent with vector gauge invariance must give the same answer.

Therefore, there are two possibilities:

(i) The divergence $\sum_\lambda k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)$ in a regularization respecting the symmetry between the three arguments vanishes. Then $\Gamma^{(3)}$ is gauge invariant and the axial current is conserved.

(ii) The divergence of the symmetric regularization does not vanish. Then it is possible to add to $\Gamma^{(3)}$ a term proportional to (15.30) to restore gauge invariance but this term breaks the symmetry between external momenta: the axial current is not conserved, an anomaly is present.

15.5.2 Divergence in the regularized theory

The calculation can be done using one of the various gauge invariant regularizations, for example Pauli–Villars's regularization or dimensional regularization with γ_5 being defined as in dimension 4 and thus no longer anticommuting with other γ matrices.

Instead we choose a regularization which preserves the symmetry between the three external arguments and global chiral symmetry, but breaks gauge invariance, replacing in the fermion propagator:

$$(\not{q})^{-1} \longmapsto (\not{q})^{-1}\rho(\varepsilon q^2),$$

where ε is the regularization parameter ($\varepsilon \to 0$), $\rho(z)$ is a positive differentiable function such that $\rho(0) = 1$, and decreasing at least like $1/z$ for $z \to +\infty$.

Then the compatibility between current conservation and gauge invariance implies that $k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k; p_1, p_2)$ vanishes.

It is convenient to consider directly the contribution $C^{(2)}(k)$ of order A^2 to $\langle \sum_\lambda k_\lambda J^5_\lambda(k) \rangle$ which sums the two diagrams:

$$C^{(2)}(k) = e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2 \sum_{\mu,\nu} A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\big(\varepsilon(q+k)^2\big)$$
$$\times\, \rho\big(\varepsilon(q-p_2)^2\big)\rho\big(\varepsilon q^2\big)\, \mathrm{tr}\left[\gamma_5 \not{k}(\not{q}+\not{k})^{-1}\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu \not{q}^{-1}\right].$$

The calculation relies on the cyclic property of the trace and the anticommutation of γ_5.

We transform the expression, using the identity

$$(\not{q})^{-1}\not{k}(\not{q}+\not{k})^{-1}=(\not{q})^{-1}-(\not{q}+\not{k})^{-1}, \tag{15.31}$$

and obtain

$$C^{(2)}(k)=e^2\int \mathrm{d}^4p_1\,\mathrm{d}^4p_2\sum_{\mu,\nu}A_\mu(p_1)A_\nu(p_2)\int\frac{\mathrm{d}^4q}{(2\pi)^4}\rho\bigl(\varepsilon(q+k)^2\bigr)\rho\bigl(\varepsilon(q-p_2)^2\bigr)$$
$$\times\,\rho\bigl(\varepsilon q^2\bigr)\,\mathrm{tr}\left\{\gamma_5\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu\left[\not{q}^{-1}-(\not{q}+\not{k})^{-1}\right]\right\}. \tag{15.32}$$

We separate the two contributions in the right hand side. In the second contribution, proportional to $(\not{q}+\not{k})^{-1}$ we interchange $(p_1,\,\mu)$ and $(p_2,\,\nu)$ and shift $q\mapsto q+p_1$. Combining again the two contributions, we find,

$$C^{(2)}(k)=e^2\int \mathrm{d}^4p_1\,\mathrm{d}^4p_2\sum_{\mu,\nu}A_\mu(p_1)A_\nu(p_2)\int\frac{\mathrm{d}^4q}{(2\pi)^4}\rho\bigl(\varepsilon(q-p_2)^2\bigr)\rho\bigl(\varepsilon q^2\bigr)$$
$$\times\,\mathrm{tr}\left[\gamma_5\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu\not{q}^{-1}\right]\left[\rho\bigl(\varepsilon(q+k)^2\bigr)-\rho\bigl(\varepsilon(q+p_1)^2\bigr)\right].$$

We see that the two terms would cancel in the absence of regulators. This corresponds to the formal proof of current conservation. However, without regularization the integrals diverge and previous manipulations are not legitimate.

By contrast, here we find a non-vanishing sum because the regulating factors which are different. After evaluation of the trace, the sum becomes

$$C^{(2)}(k) = 4e^2 \int \mathrm{d}^4 p_1 \, \mathrm{d}^4 p_2 \sum_{\mu,\nu,\rho,\sigma} A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\big(\varepsilon(q-p_2)^2\big) \rho\big(\varepsilon q^2\big) \times \epsilon_{\mu\nu\rho\sigma} \frac{p_{2\rho} q_\sigma}{q^2 (q-p_2)^2} \left[\rho\big(\varepsilon(q+p_1)^2\big) - \rho\big(\varepsilon(q+k)^2\big)\right].$$

Contributions coming from finite values of q cancel in the $\varepsilon \to 0$ limit. Due to the cut-off, the relevant values of q are of order $\varepsilon^{-1/2}$. Therefore, we can simplify the q integrand:

$$\int \frac{\mathrm{d}^4 q}{(2\pi)^4 q^4} p_{2\rho} q_\sigma \rho^2(\varepsilon q^2) \rho'(\varepsilon q^2) \sum_\lambda \left[2\varepsilon q_\lambda (p_1 - k)_\lambda\right].$$

The identity

$$\int \mathrm{d}^4 q\, q_\alpha q_\beta f(q^2) = \tfrac{1}{4}\delta_{\alpha\beta} \int \mathrm{d}^4 q\, q^2 f(q^2),$$

transforms the integral into

$$\tfrac{1}{2} p_{2\rho}(2p_1 + p_2)_\sigma \int \frac{\varepsilon \mathrm{d}^4 q}{(2\pi)^4 q^2} \rho^2(\varepsilon q^2)\rho'(\varepsilon q^2).$$

The remaining integral can be calculated explicitly (we recall $\rho(0) = 1$)

$$\int \frac{\varepsilon \mathrm{d}^4 q}{(2\pi)^4 q^2} \rho^2(\varepsilon q^2)\rho'(\varepsilon q^2) = \frac{1}{8\pi^2} \int_0^\infty \varepsilon q \mathrm{d}q\, \rho^2(\varepsilon q^2)\rho'(\varepsilon q^2) = -\frac{1}{48\pi^2},$$

and yields a result independent of the function ρ. We finally obtain

$$\sum_\lambda \left\langle k_\lambda J^5_\lambda(k) \right\rangle = -\frac{e^2}{12\pi^2} \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, p_{1\mu} A_\nu(p_1) p_{2\rho} A_\sigma(p_2). \tag{15.33}$$

From the definition (15.29) we conclude

$$\sum_{\lambda} k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{6\pi^2} \sum_{\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma} \,.$$

This non-vanishing result implies that any definition of the determinant $\det \not{D}$ breaks at least either current conservation or gauge invariance. Since gauge invariance is essential to the construction of QED we choose to break current conservation. Exchanging arguments, we obtain the value of $\sum_\mu p_{1\mu} \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)$:

$$\sum_{\mu} p_{1\mu} \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{6\pi^2} \sum_{\rho,\sigma} \epsilon_{\lambda\nu\rho\sigma} k_\rho p_{2\sigma} \,.$$

By contrast, if we had used a gauge invariant regularization, the result for $\Gamma^{(3)}$ would have differed by a term $\delta\Gamma^{(3)}$ proportional to (15.30):

$$\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = K \sum_{\rho} \epsilon_{\lambda\mu\nu\rho} (p_1 - p_2)_\rho \,.$$

The constant K then is determined by the condition of gauge invariance

$$\sum_\mu p_{1\mu}\left[\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)+\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)\right]=0\,,$$

which yields

$$\sum_\mu p_{1\mu}\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)=-\frac{e^2}{6\pi^2}\sum_{\rho,\sigma}\epsilon_{\lambda\nu\rho\sigma}k_\rho p_{2\sigma}\ \Rightarrow\ K=e^2/(6\pi^2).$$

This gives an additional contribution to the divergence of the current

$$\sum_\lambda k_\lambda\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)=\frac{e^2}{3\pi^2}\sum_{\rho,\sigma}\epsilon_{\mu\lambda\rho\sigma}p_{1\rho}p_{2\sigma}\,.$$

Therefore, in a QED-like gauge invariant field theory with massless fermions the axial current is not conserved: this is called the chiral anomaly.

For any gauge invariant regularization, one finds

$$\sum_{\lambda} k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \left(\frac{e^2}{2\pi^2} \equiv \frac{2\alpha}{\pi}\right) \sum_{\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma} \,. \tag{15.34}$$

Equation (15.34) can be rewritten, after Fourier transformation, as a non-conservation equation for the axial current:

$$\sum_{\lambda} \partial_\lambda J^5_\lambda(x) = -i\frac{\alpha}{4\pi} \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \,. \tag{15.35}$$

Since global chiral symmetry is not broken, the integral over the whole space of the anomalous term must vanish. This condition is indeed verified since the anomaly can immediately be written as a total derivative (equation (15.28). The space integral of the anomalous term depends only on the behaviour of the gauge field at boundaries, and this property indicates a relation between topology and anomalies.

Equation (15.35) also implies

$$\ln\det\left[\mathrm{e}^{i\gamma_5\theta(x)}\,\not{D}\,\mathrm{e}^{i\gamma_5\theta(x)}\right]=\ln\det\not{D}$$
$$-\,i\frac{\alpha}{4\pi}\int \mathrm{d}^4x\,\theta(x)\sum_{\mu,\nu,\rho,\sigma}\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}(x)F_{\rho\sigma}(x)+O(\theta^2).$$

Chiral gauge theory. A gauge theory is consistent only if the gauge field is coupled to a conserved current. The anomaly thus prevents the construction of a theory which would have both an Abelian gauge vector and axial symmetry, where the action in the fermion sector would read

$$\mathcal{S}(\bar\psi,\psi)=-\int \mathrm{d}^4x\,\bar\psi(x)\,(\not\partial+ie\not{A}+i\gamma_5\not{B})\,\psi(x).$$

Current conservation is a WT identity in the gauge theory for the AAB correlation function.

In such a theory the one-loop diagrams contributing to the BBB correlation function are formally identical to those contributing to AAB, because two γ_5 cancel. Therefore, they also yield an anomaly that cannot be removed since the correlation function by definition is symmetric in its three arguments. This prevents the construction even of a theory with a purely axial gauge symmetry ($e = 0$).

A way to solve both problems is to cancel the anomaly by introducing another fermion of opposite chiral coupling. With more fermions other coupling combinations are possible. However, note that in the purely axial case it is simple to show that a theory with two fermions of opposite chiral charges can be rewritten as a vector theory by combining differently the chiral components of both fermions.

15.5.3 Non-Abelian vector gauge theories and Abelian axial current

We still consider an Abelian axial current but now in the framework of a non-Abelian gauge theory. The fermion fields transform non-trivially under a gauge group G and $\mathbf{A}_\mu$ is the corresponding gauge field. The action is

$$\mathcal{S}(\bar{\psi}, \psi) = -\int \mathrm{d}^4 x \, \bar{\psi}(x) \, \not{\mathbf{D}} \, \psi(x)$$

with

$$\not{\mathbf{D}} = \not{\partial} + \not{\mathbf{A}} \,. \tag{15.36}$$

The axial current

$$J_\mu^5(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x),$$

is still gauge invariant. Therefore, no new calculation is needed.

Indeed, the result is completely determined by dimensional analysis, gauge invariance and the previous calculation which yields the term of order $\mathbf{A}^2$:

$$\sum_{\lambda} \partial_\lambda J^5_\lambda(x) = -\frac{i}{16\pi^2} \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} \operatorname{tr} \mathbf{F}_{\mu\nu} \mathbf{F}_{\rho\sigma} \,, \qquad (15.37)$$

in which $\mathbf{F}_{\mu\nu}$ is now the corresponding curvature tensor. Again this expression must be a total derivative. One verifies that

$$\sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} \operatorname{tr} \mathbf{F}_{\mu\nu} \mathbf{F}_{\rho\sigma} = 4 \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} \partial_\mu \operatorname{tr} \left(\mathbf{A}_\nu \partial_\rho \mathbf{A}_\sigma + \tfrac{2}{3} \mathbf{A}_\nu \mathbf{A}_\rho \mathbf{A}_\sigma \right). \qquad (15.38)$$

15.5.4 Anomaly and eigenvalues of the Dirac operator

We assume that the spectrum of $\not{\mathbf{D}}$, the Dirac operator in a non-Abelian gauge field (equation (15.36)), is discrete (confining temporarily the fermions in a box if necessary) and call d_n and $\varphi_n(x)$ the corresponding eigenvalues and eigenvectors:

$$\not{\mathbf{D}}\varphi_n = d_n\varphi_n\,.$$

The eigenvalues are gauge invariant, because in a gauge transformation of unitary matrix $\mathbf{g}(x)$ the Dirac operator becomes

$$\not{\mathbf{D}} \mapsto \mathbf{g}^{-1}(x)\not{\mathbf{D}}\mathbf{g}(x) \ \Rightarrow\ \varphi_n(x) \mapsto \mathbf{g}(x)\varphi_n(x).$$

For a unitary or orthogonal group, the massless Dirac operator is anti-hermitian; therefore, the eigenvalues are imaginary and the eigenvectors orthogonal. In addition we choose them with unit norm.

The anticommutation $\not{\mathbf{D}}\gamma_5 + \gamma_5\not{\mathbf{D}} = 0$ implies

$$\not{\mathbf{D}}\gamma_5\varphi_n = -d_n\gamma_5\varphi_n\,.$$

Therefore, either d_n is different from zero, and $\gamma_5 \varphi_n$ is an eigenvector of $\not{D}$ with eigenvalue $-d_n$, or d_n vanishes. The eigenspace corresponding to the eigenvalue 0 then is invariant under γ_5, which can be diagonalized: the eigenvectors of $\not{D}$ can be chosen eigenvectors of definite chirality, that is, eigenvectors of γ_5 with eigenvalue ± 1,

$$\not{D} \varphi_n = 0\,, \quad \gamma_5 \varphi_n = \pm \varphi_n\,.$$

We denote by n_+ and n_- the dimensions of the eigenspaces of positive and negative chirality, respectively.

We now consider the determinant of the operator $\not{D} + m$ regularized by mode truncation (mode regularization):

$$\det{}_N(\not{D} + m) = \prod_{n \le N} (d_n + m),$$

keeping the N lowest eigenvalues of $\not{D}$ (in modulus), with $N - n_+ - n_-$ even, in such a way that the corresponding subspace is γ_5 invariant.

The regularization is gauge invariant because the eigenvalues of $\not{D}$ are gauge invariant.

Note that in the truncated space the trace of γ_5 is the index of the Dirac operator:

$$\operatorname{tr}\gamma_5 = n_+ - n_- \,. \tag{15.39}$$

It does not vanish if $n_+ \neq n_-$, a situation which endangers axial current conservation.

In a chiral transformation (15.27) with θ constant, the determinant of $(\not{D} + m)$ becomes

$$\det{}_N(\not{D} + m) \mapsto \det{}_N\left(\mathrm{e}^{i\theta\gamma_5}(\not{D} + m)\,\mathrm{e}^{i\theta\gamma_5}\right).$$

We now consider the various eigenspaces.

If $d_n \neq 0$ the matrix γ_5 is represented by the Pauli matrix σ_1 in the sum of eigenspaces corresponding to the two eigenvalues $\pm d_n$ and $\not{D} + m$ by $d_n\sigma_3 + m$. The determinant in the subspace is then

$$\det\left(\mathrm{e}^{i\theta\sigma_1}(d_n\sigma_3 + m)\,\mathrm{e}^{i\theta\sigma_1}\right) = \det \mathrm{e}^{2i\theta\sigma_1}\det(d_n\sigma_3 + m) = m^2 - d_n^2,$$

because σ_1 is traceless.

In the eigenspace of vanishing eigenvalue $d_n = 0$ with positive chirality, of dimension n_+, γ_5 is diagonal with eigenvalue 1 and thus

$$m^{n_+} \mapsto m^{n_+}\,\mathrm{e}^{2i\theta n_+}\,.$$

Similarly, in the eigenspace $d_n = 0$ of chirality -1

$$m^{n_-} \mapsto m^{n_-}\,\mathrm{e}^{-2i\theta n_-}\,.$$

We conclude

$$\det_N\left(\mathrm{e}^{i\theta\gamma_5}(\not{D}+m)\,\mathrm{e}^{i\theta\gamma_5}\right)=\mathrm{e}^{2i\theta(n_+-n_-)}\det_N(\not{D}+m).$$

The ratio of both determinants is independent of N. Taking the limit $N\to\infty$, we find

$$\det\left[\left(\mathrm{e}^{i\gamma_5\theta}(\not{D}+m)\,\mathrm{e}^{i\gamma_5\theta}\right)(\not{D}+m)^{-1}\right]=\mathrm{e}^{2i\theta(n_+-n_-)}\,. \qquad (15.40)$$

The left hand side of equation (15.40) is 1 when $\theta=n\pi$, which implies that the coefficient of 2θ in the right hand side must indeed be an integer.

The variation of $\ln\det(\not{D}+m)$,

$$\ln\det\left[\left(\mathrm{e}^{i\gamma_5\theta}(\not{D}+m)\,\mathrm{e}^{i\gamma_5\theta}\right)(\not{D}+m)^{-1}\right]=2i\theta(n_+-n_-),$$

at first order in θ is related to the variation of the action (15.26) and thus to the expectation value of the integral of the divergence of the axial current $\left\langle\int \mathrm{d}^4x\sum_\mu \partial_\mu J^5_\mu(x)\right\rangle$.

In the limit $m = 0$ it is thus related to the space integral of the chiral anomaly (15.37):

$$-\frac{1}{32\pi^2} \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} \int \mathrm{d}^4 x \,\operatorname{tr} \mathbf{F}_{\mu\nu} \mathbf{F}_{\rho\sigma} = n_+ - n_- \,. \tag{15.41}$$

Concerning this result several comments can be made:

(i) At first order in θ in the absence of regularization we have calculated ($\ln\det = \operatorname{tr}\ln$)

$$\ln\det\left[1 + i\theta\left(\gamma_5 + (\not{D} + m)\gamma_5(\not{D} + m)^{-1}\right)\right] \sim 2i\theta \operatorname{tr}\gamma_5 \,,$$

where we have used the cyclic property of the trace. Since the trace of the matrix γ_5 vanishes we could expect naively a vanishing result. But trace here means trace in γ space and in coordinate space, and γ_5 really stands here for $\gamma_5\delta(x-y)$. The mode regularization yields a well-defined finite result for the undefined product $0 \times \delta^d(0)$.

(ii) The property that the integral (15.41) is quantized shows that the form of the anomaly is related to topological properties of the gauge field since the integral does not change when the gauge field is deformed continuously. The integral of the anomaly over the whole space thus depends only on the behaviour at large distances of the curvature tensor $F_{\mu\nu}$ and the anomaly must be a total derivative as equation (15.38) confirms.

(iii) Gauge field configurations exist for which the right hand side of equation (15.41) does not vanish, for example, instantons, which are finite action solutions of imaginary time classical field equations, related to barrier penetration effects. We have shown above that if massless fermions are coupled to such gauge fields the determinant resulting from the fermion integration necessarily vanishes. This has some physical implications which are examined in section 15.7.

(iv) One might be surprised that $\det \not{D}$ is not invariant under global chiral transformations. However, we have just established that when the integral of the anomaly does not vanish, $\det \not{D}$ vanishes. This explains that to give a meaning to the right hand side of equation (15.40) we have been forced to introduce a mass to find a non-trivial result. The determinant of $\not{D}$ in the subspace orthogonal to eigenvectors with vanishing eigenvalue, even in presence of a mass, is chiral invariant by parity doubling, but for $n_+ \neq n_-$ not the determinant in the eigenspace of eigenvalue zero because the trace of γ_5 does not vanish in the eigenspace (equation (15.39)). In the limit $m \to 0$ the complete determinant vanishes but not the ratio of determinants for different values of θ because the powers of m cancel.

15.6 Non-Abelian anomaly

We first consider the problem of conservation of a general axial current in a non-Abelian vector gauge theory, and then the issue of obstruction to gauge invariance in chiral gauge theories.

15.6.1 General axial current

We now discuss the problem of the conservation of a general axial current in the example of a fermion action which has a $G \times G$ chiral symmetry (subgroup of $U(N) \times U(N)$), in the background of non-Abelian vector gauge fields. The generators of the gauge group may or may not be related to the diagonal subgroup G of $G \times G$ which correspond to vector currents.

We denote by t^α the generators of G. The current then has the form

$$J_\mu^{5\alpha}(x) = -\bar{\psi}\gamma_5\gamma_\mu t^\alpha \psi \,.$$

When the gauge group is connected with the chiral group, the current conservation equation involves the gauge covariant derivative ($\mathbf{D}_\mu = \partial_\mu + [A_\mu, \bullet]$):

$$\sum_\mu \mathbf{D}_\mu \mathbf{J}^5_\mu = 0\,.$$

In the calculation of the contribution to the anomaly coming from terms quadratic in the gauge fields the only modification in the previous results is the appearance of a different geometrical factor. Then the complete form of the anomaly is dictated by gauge covariance. One finds

$$\sum_\lambda \mathbf{D}_\lambda J^{5\alpha}_\lambda(x) = -\frac{i}{16\pi^2} \sum_{\mu,\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma}\, \mathrm{tr}\, t^\alpha \mathbf{F}_{\mu\nu} \mathbf{F}_{\rho\sigma}\,.$$

In particular, if the gauge group is disconnected from the chiral group the anomaly is proportional to $\mathrm{tr}\, t^\alpha$ and, therefore, only different from zero for the Abelian factors of G.

15.6.2 Obstruction to gauge invariance

We now want to consider a non-Abelian gauge field coupled to left or right-handed fermions:

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}^4x\, \bar{\psi}(x)\tfrac{1}{2}(1+\gamma_5)\not{\mathbf{D}}\psi(x),$$

($\frac{1}{2}(1-\gamma_5)$ is treated in the same way). We can construct a consistent gauge theory only if the partition function

$$\mathcal{Z}(\mathbf{A}_\mu) = \int \left[\mathrm{d}\psi \mathrm{d}\bar{\psi}\right] \exp\left[-\mathcal{S}(\psi,\bar{\psi})\right]$$

is gauge invariant.

If we introduce the generators $\mathbf{t}^\alpha$ of the gauge group in the fermion representation, we can write the corresponding current $\mathbf{J}_\mu$ as

$$J^\alpha_\mu(x) = -\bar{\psi}\tfrac{1}{2}(1+\gamma_5)\gamma_\mu \mathbf{t}^\alpha \psi\,.$$

The invariance of $\mathcal{Z}(\mathbf{A}_\mu)$ under an infinitesimal gauge transformation again leads to a covariant conservation equation for the current:

$$\sum_\mu \langle \mathbf{D}_\mu \mathbf{J}_\mu \rangle = 0\,.$$

The calculation of the term of degree two in the gauge field of the anomaly is straightforward: the regularization adopted for the calculation in section 15.5.2 is also suited to the present case since the current-gauge field three-point function is symmetric in the external arguments. The group structure yields a simple geometrical factor. The global factor can be taken from the Abelian calculation. It differs from the result (15.33) by a factor $1/2$ which comes from the projector $\frac{1}{2}(1+\gamma_5)$. The general form of the term of third degree in the gauge field can also easily be found, but the calculation of the global factor is somewhat tedious. In fact, it can be derived from Wess–Zumino consistency conditions.

The complete expression reads:

$$\sum_{\mu} (\mathbf{D}_\mu \mathbf{J}_\mu(x))^\alpha = -\frac{i}{24\pi^2} \sum_{\mu,\nu,\rho,\sigma} \partial_\mu \epsilon_{\mu\nu\rho\sigma} \operatorname{tr}\left[t^\alpha \left(\mathbf{A}_\nu \partial_\rho \mathbf{A}_\sigma + \tfrac{1}{2}\mathbf{A}_\nu \mathbf{A}_\rho \mathbf{A}_\sigma\right)\right]. \tag{15.42}$$

If the projector $\frac{1}{2}(1+\gamma_5)$ is replaced by $\frac{1}{2}(1-\gamma_5)$ the sign of the anomaly changes.

Unless this term vanishes identically there is an obstruction to the construction of the gauge theory. It is easy to verify, taking into account the antisymmetry of the ϵ tensor, that the group factor is

$$d_{\alpha\beta\gamma} = \tfrac{1}{2} \operatorname{tr}\left[t^\alpha \left(t^\beta t^\gamma + t^\gamma t^\beta\right)\right].$$

For a unitary representation the generators t^α are, with our conventions, antihermitian. Therefore, the coefficients $d_{\alpha\beta\gamma}$ are purely imaginary:

$$d^*_{\alpha\beta\gamma} = \tfrac{1}{2} \operatorname{tr}\left[t^\alpha \left(t^\beta t^\gamma + t^\gamma t^\beta\right)\right]^* = -d_{\alpha\beta\gamma}.$$

For all real (the t^α antisymmetric) or 'pseudo-real' ($t^\alpha = -S\,(t^\alpha)^T S^{-1}$) representations these coefficients vanish. It follows that the only non-Abelian groups which can lead to anomalies in four dimensions are $SU(N)$ for $N \geq 3$, $SO(6)$ and E_6.

15.7 Physics implications

Weak–electromagnetic (w.e.m.) interactions and anomaly cancellation. The condition of anomaly cancellation discussed in section 15.6.2 constrains the model of w.e.m. interactions. In the Standard Model, for example, the anomalous contributions of leptons cancels the quark contributions. This cancellation occurs within each generation, as we now show, provided that for each flavour quarks exist in three states. In the w.e.m. group $SU(2) \times U(1)$, $SU(2)$ alone is a safe group. Therefore, the problems come from the $U(1)$ factor. We expect *a priori* two conditions coming from the vertices with one $U(1)$ and two $SU(2)$ gauge fields and with three $U(1)$ gauge fields. Actually one discovers that both conditions are equivalent.

If we consider two $SU(2)$ and one $U(1)$ gauge fields, only $SU(2)$ doublets contribute and equation (15.42) leads to the condition:

$$\sum_{\text{all doublets}} Y_{\mathrm{L}} \operatorname{tr} \tau^{\alpha} \tau^{\beta} = 0\,,$$

in which Y_{L} is the $U(1)$ charge (see section 15.1). This condition reduces to

$$\sum_{\text{all doublets}} Y_{\mathrm{L}} = 0\,. \tag{15.43}$$

The vertex with three $U(1)$ gauge fields yields the condition:

$$\sum_{\text{left-handed parts}} Y_{\mathrm{L}}^{3} - \sum_{\text{right-handed parts}} Y_{\mathrm{R}}^{3} = 0\,,$$

because the contributions to the anomaly of right-handed and left-handed couplings have opposite signs.

In the Standard Model the left and right charges are related (equation (15.12)). Summing the charges of one doublet and the corresponding singlets, we obtain

$$\sum_{\text{all doublets}} (Y_{\mathrm{L}} + 1)^3 + (Y_{\mathrm{L}} - 1)^3 - 2Y_{\mathrm{L}}^3 = 0\,,$$

a condition which reduces to equation (15.43).

In one generation the lepton doublet has $Y_{\mathrm{L}} = -1$ and the quark $Y_{\mathrm{L}} = 1/3$. Therefore, a cancellation requires that the quarks exist in three states. These states are provided by the colour quantum number.

Electromagnetic π_0 decay. In an effective low energy field theory for Strong Interactions (see section 12.5) based on a linearly broken $SU(2) \times SU(2)$ symmetry, where hadrons are considered as elementary fields, the non-conservation of the axial current $\mathbf{J}^5_\mu$ is at leading order expressed by the equation

$$\sum_\mu \partial_\mu \mathbf{J}^5_\mu = m_\pi^2 f_\pi \boldsymbol{\pi} \,. \tag{15.44}$$

We concentrate here on the third component $[J^5_\mu]_3$ of the current which corresponds in the right hand side to the neutral pion π_0 field. After introduction of electromagnetic interactions in the model, the relation between the divergence of the axial current and the π_0 field allows to calculate the electromagnetic decay rate of the neutral pion when the four-momentum $\mathbf{k}$ of the pion goes to zero.

In the absence of anomalies, the expectation value of relation (15.44) multiplied by two photon fields implies that the decay rate vanishes for $\mathbf{k} = 0$ in contradiction with reasonable smoothness assumptions and experimental results.

By contrast, taking into account the anomaly equation (15.35) one finds

$$\sum_{\mu} \partial_\mu [J^5_\mu]_3 = m_\pi^2 f_\pi \pi_0 - i\frac{\alpha}{8\pi} \sum_{\mu,\nu,\rho,\sigma,} \epsilon_{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \,.$$

Multiplying the equation by two photon fields, taking the expectation value and going to the limit $\mathbf{k} = 0$ to eliminate the left hand side, one now obtains a non-vanishing decay amplitude for an unphysical π_0 at zero four momentum.

In the σ-model at leading order one can extrapolate to $k^2 = -m_\pi^2$. The theoretical rate Γ is given by

$$\Gamma = \frac{\alpha^2 m_\pi^3}{64\pi^3 f_\pi^2} = 7.6 \text{ eV},$$

while $\Gamma^{\text{exp}} = (7.37 \pm 1.5)$ eV. The theoretical result is in excellent agreement with experiment. A similar estimate was first derived by Steinberger from direct Feynman graph calculation, before the relation to anomalies had been discovered.

Note that an analogous theoretical estimate is obtained in the quark model with massless quarks, for three colours.

The solution of the $U(1)$ problem. In a theory of Strong Interactions with massless quarks interacting through a colour gauge group, the action has a chiral $U(N_\mathrm{F})\times U(N_\mathrm{F})$ symmetry, in which N_F is the number of flavours. The spontaneous breaking of the chiral group to its diagonal subgroup $U(N_\mathrm{F})$ leads to expect N_F^2 Goldstone bosons associated with the axial currents. From the preceding analysis we know that the axial current corresponding to the $U(1)$ Abelian subgroup has an anomaly. Of course the WT identities that imply the existence of Goldstone bosons correspond to constant group transformations and, therefore, involve only the space integral of the divergence of the current. Since the anomaly is a total derivative one might have expected the integral to vanish.

However, non-Abelian gauge theories admit instanton solutions which give a periodic structure to the vacuum (like in the case of the one-dimensional cosine potential).

These instanton solutions correspond to gauge configurations which approach non-trivial pure gauges at infinity and give the set of discrete non-vanishing values one expects from equation (15.41) to the space integral of the anomaly (15.37). This indicates (but no satisfactory calculation of the instanton contribution has been performed) that for small, but non-vanishing, quark masses the $U(1)$ axial current is far from being conserved and, therefore, no light would-be Goldstone boson is generated. This observation resolves a long standing puzzle since experimentally no corresponding light pseudoscalar boson is found for $N_F = 2, 3$.