To order $p_\perp$, the left-handed spinors are

$$\xi(p) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \xi(k) = \begin{pmatrix} p_\perp/2(1-z)p \\ 1 \end{pmatrix}. \tag{17.88}$$

The polarization vectors for the photon are

$$\epsilon_L^{*i}(q) = \frac{1}{\sqrt{2}}\Big(1, i, -\frac{p_\perp}{zp}\Big), \qquad \epsilon_R^{*i}(q) = \frac{1}{\sqrt{2}}\Big(1, -i, -\frac{p_\perp}{zp}\Big). \tag{17.89}$$

Notice that, when these vectors are contracted with the Pauli matrix in Eq. (17.87), the first two components of the right-handed polarization vector give $(\sigma^1 - i\sigma^2) = 2\sigma^-$, which annihilates $\xi(p)$. The only remaining term comes from the $i = 3$ component, and we find

$$i\mathcal{M}(e_L^- \to e_L^- \gamma_R) = ie\frac{\sqrt{2(1-z)}}{z}p_\perp. \tag{17.90}$$

For the left-handed photon polarization, there is an additional contribution from the first two components of ϵ_L^*. These add to

$$i\mathcal{M}(e_L^- \to e_L^- \gamma_L) = ie\frac{\sqrt{2(1-z)}}{z(1-z)}p_\perp. \tag{17.91}$$

Parity invariance implies that the values of the matrix elements are unchanged if all helicities are flipped; this immediately gives the required matrix elements for the case of an initial e_R^-. The squared matrix element, averaged over initial helicities, is therefore

$$\frac{1}{2}\sum_{\text{pols.}} |\mathcal{M}|^2 = \frac{2e^2 p_\perp^2}{z(1-z)}\left[\frac{1 + (1-z)^2}{z}\right]. \tag{17.92}$$

The first term in the brackets comes from a photon with spin parallel to the electron spin; the second term comes from a photon with spin opposite to the electron spin.

The Equivalent Photon Approximation

Now we have all the pieces needed to compute the cross sections for the processes shown in Fig. 17.15. We first consider the process with a virtual photon. Call the initial state on the right-hand side of the diagram X and the final state Y, and let $\mathcal{M}_{\gamma X}$ represent the matrix element for the scattering of the photon from X. We will assume for simplicity that X is unpolarized, so that the scattering cross section does not depend on the virtual photon polarization. Then the complete diagram gives a cross section

$$\sigma = \frac{1}{(1+v_X)2p2E_X}\int \frac{d^3k}{(2\pi)^3}\frac{1}{2k^0}\int d\Pi_Y \left[\frac{1}{2}\sum |\mathcal{M}|^2\right]\Big(\frac{1}{q^2}\Big)^2 |\mathcal{M}_{\gamma X}|^2, \tag{17.93}$$

where v_X is the velocity of X and $\int d\Pi_Y$ is the phase space integral over Y.

The integral has a singularity when k is collinear with the incident electron momentum p. To isolate the singularity, substitute for k^0 and q^2 from Eqs.

(17.82) and (17.85) and rewrite the integral over k as

$$d^3k = dk^3 d^2k_\perp = pdz \cdot \pi dp_\perp^2. \qquad (17.94)$$

Then the cross section can be expressed as

$$\begin{aligned}\sigma &= \int \frac{pdzdp_\perp^2}{16\pi^2(1-z)p}\left[\frac{1}{2}\sum|\mathcal{M}|^2\right]\frac{(1-z)^2}{p_\perp^4}\frac{z}{(1+v_X)2zp2E_X}\int d\Pi_Y|\mathcal{M}_{\gamma X}|^2 \\ &= \int \frac{dzdp_\perp^2}{16\pi^2(1-z)}\left[\frac{1}{2}\sum|\mathcal{M}|^2\right]\frac{z(1-z)^2}{p_\perp^4}\cdot\sigma(\gamma X\to Y). \qquad (17.95)\end{aligned}$$

Finally, insert the spin-averaged electron emission vertex (17.92), to obtain

$$\begin{aligned}\sigma &= \int \frac{dzdp_\perp^2}{16\pi^2}\frac{z(1-z)}{p_\perp^4}\frac{2e^2p_\perp^2}{z(1-z)}\left[\frac{1+(1-z)^2}{z}\right]\cdot\sigma(\gamma X\to Y) \\ &= \int_0^1 dz\int\frac{dp_\perp^2}{p_\perp^2}\frac{\alpha}{2\pi}\left[\frac{1+(1-z)^2}{z}\right]\cdot\sigma(\gamma X\to Y). \qquad (17.96)\end{aligned}$$

The integral over $p_\perp^2$ runs from momentum transfers of order s down to the electron mass m^2, which cuts off the singularity. Thus, our final result is

$$\sigma(e^-X\to e^-Y) = \int_0^1 dz\,\frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+(1-z)^2}{z}\right]\cdot\sigma(\gamma X\to Y). \qquad (17.97)$$

The cross section on the right-hand side is computed for a real, transversely polarized photon of momentum zp. The factor $\log(s/m^2)$ represents the mass singularity. This formula is the Weizsäcker-Williams *equivalent photon approximation*, which we encountered earlier in Problems 5.5 and 6.2.

Formula (17.97) takes on a new significance when it is juxtaposed with the QCD predictions of the previous two sections. This QED formula has just the same form as a parton model expression, with the Weizsäcker-Williams distribution function

$$f_\gamma(z) = \frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+(1-z)^2}{z}\right] \qquad (17.98)$$

playing the role of the probability to find a photon of longitudinal fraction z in the incident electron.

The Electron Distribution

The first diagram of Fig. 17.15, with an emitted photon and a virtual electron. can be treated in the same way. The analogue of (17.93) is

$$\sigma = \frac{1}{(1+v_X)2p2E_X}\int\frac{d^3q}{(2\pi)^3}\frac{1}{2q^0}\int d\Pi_Y\left[\frac{1}{2}\sum|\mathcal{M}|^2\right]\left(\frac{1}{k^2}\right)^2|\mathcal{M}_{e^-X}|^2.$$

Following the steps that led to (17.97), we find

$$\begin{aligned}\sigma(e^-X\to\gamma Y) &= \int\frac{dz\,dp_\perp^2}{16\pi^2 z}\left[\frac{1}{2}\sum|\mathcal{M}|^2\right]\frac{z^2}{p_\perp^4}\cdot(1-z)\sigma(e^-X\to Y)\\ &= \int\frac{dz\,dp_\perp^2}{16\pi^2}\,\frac{z(1-z)}{p_\perp^4}\,\frac{2e^2p_\perp^2}{z(1-z)}\left[\frac{1+(1-z)^2}{z}\right]\cdot\sigma(e^-X\to Y)\\ &= \int_0^1 dz\int\frac{dp_\perp^2}{p_\perp^2}\,\frac{\alpha}{2\pi}\left[\frac{1+(1-z)^2}{z}\right]\cdot\sigma(e^-X\to Y),\end{aligned}\tag{17.99}$$

where the intermediate electron carries a longitudinal fraction $(1-z)$.

It is tempting to substitute $x=(1-z)$ and interpret the factor multiplying the cross section under the integral in (17.99) as the parton distribution for finding an electron parton in the electron. This would give

$$f_e^{(1)}(x)=\frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+x^2}{1-x}\right].\tag{17.100}$$

However, this expression is not adequate. Most obviously, it does not take into account the processes without radiation, in which the electron remains an electron at the full energy. This is easily remedied by considering (17.100) as the order-α correction to the most naive expectation,

$$f_e^{(0)}(x)=\delta(1-x),\tag{17.101}$$

in which we consider the electron to contain only an electron at the full energy. Unfortunately, the sum of (17.101) and (17.100) still does not give an adequate description of the electron distribution, for two reasons. First, Eq. (17.100) diverges near $x=1$, and we need a prescription for treating this singularity. Second, while Eq. (17.100) takes into account the virtual electrons moved to longitudinal fraction x from $x=1$ by radiation, it does not take into account the concomitant loss of electrons from the delta function peak at $x=1$.

The divergence of (17.100) at $x=1$ corresponds to the emission of soft photons. We saw in Section 6.5 that the emission of soft photons does not affect the rate of a QED reaction. Order by order in α, one finds that infrared-divergent positive contributions to the total rate from the emission of soft photons are balanced by negative contributions from diagrams with soft virtual photons. In the present example, the negative contribution must decrease the weight of the process in which no photon is emitted. Thus, to order α, the parton distribution for electrons in the electron should have the form

$$f_e(x)=\delta(1-x)+\frac{\alpha}{2\pi}\log\frac{s}{m^2}\left(\frac{1+x^2}{(1-x)}-A\delta(1-x)\right).\tag{17.102}$$

The coefficient A results from diagrams with virtual photons that we have not computed. However, the effect of these diagrams is easy to understand; they subtract from the delta function the probability that has been moved to lower x by radiation, so that the integral over the full term of order α is zero. Another way of expressing this criterion is that A is determined by the condition that the electron contain exactly one electron parton,

$$\int_0^1 dx\, f_e(x) = 1. \tag{17.103}$$

(This equation will be modified below, when we include pair-creation processes.)

It is not so clear how to integrate over the singular denominator in (17.100) to determine A explicitly. It is conventional to define a distribution that can be integrated by subtracting a delta function from the singular term. Define the distribution

$$\frac{1}{(1-x)_+} \tag{17.104}$$

to agree with the function $1/(1-x)$ for all values of x less than 1, and to have a singularity at $x = 1$ such that the integral of this distribution with any smooth function $f(x)$ gives

$$\int_0^1 dx\, \frac{f(x)}{(1-x)_+} = \int_0^1 dx\, \frac{f(x) - f(1)}{(1-x)}. \tag{17.105}$$

Less formally,

$$\frac{1}{(1-x)_+} = \lim_{\epsilon\to 0}\left[\frac{1}{(1-x)}\theta(1-x-\epsilon) - \delta(1-x)\int_0^{1-\epsilon} dx' \frac{1}{(1-x')}\right]. \tag{17.106}$$

The more formal definition (17.105) is often easier to use in practice.

Using this definition, we can bring a piece of the delta function into the singular term of (17.102) by changing the denominator $(1-x)$ to $(1-x)_+$. Then, to normalize (17.102), we need the integral

$$\int_0^1 dx\, \frac{1+x^2}{(1-x)_+} = \int_0^1 dx\, \frac{x^2-1}{(1-x)} = -\frac{3}{2}.$$

Our final form of the electron distribution, to order α, is

$$f_e(x) = \delta(1-x) + \frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x)\right]. \tag{17.107}$$

This distribution is now properly normalized, but it is still highly singular near $x = 1$. Thus, we should expect higher-order corrections to the electron distribution function to be important in this region. We must, then, think about how to treat the emission of many collinear photons.

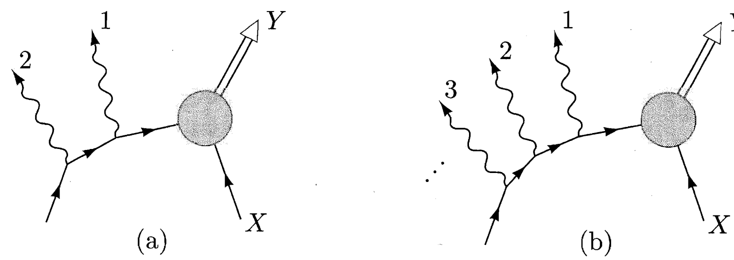


Figure 17.17. Higher-order diagrams with collinear photon emission: (a) two collinear photons; (b) many collinear photons.

Multiple Splittings

In fact, it is not difficult to extend the analysis we have just completed to account for emission of many collinear photons. Consider the process shown in Fig. 17.17(a), in which photon 1 is radiated with a transverse momentum $p_{1\perp}$ and photon 2 with a transverse momentum $p_{2\perp}$. The emission of photon 2 can be computed just as we did above. If $p_{2\perp} \ll p_{1\perp}$, the first virtual electron is very close to mass shell compared to $p_{1\perp}^2$ and so we can ignore its virtuality in computing the emission of the photon 1. The double photon emission gives a contribution of order

$$\left(\frac{\alpha}{2\pi}\right)^2 \int_{m^2}^{s} \frac{dp_{1\perp}^2}{p_{1\perp}^2} \int_{m^2}^{p_{1\perp}^2} \frac{dp_{2\perp}^2}{p_{2\perp}^2} = \frac{1}{2}\left(\frac{\alpha}{2\pi}\right)^2 \log^2 \frac{s}{m^2}.$$

In the opposite limit, $p_{2\perp} \gg p_{1\perp}$, there is no denominator of order $p_{1\perp}^2$, and so we do not find a double logarithm. Only in the case $p_{2\perp} \ll p_{1\perp}$ can the contribution of order α^2 compete with the contribution of order α.

This argument extends to the emission of arbitrarily many collinear photons, Fig. 17.17(b). The region of integration over the photon phase space corresponding to the ordering

$$p_{1\perp} \gg p_{2\perp} \gg p_{3\perp} \gg \cdots \tag{17.108}$$

gives a contribution that contains the factor

$$\frac{1}{n!}\left(\frac{\alpha}{2\pi}\right)^n \log^n \frac{s}{m^2}. \tag{17.109}$$

If the photon transverse momenta are ordered in any other way, the contribution from that region contains at least one less power of the large logarithm at the same order in α. If condition (17.108) holds, the virtual electron momenta are increasingly off-mass-shell as one proceeds from the outside of the diagram toward the hard collision. In this case, the electron momenta are said to be *strongly ordered.*

This set of conclusions has an interesting physical interpretation. Since the intermediate electrons are increasingly virtual as we go into the diagram, it is natural to interpret them as components of the physical electron when this particle is analyzed at successively smaller distance scales. The intermediate electron with $k^2 \sim p_\perp^2$ may be thought of as a constituent of the electron made visible when the wavefunction of the physical electron is probed with a resolution $\Delta r \sim (p_\perp)^{-1}$. In this picture, the electron seen at one resolution can be resolved at a finer scale into a more virtual electron and a number of photons.

From either the perspective of computing Feynman diagrams or the grander perspective of the electron structure, it is useful to imagine the splitting of the electron into a virtual electron plus photons to be a continuous evolution process as a function of the transverse momentum of the electron constituent. To describe this process mathematically, we introduce an explicit $p_\perp$ dependence of the electron and photon distribution functions. We define the functions $f_\gamma(x, Q)$ and $f_e(x, Q)$ to give the probabilities of finding a photon or an electron of longitudinal fraction x in the physical electron, taking into account collinear photon emissions with transverse momenta $p_\perp < Q$. If Q is slightly increased to $Q + \Delta Q$, we must take into account the possibility that an electron constituent in $f_e(x, Q)$ will radiate a photon with $Q < p_\perp < Q + \Delta Q$. The differential probability for an electron to split off a photon that carries away a fraction z of its energy is

$$\frac{\alpha}{2\pi}\frac{dp_\perp^2}{p_\perp^2}\frac{1+(1-z)^2}{z}. \tag{17.110}$$

The new photon distribution can therefore be computed as follows:

$$\begin{aligned} f_\gamma(x, Q{+}\Delta Q) & \\ &= f_\gamma(x,Q) + \int_0^1 dx' \int_0^1 dz \left[\frac{\alpha}{2\pi}\frac{\Delta Q^2}{Q^2}\frac{1+(1-z)^2}{z}\right] f_e(x', p_\perp)\delta(x - zx') \\ &= f_\gamma(x,Q) + \frac{\Delta Q}{Q}\int_x^1 \frac{dz}{z}\left[\frac{\alpha}{\pi}\frac{1+(1-z)^2}{z}\right] f_e(\frac{x}{z}, p_\perp). \end{aligned} \tag{17.111}$$

Passing to a continuous evolution, we find that the function $f_\gamma(x, Q)$ is determined by the integral-differential equation

$$\frac{d}{d\log Q} f_\gamma(x,Q) = \int_x^1 \frac{dz}{z}\left[\frac{\alpha}{\pi}\frac{1+(1-z)^2}{z}\right] f_e(\frac{x}{z}, Q). \tag{17.112}$$

Similarly, the distribution of component electrons in the physical electron will evolve with Q, reflecting the appearance of electrons at lower values of x due to photon radiation, and the disappearance of these electrons at higher x. The term in brackets in Eq. (17.107) gives a correct accounting of both effects for the radiation of a single photon. Thus, the electron distribution evolves according to

$$\frac{d}{d\log Q} f_e(x,Q) = \int_x^1 \frac{dz}{z}\left[\frac{\alpha}{\pi}\left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right)\right] f_e(\frac{x}{z},Q). \qquad (17.113)$$

By integrating these integral-differential equations using appropriate initial conditions, we sum all of the logarithmically enhanced terms of the form (17.109). The initial conditions should be fixed at a point that will reproduce the correct denominator of the logarithms in Eqs. (17.98) and (17.107). Thus, we should set

$$f_e(x,Q) = \delta(1-x), \qquad f_\gamma(x,Q) = 0, \qquad (17.114)$$

at $Q^2 \sim m^2$.

The resulting distribution functions can be used to compute the cross sections for electron hard scattering from an arbitrary target. Then Eqs. (17.97) and (17.99) should be replaced by

$$\begin{aligned} \sigma(e^- X \to e^- + n\gamma + Y) &= \int_0^1 dx\, f_\gamma(x,Q)\sigma(\gamma X \to Y), \\ \sigma(e^- X \to n\gamma + Y) &= \int_0^1 dx\, f_e(x,Q)\sigma(e^- X \to Y), \end{aligned} \qquad (17.115)$$

where the cross sections under the integrals are computed for a photon or electron carrying a fraction x of the original electron momentum, the functions $f_\gamma(x,Q)$, $f_e(x,Q)$ are the solutions to Eqs. (17.112) and (17.113), and the momentum Q is chosen as a characteristic momentum transfer of the γX or $e^- X$ subprocess.

Photon Splitting to Pairs

The evolution equations for $f_\gamma(x)$ and $f_e(x)$ need one more modification before they can be considered complete. As written, these equations account for the radiation of photons by electrons to all orders. However, they omit another process that is of the same order in α: the splitting of a photon into an electron-positron pair. We must include this process in our evolution equations, because the process shown in Fig. 17.18, for example, has the same logarithmic enhancement as that shown in Fig. 17.17(a).

We can compute the effects of photon splitting in the same way that we computed with effects of photon radiation. The basic kinematics of the process is very similar, as shown in Fig. 17.19; the only difference is that the photon is now in the initial state, while the final state consists of an almost collinear electron-positron pair. We need to work out the analogue of Eq. (17.92) for this process.

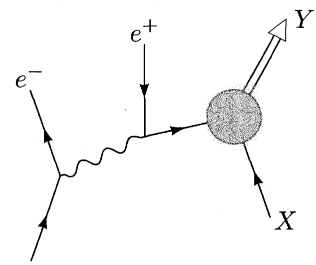


Figure 17.18. A process that involves e^+e^- pair creation enhanced by a collinear mass singularity.

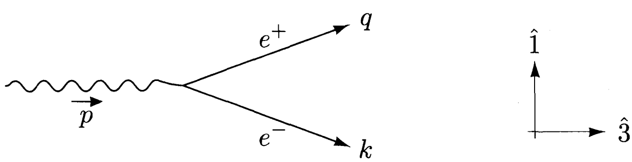


Figure 17.19. Kinematics of the vertex for photon conversion to a collinear electron-positron pair.

Consider the case in which the outgoing electron is left-handed. Then the outgoing positron must be right-handed, by helicity conservation; its spin wavefunction will contain a left-handed spinor. Let us take the electron momentum to be k, given by Eq. (17.82), and the positron momentum to be q. Then the vertex gives a matrix element

$$i\mathcal{M} = -ie\bar{u}_L(k)\gamma_\mu v_L(q)\epsilon_T^\mu(p), \tag{17.116}$$

where the photon polarization vector can be either left- or right-handed. When we insert the explicit forms for the massless spinors, we obtain

$$i\mathcal{M} = ie\sqrt{2(1-z)p}\sqrt{2zp}\,\xi^\dagger(k)\sigma^i\xi(q)\cdot\epsilon_T^i(p),$$

where the electron and positron spinors are given, to order $p_\perp$, by

$$\xi(q) = \begin{pmatrix} -p_\perp/2zp \\ 1 \end{pmatrix}, \qquad \xi(k) = \begin{pmatrix} p_\perp/2(1-z)p \\ 1 \end{pmatrix}.$$

The polarization vectors for the photon are

$$\epsilon_L^i(p) = \frac{1}{\sqrt{2}}(1,-i,0), \qquad \epsilon_R^i(p) = \frac{1}{\sqrt{2}}(1,i,0).$$

Dotting these vectors into σ^i, we find for the polarized matrix elements

$$i\mathcal{M}(\gamma_L \to e_L^- e_R^+) = -ie\frac{\sqrt{2z(1-z)}}{z}p_\perp,$$

and

$$i\mathcal{M}(\gamma_R \to e_L^- e_R^+) = +ie\frac{\sqrt{2z(1-z)}}{(1-z)}p_\perp.$$

Again, the matrix elements are unchanged if all helicities are flipped. Thus the squared matrix element, averaged over initial photon polarizations, is

$$\frac{1}{2}\sum_{\text{pols.}} |\mathcal{M}|^2 = \frac{2e^2p_\perp^2}{z(1-z)}\left[z^2 + (1-z)^2\right], \tag{17.117}$$

where z is the momentum fraction carried by the positron. The first term in the brackets comes from processes in which the spin of the positron is parallel to the spin of the photon; the second term comes from processes where the electron spin is parallel to the photon spin.

The squared matrix element (17.117) generates an evolution of constituent photons into electrons and positrons. The form of the evolution equation is similar to (17.113), but with the photon distribution on the right-hand side, and with the expression in parentheses replaced by

$$\left(z^2 + (1-z)^2\right). \tag{17.118}$$

When we create an electron-positron pair, we must remove a photon; this requires a negative term in the evolution equation for the photon distribution (17.112) that contains a delta function multiplying the normalization of (17.118):

$$\int_0^1 dz\left(z^2 + (1-z)^2\right) = \frac{2}{3}. \tag{17.119}$$

Evolution Equations for QED

Including the effects of pair creation, we find the complete evolution equations for electron, positron, and photon distributions in QED. These equations, originally derived by Gribov and Lipatov, sum the leading logarithms from collinear singularities to all orders in α. The evolution equations take the form

$$\begin{aligned}
\frac{d}{d\log Q}f_\gamma(x,Q) &= \frac{\alpha}{\pi}\int_x^1 \frac{dz}{z}\Big\{P_{\gamma\leftarrow e}(z)\left[f_e(\tfrac{x}{z},Q) + f_{\bar e}(\tfrac{x}{z},Q)\right] \\
&\qquad\qquad + P_{\gamma\leftarrow\gamma}(z)f_\gamma(\tfrac{x}{z},Q)\Big\}, \\
\frac{d}{d\log Q}f_e(x,Q) &= \frac{\alpha}{\pi}\int_x^1 \frac{dz}{z}\Big\{P_{e\leftarrow e}(z)f_e(\tfrac{x}{z},Q) + P_{e\leftarrow\gamma}(z)f_\gamma(\tfrac{x}{z},Q)\Big\}, \\
\frac{d}{d\log Q}f_{\bar e}(x,Q) &= \frac{\alpha}{\pi}\int_x^1 \frac{dz}{z}\Big\{P_{e\leftarrow e}(z)f_{\bar e}(\tfrac{x}{z},Q) + P_{e\leftarrow\gamma}(z)f_\gamma(\tfrac{x}{z},Q)\Big\}.
\end{aligned} \tag{17.120}$$

The *splitting functions* $P_{i\leftarrow j}(z)$ are given by

$$\begin{aligned}
P_{e\leftarrow e}(z) &= \frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z), \\
P_{\gamma\leftarrow e}(z) &= \frac{1+(1-z)^2}{z}, \\
P_{e\leftarrow\gamma}(z) &= z^2 + (1-z)^2, \\
P_{\gamma\leftarrow\gamma}(z) &= -\frac{2}{3}\delta(1-z).
\end{aligned} \tag{17.121}$$

To obtain the distribution functions for an electron relevant to a given momentum transfer Q, we should integrate these equations with the initial conditions

$$f_e(x,Q) = \delta(1-x), \qquad f_{\bar{e}}(x,Q) = 0, \qquad f_\gamma(x,Q) = 0, \tag{17.122}$$

at $Q = m$. With different initial conditions, the same equations give the distribution functions for a physical positron or photon. The solutions to these equations are used as in Eq. (17.115) to compute cross sections involving processes induced by electrons, positrons, or photons that involve large momentum transfer.

The evolution equations (17.120) are constructed in such a way as to conserve electron number and longitudinal momentum. Thus, the basic sum rules (17.36) and (17.39) satisfied by the parton distributions of hadrons also apply to the QED distribution functions. Specifically, the distribution functions of the electron contain one net electron constituent,

$$\int_0^1 dx\big[f_e(x,Q) - f_{\bar{e}}(x,Q)\big] = 1, \tag{17.123}$$

and account for the total momentum of the physical electron,

$$\int_0^1 dx\, x\big[f_e(x,Q) + f_{\bar{e}}(x,Q) + f_\gamma(x,Q)\big] = 1. \tag{17.124}$$

It is an instructive exercise to verify explicitly, using Eq. (17.120), that the values of these integrals do not depend on Q.

The Altarelli-Parisi Equations

If we encounter mass singularities in QED associated with collinear photon emission, we must also encounter mass singularities in QCD associated with collinear gluon and quark emission. If we compute the corrections of order α_s to the leading-order parton cross sections discussed in Sections 17.3 and 17.4. using massless quarks and gluons, we will find that these correction terms diverge when we integrate over the collinear configurations. Thus the parton-model expressions, at least in their simplest form, break down already at the next-to-leading order in α_s.

However, assuming that the singularities of QCD are no worse than those of QED, the considerations of the previous section tell us how to treat these singular terms. In QED, we found it natural to include the large corrections associated with the mass singularities in the parton distributions rather than in the hard-scattering cross sections. Viewed in this way, the singular terms supply the kernel of an evolution equation for the parton distributions as a function of the logarithm of the momentum scale. Hard scattering with a momentum transfer Q probes the electron at a distance of order Q^{-1}. When the electron wavefunction is resolved to very small scales, it appears as a constituent electron, carrying only a fraction of the total longitudinal momentum, plus a number of constituent photons and electron-positron pairs. Any one of these constituents that carries a substantial fraction of the total electron momentum can initiate a hard-scattering process.

Precisely the same logic applies to the calculation of QCD cross sections. The contributions from the region of collinear gluon or quark emission should be associated with the parton distribution functions rather than with the hard-scattering cross sections. If we make this association, we find that the parton distributions are no longer independent of the momentum Q that characterizes the hard-scattering process; rather, they now evolve logarithmically with Q. For example, the basic equation (17.30) for deep-inelastic scattering will become

$$\frac{d^2\sigma}{dxdy}(e^-p \to e^-X) = \Big(\sum_f x f_f(x,Q) Q_f^2\Big) \cdot \frac{2\pi\alpha^2 s}{Q^4}\big[1+(1-y)^2\big], \qquad (17.125)$$

and so Bjorken scaling will be violated. Since this violation takes place only on a logarithmic scale in Q^2, it will be a subtle effect, and *approximate* Bjorken scaling will still be a prediction of QCD. But the violation of Bjorken scaling is inevitable, since QCD is a quantum field theory with degrees of freedom at all momentum scales. As we probe the proton wavefunction at increasingly short distances, we excite the high-momentum degrees of freedom and resolve the wavefunction into an increasing number of quarks, antiquarks, and gluons.

The evolution of the QED parton distributions, governed by Eq. (17.120), is characterized by the parameter α/π, so the parton distributions change by $\sim$1% as Q is changed by a factor of 10. In QCD, the corresponding factor governing the rate of evolution should be $\alpha_s(Q)/\pi$. Thus, when Q is very small, the evolution is rapid and contributions of higher order in perturbation theory are important. Ultimately, the initial conditions for the evolution are determined by the form of the proton wavefunction at large distance scales, which cannot be calculated using Feynman diagrams. On the other hand, when Q is large, well above 1 GeV in practice, the evolution becomes slow and is dominated by the leading order in perturbation theory. In that case,

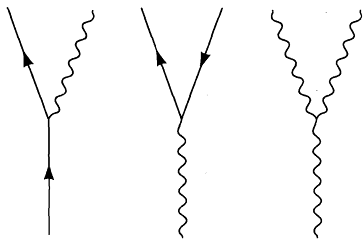


Figure 17.20. The three vertices that contribute to parton evolution in QCD.

QCD perturbation theory makes precise predictions for the form of the evolution of the parton distributions, and these predictions can be tested against experiment.

To derive the evolution equations of parton distributions in QCD we can use the same techniques and logic that we used above for QED. There is a subtlety, that the reduction of the gluon propagator to transverse polarization states in the limit $q^2 \to 0$, Eq. (17.81), cannot be proved so simply as in QED. However, the result is correct also in the non-Abelian case.* Once this technical point is resolved, the kinematics of collinear emission is exactly the same as in QED. Thus we find evolution equations of the same form as in QED, modified only by the replacement of α by α_s, the insertion of appropriate color factors, and the accounting of the effects of the three-gluon vertex.

Collinear emission processes in QCD involve the three vertices shown in Fig. 17.20. Of these, the first two have the same Lorentz structure as those shown in Figs. 17.16 and 17.19. The only difference, aside from the strength of the coupling constant, comes in the color indices. We will treat color just as we treated spin in the preceding analysis: We average over initial colors, and sum over final colors. Then the first vertex of Fig. 17.20, representing the splitting of a quark into a quark and a gluon, receives the color factor

$$\tfrac{1}{3}\,\mathrm{tr}[t^a t^a] = C_2(r) = \tfrac{4}{3}. \tag{17.126}$$

The second vertex, representing the splitting of a gluon into a quark-antiquark pair, receives the factor

$$\tfrac{1}{8}\,\mathrm{tr}[t^a t^a] = \tfrac{1}{2}. \tag{17.127}$$

The third vertex in Fig. 17.20 represents the splitting of a gluon to two gluons, an effect that is new to the non-Abelian case. It is straightforward to compute the contribution of this vertex to the evolution equations by taking the matrix elements of the vertex between transverse gluon states of definite helicity. This calculation is the subject of Problem 17.4.

*See, for example, J. Collins and D. Soper, in A. Mueller, *Quantum Chromodynamics* (World Scientific, Singapore, 1991).

By accounting for all of these effects, we can modify the QED evolution equations (17.120) into the correct set of evolution equations for parton distributions in QCD. These are known as the *Altarelli-Parisi equations.* They describe the coupled evolution of parton distributions $f_f(x,Q)$, $f_{\bar{f}}(x,Q)$ for each flavor of quark and antiquark that can be treated as massless at the scale Q, together with the parton distribution of gluons, $f_g(x,Q)$. Explicitly,

$$
\begin{aligned}
\frac{d}{d\log Q} f_g(x,Q) &= \frac{\alpha_s(Q^2)}{\pi}\int_x^1 \frac{dz}{z}\Big\{P_{g\leftarrow q}(z)\sum_f \big[f_f(\tfrac{x}{z},Q)+f_{\bar{f}}(\tfrac{x}{z},Q)\big] \\
&\qquad\qquad + P_{g\leftarrow g}(z) f_g(\tfrac{x}{z},Q)\Big\}, \\
\frac{d}{d\log Q} f_f(x,Q) &= \frac{\alpha_s(Q^2)}{\pi}\int_x^1 \frac{dz}{z}\Big\{P_{q\leftarrow q}(z) f_f(\tfrac{x}{z},Q) + P_{q\leftarrow g}(z) f_g(\tfrac{x}{z},Q)\Big\}, \\
\frac{d}{d\log Q} f_{\bar{f}}(x,Q) &= \frac{\alpha_s(Q^2)}{\pi}\int_x^1 \frac{dz}{z}\Big\{P_{q\leftarrow q}(z) f_{\bar{f}}(\tfrac{x}{z},Q) + P_{q\leftarrow g}(z) f_g(\tfrac{x}{z},Q)\Big\}.
\end{aligned}
\tag{17.128}
$$

The first three splitting functions can be taken from Eqs. (17.121), multiplied by the color factors computed in Eqs. (17.126) and (17.127):

$$
\begin{aligned}
P_{q\leftarrow q}(z) &= \frac{4}{3}\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right], \\
P_{g\leftarrow q}(z) &= \frac{4}{3}\left[\frac{1+(1-z)^2}{z}\right], \\
P_{q\leftarrow g}(z) &= \frac{1}{2}\left[z^2+(1-z)^2\right].
\end{aligned}
\tag{17.129}
$$

The fourth splitting function requires also the computation of Problem 17.4; the result is

$$
P_{g\leftarrow g}(z) = 6\left[\frac{(1-z)}{z} + \frac{z}{(1-z)_+} + z(1-z) + \Big(\frac{11}{12} - \frac{n_f}{18}\Big)\delta(1-z)\right]. \tag{17.130}
$$

The final term in this expression, which is proportional to n_f, the number of light quark flavors, is the subtraction term associated with gluon splitting into $q\bar{q}$ pairs. The Altarelli-Parisi equations describe the evolution of parton distributions for any hadron, or any hadronic constituent, up to corrections of order α_s that are not enhanced by large logarithms.

Our derivation of the Altarelli-Parisi equations respects the conservation laws of QCD for quark numbers and longitudinal momentum. Thus, the equations must respect the parton-model sum rules (17.36) and (17.39). As in the QED case, it is instructive to verify explicitly that these integrals are independent of Q.

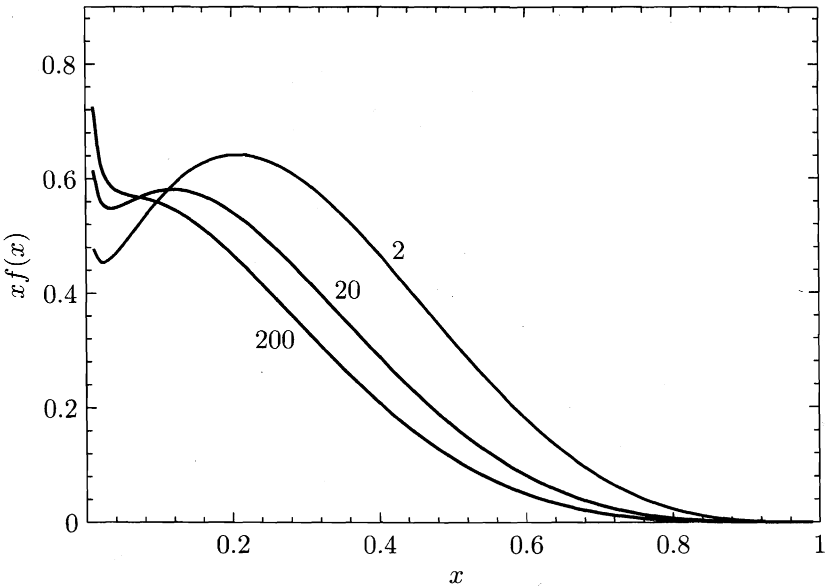


Figure 17.21. The u quark parton distribution function $xf_u(x,Q)$ at $Q =$ 2, 20, and 200 GeV, showing the effects of parton evolution according the Altarelli-Parisi equations. These curves are taken from the CTEQ fit to deep inelastic scattering data described in Fig. 17.6.

In QED, we could use the evolution equations to explicitly compute the structure function of the electron. In QCD, this is no longer possible, because the initial conditions required to integrate the equations are determined by the strong-coupling region of QCD and so are not known *a priori*. However, one can determine the initial conditions of the proton structure experimentally, by measuring the cross section for deep inelastic scattering at a given value of Q^2. One can then predict the structure functions, and thus the deep inelastic cross sections, at higher values of Q^2. There is one subtlety in this analysis: The gluon distribution is not directly measured in deep inelastic scattering, but it does enter the evolution equation for the quark distributions. Thus, some of the information on the Q^2 dependence of deep inelastic scattering simply goes into determining the gluon distribution. However, the gluon distribution is absolutely normalized by the momentum sum rule (17.39), so the evolution equations have predictive power even if this distribution must be fit from the data.

The Altarelli-Parisi equations predict a characteristic form for the evolution of parton distribution functions, shown in Fig. 17.21. Partons at high x tend to radiate and drop down to lower values of x. Meanwhile, new partons are formed at low x as products of this radiation. Thus, the parton distributions decrease at large x and increase much more rapidly at small x as Q^2 increases. We can picture the proton as having more and more constituents,

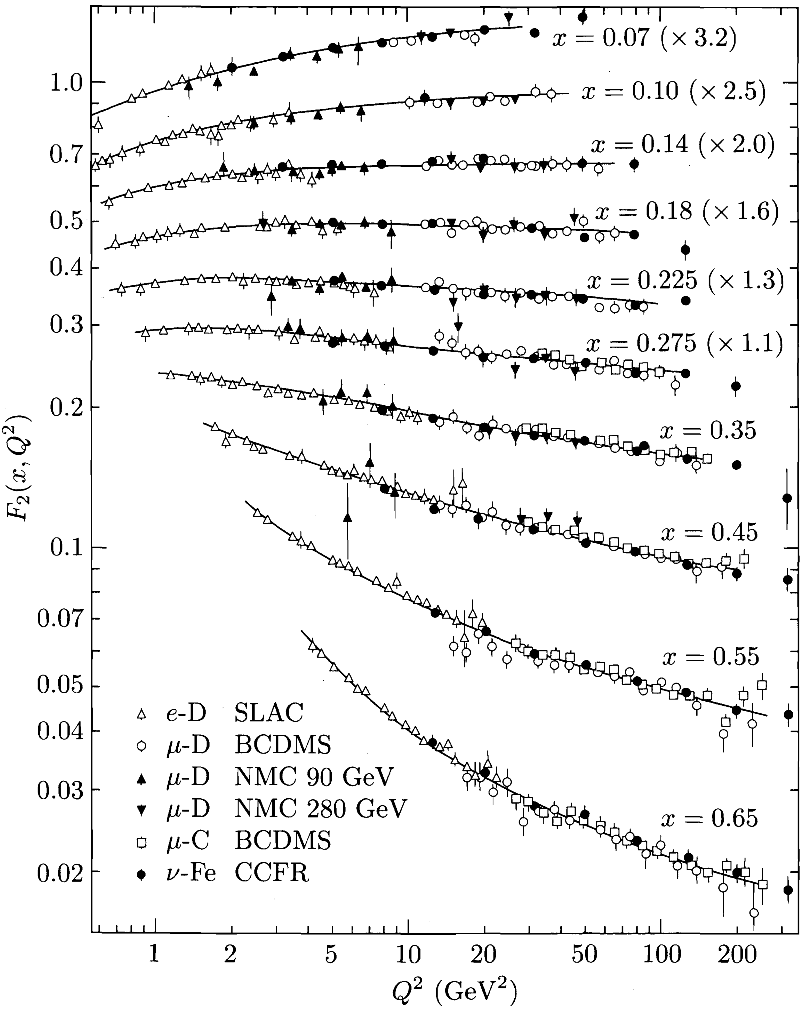


Figure 17.22. Dependence on Q^2 of the combination of quark distribution functions $F_2 = \sum_f xQ_f^2 f_f(x, Q^2)$ measured in deep inelastic electron-proton scattering. The various curves show the variation of F_2 for fixed values of x, and the comparison of this variation to a model evolved with the Altarelli-Parisi equations. The upper six data sets have been multiplied by the indicated factors to separate them on the plot. The data were compiled by M. Virchaux and R. Voss for the Particle Data Group, *Phys. Rev.* **D50**, 1173 (1994), Fig. 32.2. The complete references to the original experiments are given there.

which share its total momentum, as its wavefunction is probed on finer and finer distance scales.

Figure 17.22 shows the evolution of the combination of distribution functions that is measured in deep inelastic scattering, as a function of Q^2. We see the characteristic decrease of the distribution functions at large x and the increase at small x. The data are compared to a model evolved according to the Altarelli-Parisi equations; this model apparently describes the data quite well.

17.6 Measurements of α_s

Before concluding our introductory survey of QCD, we should summarize the quantitative verification of the theory. We discussed precision tests of QED in Section 6.3, bringing together various measurements of the coupling α; the best determinations agree to eight significant figures. Since QCD perturbation theory works only for hard-scattering processes, with uncertainties due to soft processes that are difficult to estimate, this theory has not been tested to such extreme accuracy. Nevertheless, it is interesting to bring together the best available determinations of α_s, to see how well they agree.

In order to compare values of α_s, it is necessary to express these using a common set of conventions. First, one must set the renormalization scale; a useful choice is the mass of the neutral weak boson Z^0: $m_Z = 91.19$ GeV. Second, one must fix the renormalization scheme that defines the QCD coupling constant at this scale. It has become conventional to use as a standard the bare coupling after regularization by modified minimal subtraction, Eq. (11.77). The resulting standard coupling constant is called $\alpha_{s\overline{\text{MS}}}(m_Z^2)$.

Measurements of α_s from a number of types of experiments are summarized in Table 17.1. In Section 17.2 we saw that one can obtain a value of α_s from the measurement of the total cross section for e^+e^- annihilation to hadrons or, equivalently, the ratio R of the number of observed hadronic and leptonic events. An independent measurement of α_s can be obtained from the fraction of e^+e^- annihilation events with three-jet final states or, equivalently, from the transverse momentum distribution of produced hadrons relative to the jet axis. A number of measurements of this type are collected and averaged under the heading 'Event shapes'. A similar measurement of α_s is obtained from the measurement of the transverse momentum spectrum of W bosons produced from quark-antiquark annihilation at high-energy $p\bar{p}$ colliders. The gluon radiative correction to the vertex in deep inelastic neutrino scattering can also be used to extract α_s. The rate of Bjorken scaling violation in deep inelastic scattering is controlled by α_s, and so this effect provides another α_s measurement. The decays of the lightest $b\bar{b}$ bound state Υ and the $c\bar{c}$ bound state ψ are governed by QCD and yield a measurement of α_s. Finally, the spectrum of $c\bar{c}$ and $b\bar{b}$ bound states can be computed numerically in terms of the QCD coupling constant, and the comparison with experiment gives a determination of α_s.

Table 17.1. Values of $\alpha_s(m_Z)$ Obtained from QCD Experiments

Process:	$\alpha_s(m_Z)$	Q (GeV)
Deep inelastic scattering	0.118 (6)	1.7
R in τ lepton decay	0.123 (4)	1.8
ψ, Υ spectroscopy	0.110 (6)	2.3
Transverse momentum of W production	0.121 (24)	4.
Deep inelastic scattering (evolution)	0.112 (4)	5.
Event shapes in e^+e^- annihilation	0.121 (6)	5.8,9.1
Rate for ψ, Υ decay	0.108 (10)	9.5
R in e^+e^- annihilation (20–65 GeV)	0.124 (21)	35.
R in Z^0 decay	0.124 (7)	91.2

The values of $\alpha_s(m_Z)$ displayed in this table are obtained by fitting experimental results to the theoretical expressions given by perturbative QCD using minimal subtraction. The values of α_s have been evolved to $Q = m_Z$ using the renormalization group equation. R refers to the ratio of cross sections or partial widths to hadrons versus leptons. The numbers in parentheses are the standard errors in the last displayed digits. The column labelled 'Q' gives an idea of the value of Q at which the measurement was made. (Typically, these measurements average over a range of Q, and that averaging is taken into account in the quoted values of α_s.) This table is based on the results compiled by I. Hinchliffe in his article for the Particle Data Group, *Phys. Rev.* **D50**, 1297 (1994). This article contains a full set of references and a discussion of the sources of uncertainty in these determinations.

The table shows the values of α_s extracted from each of these measurements, expressed in terms of the value in the reference conventions, $\alpha_{s\overline{\text{MS}}}(m_Z^2)$. We see that several of the experiments determine α_s to an accuracy of 5%, and that the various determinations are consistent with one another at this level. In Fig. 17.23, we have plotted the original values of α_s represented in Table 17.1, before conversion to a common scale, versus the momentum scale Q at which each was obtained. This comparison gives a striking direct verification of the running of α_s.

At the beginning of this chapter, we wrote down a candidate for the fundamental theory of strong interactions using only a few simple principles: the existence of quarks and the identification of their quantum numbers, and the idea that the theory of the quark interactions should be an asymptotically free gauge theory. It is remarkable that these simple considerations have led us to a description of strong interactions that is quantitatively correct for a broad range of phenomena in the hard-scattering regime where asymptotic freedom can be used as a tool for calculation.

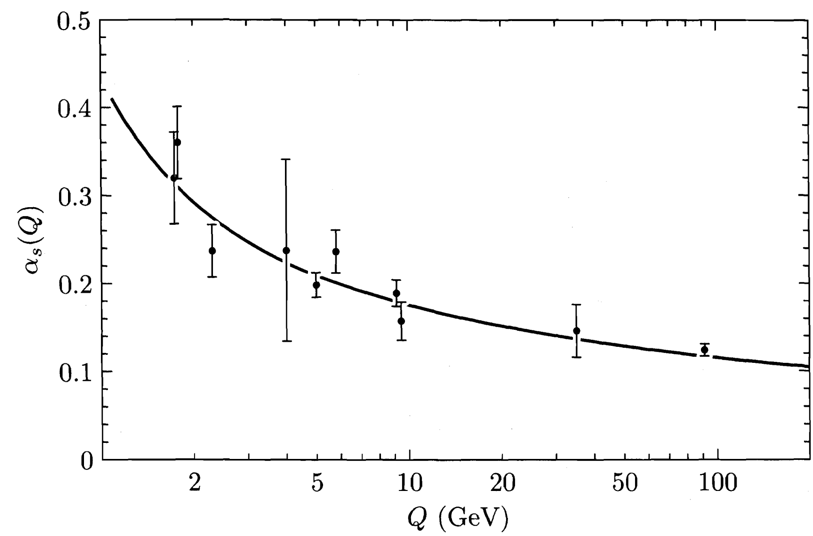


Figure 17.23. Measurements of α_s, plotted against the momentum scale Q at which the measurement was made. This figure was constructed by evolving the values of $\alpha_s(m_Z)$ listed in Table 17.1 back to the values of Q indicated in the table. The value for e^+e^- event shapes has been split into two points corresponding to experiments at the TRISTAN and LEP accelerators. These values are compared to the theoretical expectation from the renormalization group evolution with the initial condition $\alpha_s(m_Z) = 0.117$.

Problems

17.1 Two-loop renormalization group relations.

(a) In higher orders of perturbation theory, the expression for the QCD β function will be a series

$$\beta(g) = -\frac{b_0}{(4\pi)^2}g^3 - \frac{b_1}{(4\pi)^4}g^5 - \frac{b_2}{(4\pi)^6}g^7 + \cdots .$$

Integrate the renormalization group equation and show that the running coupling constant is now given by

$$\alpha_s(Q^2) = \frac{4\pi}{b_0}\left[\frac{1}{\log(Q^2/\Lambda^2)} - \frac{b_1}{b_0^2}\frac{\log\log(Q^2/\Lambda^2)}{(\log(Q^2/\Lambda^2))^2} + \cdots\right],$$

where the omitted terms decrease as $(\log(Q^2/\Lambda^2))^{-2}$.

(b) Combine this formula with the perturbation series for the e^+e^- annihilation cross section:

$$\sigma(e^+e^- \to \text{hadrons}) = \sigma_0 \cdot \Big(3\sum_f Q_f^2\Big) \cdot \Big[1 + \frac{\alpha_s}{\pi} + a_2\Big(\frac{\alpha_s}{\pi}\Big)^2 + \mathcal{O}(\alpha_s^3)\Big].$$

The coefficient a_2 depends on the details of the renormalization conditions defining α_s. Show that the leading two terms in the asymptotic behavior of $\sigma(s)$ for large s depend only on b_0 and b_1 and are independent of a_2 and b_2. Thus the first *two* coefficients of the QCD β function are independent of the renormalization prescription.

17.2 A direct test of the spin of the gluon. In this problem, we compare the predictions of QCD with those of a model in which the interaction of quarks is mediated by a scalar boson. Let the coupling of the scalar gluon to quarks be given by

$$\delta\mathcal{L} = gS\bar{q}q,$$

and define $\alpha_g = g^2/4\pi$.

(a) Using the technique described in parts (b) and (c) of the Final Project of Part I, compute the cross section for $e^+e^- \to q\bar{q}S$ to the leading order of perturbation theory. This cross section depends on the energies of the q, $\bar{q}$, and S, which we represent as fractions x_1, x_2, x_3 of the electron beam energy, as in Eq. (17.18). Show that

$$\frac{d^2\sigma}{dx_1 dx_2}(e^+e^- \to q\bar{q}S) = \frac{4\pi\alpha^2 Q_q^2}{3s} \cdot \frac{\alpha_g}{4\pi} \frac{x_3^2}{(1-x_q)(1-x_{\bar{q}})}.$$

(b) In practice, it is very difficult to tell quarks from gluons experimentally, since both particles appear as jets of hadrons. Therefore, let x_a be the largest of x_1, x_2, x_3, let x_b be the second largest, and let x_c be the smallest. Sum over the various possibilities to derive an expression for $d^2\sigma/dx_a dx_b$, both in QCD, using Eq. (17.18), and in the scalar gluon model. Show that these models can be distinguished by their distributions in the x_a, x_b plane.

17.3 Quark-gluon and gluon-gluon scattering.

(a) Compute the differential cross section

$$\frac{d\sigma}{d\hat{t}}(q\bar{q} \to gg)$$

for quark-antiquark annihilation in QCD to the leading order in α_s. This is most easily done by computing the amplitudes between states of definite quark and gluon helicity. Ignore all masses. Use explicit polarization vectors and spinors, for example,

$$\epsilon^\mu = \frac{1}{\sqrt{2}}(0, 1, i, 0)$$

for a right-handed gluon moving in the $+\hat{3}$ direction. You need only consider transversely polarized gluons. By helicity conservation, only the initial states $q_L\bar{q}_R$ and $q_R\bar{q}_L$ can contribute; by parity, these two states give identical cross sections. Thus it is necessary only to compute the amplitudes for the three processes

$$q_L\bar{q}_R \to g_R g_R,$$
$$q_L\bar{q}_R \to g_R g_L,$$
$$q_L\bar{q}_R \to g_L g_L.$$

In fact, by CP invariance, the first and third processes have equal cross sections. After computing the amplitudes, square them and combine them properly with color factors to construct the various helicity cross sections. Finally, combine these to form the total cross section averaged over initial spins and colors.

(b) Compute the differential cross section

$$\frac{d\sigma}{d\hat{t}}(gg \to gg)$$

for gluon-gluon scattering. There are 16 possible combinations of helicities, but many of them are related to each other by parity and crossing symmetry. All 16 can be built up from the three amplitudes for

$$g_R g_R \to g_R g_R,$$
$$g_R g_R \to g_R g_L,$$
$$g_R g_R \to g_L g_L.$$

Show that the last two of these amplitudes vanish. The first can be dramatically simplified using the Jacobi identity. When the smoke clears, only three of the 16 polarized gluon scattering cross sections are nonzero. Combine these to compute the spin- and color-averaged differential cross section.

17.4 The gluon splitting function. Compute the gluon splitting function (17.130) for the Altarelli-Parisi equations. To carry out this computation, first compute the matrix elements of the three-gluon vertex shown in Fig. 17.20 between gluon states of definite helicity. Combine these to derive the splitting function in the region $x < 1$. Then fix the singularity of the splitting function at $x = 1$ to give this function the correct overall normalization.

17.5 Photoproduction of heavy quarks. Consider the process of heavy quark pair photoproduction, $\gamma + p \to Q\overline{Q} + X$, for a heavy quark of mass M and electric charge Q. If M is large enough, any diagram contributing to this process must involve a large momentum transfer; thus a perturbative QCD analysis should apply. This idea applies in practice already for the production of c quark pairs. Work out the cross section to the leading order in QCD. Choose the parton subprocess that gives the leading contribution to this reaction, and write the parton-model expression for the cross section. You will need to compute the relevant subprocess cross section, but this can be taken directly from one of the QED calculations in Chapter 5. Then use this result to write an expression for the cross section for γ-proton scattering.

17.6 Behavior of parton distribution functions at small x. It is possible to solve the Altarelli-Parisi equations analytically for very small x, using some physically motivated approximations. This discussion is based on a paper of Ralston.†

(a) Show that the Q^2 dependence of the right-hand side of the A-P equations can be expressed by rewriting the equations as differential equations in

$$\xi = \log\log\left(\frac{Q^2}{\Lambda^2}\right),$$

where Λ is the value of Q^2 at which $\alpha_s(Q^2)$, evolved with the leading-order β function, formally goes to infinity.

---

†J. P. Ralston, *Phys. Lett.* **172B**, 430 (1986).

(b) Since the branching functions to gluons are singular as z^{-1} as $z \to 0$, it is reasonable to guess that the gluon distribution function will blow up approximately as x^{-1} as $x \to 0$. The resulting distribution

$$dx\, f_g(x) \sim \frac{dx}{x}$$

is approximately scale invariant, and so its form should be roughly preserved by the A-P equations. Let us, then, make the following two approximations: (1) the terms involving the gluon distribution completely dominate the right-hand sides of the A-P equations; and (2) the function

$$\tilde{g}(x, Q^2) = x f_g(x, Q^2)$$

is a slowly varying function of x. Using these approximations, and the limit $x \to 0$, show that the A-P equation for $f_g(x)$ can be converted to the following differential equation:

$$\frac{\partial^2}{\partial w \partial \xi} \tilde{g}(x, \xi) = \frac{12}{b_0} \tilde{g}(x, \xi),$$

where $w = \log(1/x)$ and $b = (11 - \frac{2}{3} n_f)$. Show that if $w\xi \gg 1$, this equation has the approximate solution

$$\tilde{g} = K(Q^2) \cdot \exp\left(\left[\frac{48}{b_0} w(\xi - \xi_0) \right]^{1/2} \right),$$

where $K(Q^2)$ is an initial condition.

(c) The quark distribution at very small x is mainly created by branching of gluons. Using the approximations of part (b), show that, for any flavor of quark, the right-hand side of the A-P equation for $f_q(x)$ can be approximately integrated to yield an equation for $\tilde{q}(x) = x f_q(x)$:

$$\frac{\partial}{\partial \xi} \tilde{q}(x, \xi) = \frac{2}{3b_0} \tilde{g}(x, \xi).$$

Show, again using $w\xi \gg 1$, that this equation has as its integral

$$\tilde{q} = \left(\frac{\xi - \xi_0}{27 b_0 w} \right)^{1/2} K(Q^2) \cdot \exp\left(\left[\frac{48}{b_0} w(\xi - \xi_0) \right]^{1/2} \right).$$

(d) Ralston suggested that the initial condition

$$K(Q^2) = 50.36(\exp(\xi - \xi_0) - 0.957) \cdot \exp\left[-7.597(\xi - \xi_0)^{1/2} \right],$$

with $Q_0^2 = 5\ \text{GeV}^2$, $\Lambda = 0.2$ GeV, and $n_f = 5$, gave a reasonable fit to the known properties of parton distributions, extrapolated into the small x region. Use this function and the results above to sketch the behavior of the quark and gluon distributions at small x and large Q^2.

Chapter 18

Operator Products and Effective Vertices

Our analysis of QCD in Chapter 17 was founded on the principle of asymptotic freedom, which told us that strong interaction processes with large momentum transfer might reliably be treated in weak-coupling perturbation theory. So far, however, we have made little use in QCD of the more powerful tools of the renormalization group. In this chapter, we will work out some implications of the Callan-Symanzik equation in QCD. We will see that asymptotically free theories have their own characteristic scaling behavior, with corrections in the form of anomalous powers of logarithms of the momentum scale. Though these corrections are generally weaker than those in the scalar field theories studied in Chapter 13, they nevertheless have important qualitative effects on the strong interactions.

We begin by considering the scaling law for mass terms in QCD, taking over directly the formalism that we used to describe the mass term of ϕ^4 theory in Sections 12.4 and 12.5. Other applications, however, require a more powerful theoretical tool, the *operator product expansion.* Section 18.3 introduces a general description of products of operators in quantum field theory and explains how such operator products are constrained by the Callan-Symanzik equation. The last two sections use this tool to develop a new viewpoint toward deep inelastic scattering and other hard processes in QCD.

18.1 Renormalization of the Quark Mass Parameter

Up to this point, we have always assumed that quark masses are small enough that they can be ignored in high-energy processes. This is not always an adequate assumption even for the light quarks u, d, s; for the heavier quarks c, b, t, the masses can have very important effects. However, since isolated quarks do not exist, it is not possible to define the mass of a quark unambiguously. In the discussion to follow, we will consider the quark mass to be a parameter of QCD perturbation theory, defined by a renormalization prescription at some renormalization scale M.

Because we define the quark mass as we would a coupling constant, by a renormalization convention, we should expect that this parameter will run according to a renormalization group evolution, so that different values of the mass parameter apply to different processes. We say that our original prescription leads to an *effective* quark mass, which depends on the momentum scale at which it is evaluated. In this section, we will work out the leading dependence of this effective mass on the momentum scale.

The basic formalism for effective mass terms was set out in Section 12.5. To add a mass term to the QCD Lagrangian, we must first define the mass operator $(\bar{q}q)$ by a renormalization prescription at a scale M. Then we can define the quark mass by adding to the Lagrangian the term

$$\Delta\mathcal{L}_m = -m(\bar{q}q)_M. \tag{18.1}$$

In this discussion, we will assume that the quark mass m is small enough that we need only keep terms of leading order in m. We will also assume, for simplicity, that we have such a mass term for only one quark flavor.

In the zero-mass limit, Green's functions of the operator $(\bar{q}q)$ with quark fields,

$$\begin{aligned} G^{(n,k)}&(x_1,\ldots,x_n,y_1,\ldots,y_n,z_1,\ldots,z_k) \\ &= \langle q(x_1)\cdots q(x_n)\bar{q}(y_1)\cdots\bar{q}(y_n)\bar{q}q(z_1)\cdots\bar{q}q(z_k)\rangle, \end{aligned} \tag{18.2}$$

obey the Callan-Symanzik equation

$$\left[M\frac{\partial}{\partial M}+\beta\frac{\partial}{\partial g}+2n\gamma+k\gamma_{\bar{q}q}\right]G^{(n,k)}(\{x_i\},\{y_i\},\{z_j\},g,M)=0, \tag{18.3}$$

where γ is the anomalous dimension of the quark field and $\gamma_{\bar{q}q}$ is the anomalous dimension of the operator $\bar{q}q$. If we include the mass terms in the Lagrangian according to (18.1), the Green's function of n quark fields and n antiquark fields satisfies

$$\left[M\frac{\partial}{\partial M}+\beta\frac{\partial}{\partial g}+2n\gamma+\gamma_{\bar{q}q}m\frac{\partial}{\partial m}\right]G^{(n)}(\{x_i\},\{y_i\},g,m,M)=0. \tag{18.4}$$

The derivative with respect to m counts the number of times the mass operator is used. In Section 12.5, we traded the variable m, with the dimensions of mass, for a dimensionless variable. However, in QCD, it is just as convenient to consider the dimensionful parameter m as a coupling constant. The solution of the Callan-Symanzik equation will then contain a running mass parameter $\overline{m}(Q)$, which depends on a typical momentum Q of the Green's function. This parameter is defined as the solution to a renormalization group equation analogous to Eq. (12.126). For this case, the equation is

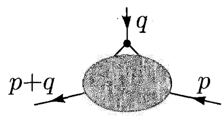
$$\frac{d}{d\log(Q/M)}\overline{m}=\gamma_{\bar{q}q}(\bar{g})\cdot\overline{m}, \tag{18.5}$$

with the initial condition

$$\overline{m}(M)=m. \tag{18.6}$$

The quantity $\overline{m}(Q)$ is the *effective mass*, which should be used to compute the mass effects on quark production or scattering processes with the momentum transfer Q.

To compute $\overline{m}(Q)$ explicitly, we need to work out the anomalous dimension of the mass operator $\gamma_{\bar{q}q}$. This can be done as explained in Section 12.4. We define the normalization of the operator explicitly by the prescription that the vertex function of $(\bar{q}q)$ between renormalized quark fields should satisfy



$$= 1 \tag{18.7}$$

for $p^2 = q^2 = (p+q)^2 = -M^2$. To preserve (18.7), we will need a counterterm vertex $\delta_{\bar{q}q}$ with the structure of the operator insertion. Then, as in Eq. (12.112), the anomalous dimension is given to one-loop order by

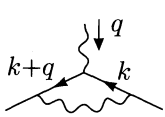
$$\gamma_{\bar{q}q} = M\frac{\partial}{\partial M}\left(-\delta_{\bar{q}q} + \delta_2\right), \tag{18.8}$$

where δ_2 is the counterterm for the quark field strength renormalization, defined in Fig. 16.8. Correlation functions of the gauge-invariant operator $(\bar{q}q)$ are gauge invariant, and so the various terms in the Callan-Symanzik equation for this function must sum to a gauge-invariant result. Since the leading coefficient of $\beta(g)$ is independent of the gauge and of other conventions, it follows from (18.3) that the leading coefficient of $\gamma_{\bar{q}q}$ is also convention independent. The counterterms δ_2 and $\delta_{\bar{q}q}$ both depend on the gauge. This argument shows that the gauge dependence must cancel in (18.8). In the calculation to follow, and in the other anomalous dimension calculations in this chapter, we will work consistently in Feynman-'t Hooft gauge.

We have already computed the divergent part of the counterterm δ_2 in Feynman-'t Hooft gauge in Section 16.4. Evaluating the group-theory factor in the result (16.77) for QCD, we find

$$\delta_2 = -\frac{4}{3}\frac{g^2}{(4\pi)^2}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}. \tag{18.9}$$

To compute $\delta_{\bar{q}q}$, we must work out the one-loop correction to the vertex (18.7). This is given by the diagram



$$= \int\frac{d^4k}{(2\pi)^4}(ig)^2 t^a\gamma^\mu\frac{i(\not{k}+\not{q})}{(k+q)^2}\cdot 1\cdot\frac{i\not{k}}{k^2}t^a\gamma_\mu\frac{-i}{(k-p)^2}. \tag{18.10}$$

In the expression for this diagram, the factor 1 represents the $\bar{q}q$ operator insertion. In the corresponding diagram for the renormalization of the quark number current $j^\nu = \bar{q}\gamma^\nu q$, this factor would be replaced by γ^ν. Since we need only the divergent part of the vertex renormalization (18.10), we can approximate the integrand by its value for large k. Then this diagram becomes

$$\begin{aligned}&\sim \int \frac{d^4k}{(2\pi)^4}\,(ig)^2 t^a \gamma^\mu \frac{i\not{k}}{k^2}\cdot 1\cdot \frac{i\not{k}}{k^2} t^a \gamma_\mu \frac{-i}{k^2}\\ &\sim -i\frac{4}{3}g^2 \int \frac{d^4k}{(2\pi)^4}\frac{d\cdot k^2}{(k^2)^3}\\ &\sim \frac{4}{3}g^2\cdot 4\cdot \frac{1}{(4\pi)^2}\Gamma(2-\tfrac{d}{2}).\end{aligned} \tag{18.11}$$

To preserve the normalization condition (18.7), we must add the counterterm

$$\delta_{\bar{q}q} = -4\cdot \frac{4}{3}\frac{g^2}{(4\pi)^2}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}. \tag{18.12}$$

Assembling (18.8), (18.9), and (18.12), we find

$$\gamma_{\bar{q}q} = -8\frac{g^2}{(4\pi)^2}. \tag{18.13}$$

As we have noted in the previous paragraph, the anomalous dimension γ_j of the quark number current can be found by a very similar calculation. This will give a good check on our formalism, since, as we have argued above Eq. (12.110), a conserved current is unambiguously normalized by its integral, the conserved charge, and so must have zero anomalous dimension. If we substitute γ^ν for 1 in (18.10) and use the same set of approximations to reduce the integral, we find in the numerator the Dirac matrix structure

$$\begin{aligned}\gamma^\mu \not{k}\gamma^\nu \not{k}\gamma_\mu &= \frac{1}{d}k^2 \gamma^\mu\gamma^\lambda\gamma^\nu\gamma_\lambda\gamma_\mu\\ &= \frac{1}{4}(-2)^2 k^2\gamma^\nu.\end{aligned} \tag{18.14}$$

Then, instead of (18.12), we need the counterterm

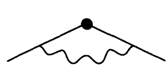
$$\delta_j = -\frac{4}{3}\frac{g^2}{(4\pi)^2}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}. \tag{18.15}$$

Combining this result with (18.9), we find

$$\gamma_j = 0, \tag{18.16}$$

in accord with our general arguments.

If we replace the gamma function in (18.11) by an explicit factor of $\log(\Lambda^2/Q^2)$, and then subtract the divergence using the counterterm (18.12), we find that the vertex diagram behaves as



$$= \frac{4}{3}\cdot 4\frac{g^2}{(4\pi)^2}\log\frac{M^2}{Q^2}. \tag{18.17}$$

This diagram gives an enhancement at small external momenta. Some of this enhancement is associated with the (gauge-dependent) rescaling of the external quark fields; relation (18.8) tells us how to extract the piece of this

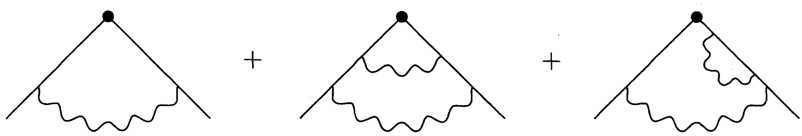


Figure 18.1. Diagrams giving the leading logarithmic contributions to the momentum dependence of the quark effective mass.

logarithm associated with the gauge-invariant enhancement of the effective mass. Thus, to order α_s,

$$\overline{m}(Q) = m \cdot \left(1 + 4\frac{g^2}{(4\pi)^2} \log \frac{M^2}{Q^2}\right). \tag{18.18}$$

To compute the momentum dependence of the effective mass more accurately, we must take two more features of the calculation into account. First, the quantity $(\alpha_s \log(M^2/Q^2))$ may become of order 1, and, in this case, we must take into account all leading logarithmic terms of the form $(\alpha_s \log(M^2/Q^2))^n$. Contributions of this type come from all the diagrams shown in Fig. 18.1. Second, the coupling constant α_s is itself a function of the momentum scale, giving a further enhancement to contributions from small Q. Both of these effects are properly accounted by solving the renormalization group equation (18.5). To the leading order in g^2, this equation takes the explicit form

$$\frac{d}{d\log(Q/M)}\overline{m} = -8\frac{\bar{g}^2}{(4\pi)^2}m = -2\frac{\alpha_s(Q^2)}{\pi}\overline{m}. \tag{18.19}$$

Inserting the solution of the renormalization group equation for $\bar{g}$ in the form (17.17), we find

$$\frac{d}{d\log(Q/M)}\overline{m} = -\frac{8}{b_0\log(Q^2/\Lambda^2)}\overline{m}, \tag{18.20}$$

where b_0 is the first coefficient of the QCD β function and Λ is now the QCD scale parameter defined in (17.16). The integral of this equation, satisfying the initial condition (18.6), is

$$\overline{m}(Q^2) = \left(\frac{\log(M^2/\Lambda^2)}{\log(Q^2/\Lambda^2)}\right)^{4/b_0} m. \tag{18.21}$$

Recall that $b_0 = 11 - \frac{2}{3}n_f$ in QCD. Another way to express (18.21) is by writing

$$\overline{m}(Q^2) = \left(\frac{\alpha_s(Q^2)}{\alpha_s(M^2)}\right)^{4/b_0} m. \tag{18.22}$$

Just as an illustration, take $n_f = 4$ and $\Lambda = 150$ MeV; then the effective masses of the light quarks increase by about a factor 2 from $Q = 100$ GeV to $Q = 1$ GeV.

The method we have just used for computing the QCD enhancement of the quark mass operator applies equally well to the matrix elements of any other gauge-invariant operator. We conclude this section by recapitulating the conclusions of the argument in their more general form.

Let $\mathcal{O}(x)$ be any gauge-invariant operator in QCD. As we saw for the mass term, the one-loop corrections to the matrix elements of this operator may contain enhancement or suppression terms proportional to $\alpha_s \log(M^2/Q^2)$, where Q is the momentum scale of a QCD process mediated by $\mathcal{O}(x)$ and M is the renormalization scale used to define the operator normalization. The part of these one-loop corrections specifically associated with the operator normalization is given by the anomalous dimension $\gamma_{\mathcal{O}}$. For an operator containing n quark or antiquark fields and k gluon fields,

$$\gamma_{\mathcal{O}} = M\frac{\partial}{\partial M}\Big(-\delta_{\mathcal{O}} + \frac{n}{2}\delta_2 + \frac{k}{2}\delta_3\Big), \tag{18.23}$$

where $\delta_{\mathcal{O}}$ is the counterterm needed to preserve the operator normalization condition and δ_2 and δ_3 are the counterterms for the quark and gluon field strength renormalization defined in Fig. 16.8. From (18.23), we can derive the explicit one-loop expression for $\gamma_{\mathcal{O}}$ in the form

$$\gamma_{\mathcal{O}} = -a_{\mathcal{O}}\frac{g^2}{(4\pi)^2}. \tag{18.24}$$

Using this result, we can solve the renormalization group equation for the coefficient of $\mathcal{O}(x)$ and find the QCD renormalization factor

$$\left(\frac{\log(M^2/\Lambda^2)}{\log(Q^2/\Lambda^2)}\right)^{a_{\mathcal{O}}/2b_0}, \tag{18.25}$$

where b_0 is the first coefficient of the QCD β function,

$$b_0 = 11 - \frac{2}{3}n_f. \tag{18.26}$$

The QCD renormalization (18.25) is an enhancement at small momenta if $a_{\mathcal{O}} > 0$.

In the remainder of this chapter, we will present further examples of this enhancement or suppression by QCD logarithms. In many cases, we will see that these factors lead to striking and nontrivial physical effects.

18.2 QCD Renormalization of the Weak Interaction

Our next example of the appearance of QCD enhancement factors occurs in the theory of the weak interactions of hadrons. In Section 17.3, we introduced the weak interaction coupling of quarks and leptons, which we described by an effective Lagrangian. For our analysis here, we will need to know a few more details of the structure of the weak interactions, so we begin this section by presenting these facts. The complete structure of the weak interactions of quarks and leptons will be discussed systematically in Chapter 20.

As we discussed in Section 17.3, the weak interactions among quarks and leptons are described by an effective Lagrangian resulting from the exchange of a virtual W vector boson. In (17.31), we wrote the effective vertex that couples quarks to leptons:

$$\Delta\mathcal{L} = \frac{g^2}{2m_W^2}\bar{\ell}\gamma^\mu\frac{(1-\gamma^5)}{2}\nu\bar{u}\gamma_\mu\frac{(1-\gamma^5)}{2}d + \text{h.c.} \tag{18.27}$$

In this chapter, we will mainly be concerned with the effects of this interaction for momentum scales much larger than 1 GeV. Thus, we will ignore quark masses. All fermion fields that appear in the weak-interaction vertices are multiplied by the left-handed projector $\frac{1}{2}(1-\gamma^5)$. In the rest of this section, we will not write this projector explicitly; rather, we will denote the projection by a subscript L. We will also introduce the Fermi constant, given by (17.32). Then (18.27) can be rewritten as

$$\Delta\mathcal{L} = \frac{4G_F}{\sqrt{2}}\big(\bar{\ell}_L\gamma^\mu\nu_L\big)\big(\bar{u}_L\gamma_\mu d_L\big) + \text{h.c.} \tag{18.28}$$

There is an analogous vertex that represents W exchange between pairs of quarks; this has the form

$$\Delta\mathcal{L} = \frac{4G_F}{\sqrt{2}}\big(\bar{d}_L\gamma^\mu u_L\big)\big(\bar{u}_L\gamma_\mu d_L\big) + \text{h.c.} \tag{18.29}$$

However, for the discussion of this chapter, we will need to write a modified, and less approximate, expression. When we discuss the theory of weak interactions in detail in Chapter 20, we will learn that the charge $+2/3$ quarks (u,c,t) couple to the charge $-1/3$ quarks (d,s,b) through the weak interactions via a unitary rotation. Thus, for example, u couples to the combination

$$\cos\theta_c\, d + \sin\theta_c\, s, \tag{18.30}$$

plus a small admixture of b, which we will ignore in this section. The mixing angle θ_c is called the *Cabibbo angle*. Because of this rotation, the weak interaction effective Lagrangian coupling quarks to quarks actually contains a number of terms, of which a particularly important one is

$$\Delta\mathcal{L} = \frac{4G_F}{\sqrt{2}}\cos\theta_c\sin\theta_c(\bar{d}_L\gamma^\mu u_L)(\bar{u}_L\gamma_\mu s_L). \tag{18.31}$$

This term allows the s quarks to decay through the process $s \to u\bar{u}d$. Similarly, the rotation of (18.28) produces the effective interaction

$$\Delta\mathcal{L} = \frac{4G_F}{\sqrt{2}} \sin\theta_c (\bar{\ell}_L \gamma^\mu \nu_L)(\bar{u}_L \gamma_\mu s_L), \tag{18.32}$$

which leads to the decay $s \to u\bar{\nu}\ell$. These weak interaction processes are referred to as *nonleptonic* and *semileptonic* decay processes, respectively. Similar expressions apply to the other heavy quarks.

Given that (18.31) and (18.32) describe the weak interaction coupling of the s quark at a fundamental level, we now discuss the modification of these couplings by QCD logarithms. We have seen in the previous section that QCD corrections have a profound effect in enhancing the strength of the quark mass term of the underlying Lagrangian. We will now investigate whether the strength of the weak interactions can receive a similar enhancement.

We first consider the semileptonic weak interaction operator (18.32). The leptonic fermion bilinear is not affected by QCD, so the QCD enhancement of this operator is just the same as that of its quark component

$$\bar{u}_L \gamma_\mu s_L. \tag{18.33}$$

However, this operator is a current and so has $\gamma = 0$. In terms of diagrams, the logarithmic enhancement resulting from the diagram shown in Fig. 18.2 is canceled by the quark field-strength renormalization, as we saw already in our discussion of the current vertex is Section 18.1. The left-handed projector $\frac{1}{2}(1-\gamma^5)$ commutes through the diagram and has no effect on the final result. The same remark applies to the semileptonic weak interaction that links u and d quarks. It implies, for that case, that the normalization of the cross sections for deep inelastic neutrino scattering given in (17.35) is not affected by QCD logarithms.

In the case of nonleptonic weak interactions, however, the effect of QCD is not so simple. Let us first compute the Feynman diagrams that give the leading corrections to the renormalization of the weak interaction vertex (18.31) and then, at a later stage, build up the renormalization group interpretation of these results.

At order α_s, the nonleptonic weak interaction vertex receives corrections from the diagrams shown in Fig. 18.3. Notice that the first diagram is precisely the current renormalization found in the semileptonic case. The second diagram gives the analogous renormalization of the second quark current. In the computation of γ, these two contributions cancel the contributions from the field-strength renormalization of the four quark fields. The remaining four diagrams of Fig. 18.3 are new contributions which contribute potentially large rescaling factors.

We now compute these diagrams, beginning with the third diagram in Fig. 18.3. As in the computation of Section 18.1, we are interested in the logarithmically divergent contribution associated with values of the loop momentum k much larger than the external momenta. The simplest way to extract this

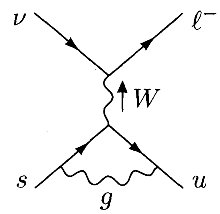


Figure 18.2. QCD correction to the strength of the semileptonic weak interaction vertex.

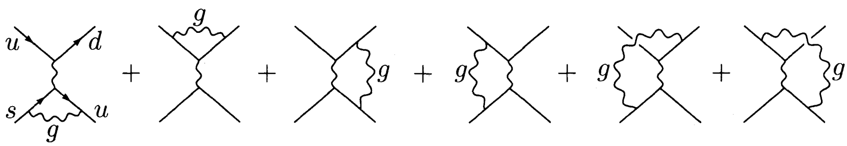


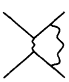
Figure 18.3. QCD corrections to the strength of the nonleptonic weak interaction vertex.

contribution is to compute each diagram in the approximation of zero external momentum. In writing the expression for these diagrams, we will omit the prefactor

$$\frac{4G_F}{\sqrt{2}}\cos\theta_c\sin\theta_c. \tag{18.34}$$

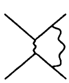
We will retain the quark fields to represent the external states, so that our final expressions will have the form of rescaled operators.

Using this notation, the third diagram in Fig. 18.3 has the value



$$= \int\frac{d^4k}{(2\pi)^4}(ig)^2\frac{-i}{k^2}\left(\bar{d}_L\gamma^\nu t^a\frac{i\not{k}}{k^2}\gamma^\mu u_L\right)\left(\bar{u}_L\gamma_\nu t^a\frac{-i\not{k}}{k^2}\gamma_\mu s_L\right). \tag{18.35}$$

Using the symmetry of the k integral, we extract the divergent piece:



$$\begin{aligned} &= ig^2\int\frac{d^4k}{(2\pi)^4}\frac{k^2/d}{(k^2)^3}\left(\bar{d}_L\gamma^\nu t^a\gamma^\lambda\gamma^\mu u_L\right)\left(\bar{u}_L\gamma_\nu t^a\gamma_\lambda\gamma_\mu s_L\right) \\ &= -\frac{g^2}{4}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}\left(\bar{d}_L\gamma^\nu t^a\gamma^\lambda\gamma^\mu u_L\right)\left(\bar{u}_L\gamma_\nu t^a\gamma_\lambda\gamma_\mu s_L\right). \end{aligned} \tag{18.36}$$

To put the product of quark fields into a more familiar form, we apply the Fierz transformation discussed at the end of Section 3.4. If the color matrices t^a were not present, the product of fermion fields would be exactly the one appearing in (3.82), and we would find

$$\left(\bar{d}_L\gamma^\nu\gamma^\lambda\gamma^\mu u_L\right)\left(\bar{u}_L\gamma_\nu\gamma_\lambda\gamma_\mu s_L\right) = 16\bar{d}_L\gamma^\nu u_L\bar{u}_L\gamma_\nu s_L. \tag{18.37}$$

The matrices t^a redirect the color quantum numbers of the quark fields. To clarify this, we need the analogue of identity (3.77) for color. To find this identity, consider the color invariant

$$(t^a)_{ij}(t^a)_{k\ell}. \tag{18.38}$$

The indices i, k transform according to the 3 representation of color; the indices j, ℓ transform according to the $\bar{3}$. Thus, (18.38) must be a linear combination of the two possible ways to contract these indices,

$$A\delta_{i\ell}\delta_{kj} + B\delta_{ij}\delta_{k\ell}. \tag{18.39}$$

The constants A and B can be determined by contracting (18.38) and (18.39) with δ_{ij} and with δ_{jk} and adjusting A and B so that the contractions of (18.39) obey the identities

$$\text{tr}[t^a](t^a)_{k\ell} = 0; \qquad (t^a t^a)_{i\ell} = \tfrac{4}{3}\delta_{i\ell}. \tag{18.40}$$

This gives the identity

$$(t^a)_{ij}(t^a)_{k\ell} = \tfrac{1}{2}\big(\delta_{i\ell}\delta_{kj} - \tfrac{1}{3}\delta_{ij}\delta_{k\ell}\big). \tag{18.41}$$

A similar relation holds for the generators of $SU(N)$ in the fundamental representation, with $(1/3)$ replaced by $(1/N)$ in that case.

Inserting (18.41) into (18.36), we find that the first term of the identity generates a new four-fermion operator,

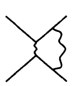
$$\big(\bar{d}_{Li}\gamma^\nu\gamma^\lambda\gamma^\mu u_{Lj}\big)\big(\bar{u}_{Lj}\gamma_\nu\gamma_\lambda\gamma_\mu s_{Li}\big), \tag{18.42}$$

where i, j are color indices. Applying the Fierz rearrangement in (18.37), and then applying the additional rearrangement (3.79), we can convert this operator to the form

$$16(\bar{d}_{Li}\gamma^\nu u_{Lj})(\bar{u}_{Lj}\gamma_\nu s_{Li}) = 16(\bar{u}_{Lj}\gamma^\nu u_{Lj})(\bar{d}_{Li}\gamma_\nu s_{Li}). \tag{18.43}$$

The minus sign in (3.79) is compensated by a minus sign from interchanging the order of fermion fields. The final result is a product of color-singlet quark currents; however, the fields in these currents are associated differently from the original operator.

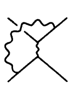
The final result of our evaluation of this diagram is



$$= -4g^2\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}\big[\tfrac{1}{2}\bar{u}_L\gamma^\mu u_L\bar{d}_L\gamma_\mu s_L - \tfrac{1}{6}\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L\big]. \tag{18.44}$$

The fourth diagram of Fig. 18.3 gives precisely the same contribution.

The evaluation of the last two diagrams in Fig. 18.3 is quite similar. The fifth diagram gives

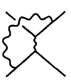


$$\begin{aligned} &= \int\frac{d^4k}{(2\pi)^4}(ig)^2\frac{-i}{k^2}\Big(\bar{d}_L\gamma^\nu t^a\frac{i\not{k}}{k^2}\gamma^\mu u_L\Big)\Big(\bar{u}_L\gamma_\mu t^a\frac{i\not{k}}{k^2}\gamma_\nu s_L\Big) \\ &= -ig^2\int\frac{d^4k}{(2\pi)^4}\frac{k^2/d}{(k^2)^3}\big(\bar{d}_L\gamma^\nu t^a\gamma^\lambda\gamma^\mu u_L\big)\big(\bar{u}_L\gamma_\mu t^a\gamma_\lambda\gamma_\nu s_L\big) \\ &= +\frac{g^2}{4}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}\big(\bar{d}_L\gamma^\nu t^a\gamma^\lambda\gamma^\mu u_L\big)\big(\bar{u}_L\gamma_\mu t^a\gamma_\lambda\gamma_\nu s_L\big). \end{aligned} \tag{18.45}$$

The four-fermion operator can be simplified as follows, by the use of the Fierz identity (3.79):

$$\begin{aligned}(\bar{d}_L\gamma^\nu\gamma^\lambda\gamma^\mu u_L)(\bar{u}_L\gamma_\mu\gamma_\lambda\gamma_\nu s_L) &= +(\bar{d}_L\gamma^\nu\gamma^\lambda\gamma^\mu\gamma_\lambda\gamma_\nu s_L)(\bar{u}_L\gamma_\mu u_L)\\ &= +(-2)^2(\bar{d}_L\gamma^\mu s_L)(\bar{u}_L\gamma_\mu u_L)\\ &= 4(\bar{d}_L\gamma^\mu u_L)(\bar{u}_L\gamma_\mu s_L).\end{aligned} \tag{18.46}$$

Again, we must reduce the product of color matrices using identity (18.41), and, again, the first term of this identity will require an additional Fierz transformation. The final result is



$$= +g^2\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}\left[\tfrac{1}{2}\bar{u}_L\gamma^\mu u_L\bar{d}_L\gamma_\mu s_L - \tfrac{1}{6}\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L\right]. \tag{18.47}$$

The last diagram in Fig. 18.3 gives an identical contribution. The sum of the contributions from these four diagrams is

$$-3g^2\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}\left[\bar{u}_L\gamma^\mu u_L\bar{d}_L\gamma_\mu s_L - \tfrac{1}{3}\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L\right]. \tag{18.48}$$

The extraction of the ultraviolet-divergent pieces of the diagrams of Fig. 18.3 is part of our formal prescription for computing the Callan-Symanzik γ function of the weak interaction vertex. However, it is useful to pause at this point and ask about the physical significance of this divergence. The diagrams of Fig. 18.3 would not be divergent if we computed them in the underlying theory with W bosons. In writing the weak interaction as an effective local vertex, we approximated the W boson propagator by a constant, assuming that the momentum k that it carried was much less than m_W:

$$\frac{1}{k^2 - m_W^2} \to \frac{-1}{m_W^2}. \tag{18.49}$$

The approximation we used to compute the QCD corrections to the effective vertex is valid only in the region of integration where $k^2 \ll m_W^2$. Outside this region we must use the full W propagator; this introduces an extra factor of k^2 in the denominator and makes the integral converge. Thus, in a direct calculation of the QCD correction, the ultraviolet-divergent terms in the evaluation of Fig. 18.3 would be replaced by logarithms cut off at m_W. The lower limit of the logarithm is set by the external momenta. In the decay of a K meson—the lightest hadron containing the s quark—these are of order m_K. Thus the correction given in (18.48) should be evaluated by replacing

$$g^2\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} \to \frac{\alpha_s}{4\pi}\log\frac{m_W^2}{m_K^2}. \tag{18.50}$$

With this interpretation, we can rewrite (18.48) as the order-α_s correction to the leading-order weak interaction vertex. The effect of this correction is the rescaling and modification of the weak interaction operator:

$$\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L \rightarrow$$
$$\left(1+\frac{\alpha_s}{4\pi}\log\frac{m_W^2}{m_K^2}\right)\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L - 3\left(\frac{\alpha_s}{4\pi}\log\frac{m_W^2}{m_K^2}\right)\bar{u}_L\gamma^\mu u_L\bar{d}_L\gamma_\mu s_L. \tag{18.51}$$

Notice that the QCD corrections not only rescale the normalization of the original operator but also introduce a new operator with a different structure. This calculation makes concrete the idea introduced in Section 12.4 that the diagrams that change the normalization of local operators may also mix together different operators with the same dimension and quantum numbers.

Since the value of the logarithm in (18.50) is about 10, the size of the leading QCD correction is of order 1 and so higher-order corrections are important. To sum the leading logarithmic corrections, we return to the renormalization group analysis. For clarity, define

$$\mathcal{O}^1 = \bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L; \qquad \mathcal{O}^2 = \bar{u}_L\gamma^\mu u_L\bar{d}_L\gamma_\mu s_L. \tag{18.52}$$

We will use the subscript 0 to denote bare operators and the subscript M to denote operators obeying renormalization conditions at the scale M. From the diagrams of Fig. 18.3, we have found that the operator whose matrix elements have the quark structure of $\mathcal{O}^1$, properly normalized at the scale M, is given by

$$\mathcal{O}^1_M = \mathcal{O}^1_0 + \delta^{11}\mathcal{O}^1_0 + \delta^{12}\mathcal{O}^2_0, \tag{18.53}$$

where the δ^{ij} are counterterms,

$$\delta^{11} = -\frac{g^2}{(4\pi)^2}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}; \qquad \delta^{12} = +3\frac{g^2}{(4\pi)^2}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}. \tag{18.54}$$

A reciprocal calculation gives $\mathcal{O}^2_M$ in terms of bare operators:

$$\mathcal{O}^2_M = \mathcal{O}^2_0 + \delta^{21}\mathcal{O}^1_0 + \delta^{22}\mathcal{O}^2_0, \tag{18.55}$$

with

$$\delta^{21} = \delta^{12}, \qquad \delta^{22} = \delta^{11}.$$

Then, in the manner than we discussed in Eq. (12.109), the operator rescaling of $\mathcal{O}^1$ and $\mathcal{O}^2$ is described in the Callan-Symanzik equation by a matrix γ^{ij} linking the two operators. Expanding this equation to first order in g^2, we see that this matrix is given to one-loop order by

$$\gamma^{ij} = M\frac{\partial}{\partial M}\left[-\delta^{ij}\right]. \tag{18.56}$$

Thus we find

$$\gamma = \frac{g^2}{(4\pi)^2}\begin{pmatrix} -2 & 6 \\ 6 & -2 \end{pmatrix}, \tag{18.57}$$

acting on the space of operators $\mathcal{O}^1$, $\mathcal{O}^2$.

The simplest way to deduce the physical effects of the rescaling described by (18.57) is to diagonalize this matrix and thus find a new basis of operators that are rescaled without mixing. For the matrix (18.57), the eigenoperators are easily seen to be

$$\begin{aligned}\mathcal{O}^{1/2} &= \tfrac{1}{2}\left[\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L - \bar{u}_L\gamma^\mu u_L\bar{d}_L\gamma_\mu s_L\right],\\ \mathcal{O}^{3/2} &= \tfrac{1}{2}\left[\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L + \bar{u}_L\gamma^\mu u_L\bar{d}_L\gamma_\mu s_L\right].\end{aligned} \tag{18.58}$$

The superscripts on these operators are their isospin quantum numbers. The operator $\mathcal{O}^{1/2}$ is antisymmetric under the interchange of the labels $\bar{d}$ and $\bar{u}$; thus, these two isospin-1/2 fields are combined to total isospin zero, and so the whole operator is isospin-1/2. This operator can mediate decays of the K meson that change the isospin by 1/2 unit, such as $K^0 \to \pi^+\pi^-$, but not processes that change the isospin by 3/2, such as $K^+ \to \pi^+\pi^0$. Experimentally, processes of the former type occur almost a thousand times faster (an observation called the $\Delta I = 1/2$ rule). Thus, it is interesting that the hard QCD corrections already make a distinction between these operators.

From the eigenvalues of (18.57), we obtain the Callan-Symanzik γ functions of the eigenoperators (18.58):

$$\gamma_{1/2} = -8\frac{g^2}{(4\pi)^2}; \qquad \gamma_{3/2} = +4\frac{g^2}{(4\pi)^2}. \tag{18.59}$$

According to Eqs. (18.24) and (18.25), this implies that the operator $\mathcal{O}^{1/2}$ receives an enhancement from hard QCD logarithms, while the operator $\mathcal{O}^{3/2}$ receives a suppression. More explicitly, we can write the operator that appears in the original nonleptonic weak interaction vertex (18.31) as

$$\left[\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L\right]\Big|_{m_W} = \left[\mathcal{O}^{1/2}\right]\Big|_{m_W} + \left[\mathcal{O}^{3/2}\right]\Big|_{m_W}. \tag{18.60}$$

As above, the subscript refers to the mass scale at which the operator is normalized. We now account for the QCD logarithms associated with evaluating the matrix element of this operator at a lower momentum scale, m_K, by replacing the operators on the right-hand side of (18.60) with operators renormalized at m_K, with the rescaling factor (18.25). This gives

$$\begin{aligned}\left[\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L\right]\Big|_{m_W} &= \left(\frac{\log(m_W^2/\Lambda^2)}{\log(m_K^2/\Lambda^2)}\right)^{4/b_0}\left[\mathcal{O}^{1/2}\right]\Big|_{m_K}\\ &\quad + \left(\frac{\log(m_W^2/\Lambda^2)}{\log(m_K^2/\Lambda^2)}\right)^{-2/b_0}\left[\mathcal{O}^{3/2}\right]\Big|_{m_K},\end{aligned} \tag{18.61}$$

where, again, $b_0 = 11 - \frac{2}{3}n_f$. This equation shows that, unlike the case of semileptonic weak interactions, the overall normalization of the effective Lagrangian for nonleptonic weak interactions is changed by QCD logarithms. In addition, the quark structure of the effective Lagrangian is altered.

Quantitatively, taking $n_f = 4$ and $\Lambda = 150$ MeV as an illustration, we find

$$\left[\bar{d}_L\gamma^\mu u_L\bar{u}_L\gamma_\mu s_L\right]\Big|_{m_W} = 2.1\left[\mathcal{O}^{1/2}\right]\Big|_{m_K} + 0.7\left[\mathcal{O}^{3/2}\right]\Big|_{m_K}. \tag{18.62}$$

Thus, the QCD logarithmic corrections from m_W to m_K give the $\Delta I = 1/2$ part of the effective vertex an enhancement of about a factor of 3.* The observed $\Delta I = 1/2$ rule in K decays requires a factor of 20 enhancement. However, part of this is expected to arise from the ratio of the matrix elements of the operators $\mathcal{O}^{1/2}_{m_K}$ and $\mathcal{O}^{3/2}_{m_K}$ between physical hadron states, which are determined by the soft, nonperturbative part of QCD dynamics.

18.3 The Operator Product Expansion

One way to describe the development of the previous section is to say that we studied an interaction that was fundamentally a product of currents by replacing this product of operators with a single local operator. We then derived the physical consequences of the original, composite, interaction by working out the QCD rescaling of this operator. The procedure of replacing a product of operators with a single effective vertex is useful in many contexts in quantum field theory. Thus, in this section, we will pause from our study of QCD to write out the general formalism governing this procedure.

Let us abstract the situation described in the previous section as follows: Consider a quantum field theory process that includes two operators $\mathcal{O}_1$, $\mathcal{O}_2$ separated by a small distance x, together with other fields $\phi(y_i)$ located much farther away, or together with external physical states. In the example above, the two operators are the quark currents that appear in the weak interaction vertex, and their separation x is a distance of order m_W^{-1}, the range of the W propagator. The external states, which contain K and π mesons, can be described by operators that create and destroy these particles. The amplitude for K decay by the weak interactions, or any more general process of this class, can then be extracted from the Green's function

$$G_{12}(x; y_1, \cdots, y_m) = \langle \mathcal{O}_1(x)\mathcal{O}_2(0)\phi(y_1)\cdots\phi(y_m)\rangle, \tag{18.63}$$

considered in the limit $x \to 0$, with the y_i fixed away from the origin. Here and in the following discussion, products of operators will be considered to be time-ordered, just as we would find by writing the product of fields under the functional integral.

The product of operators $\mathcal{O}_1(x)\mathcal{O}_2(0)$ can potentially create the most general local disturbance in the vicinity of the point 0. However, any such disturbance can be described as the effect of a local operator placed at 0. This

*M. K. Gaillard and B. W. Lee, *Phys. Rev. Lett.* **33**, 108 (1974); G. Altarelli and L. Maiani, *Phys. Lett.* **52B**, 351 (1974).

local operator must have the global symmetry quantum numbers of the product of $\mathcal{O}_1\mathcal{O}_2$, but it is otherwise unrestricted. It is useful to write this operator as a linear combination of operators from a standard basis. The coefficients in this linear combination can depend on the separation x. Typically, products of operators in quantum field theory are singular, so it is likely that some of the coefficients will have singularities as $x \to 0$. Combining these observations, Wilson proposed that the effects of the operator product could be computed by replacing the product of operators in (18.63) with a linear combination of local operators,

$$\mathcal{O}_1(x)\mathcal{O}_2(0) \to \sum_n C_{12}{}^n(x)\mathcal{O}_n(0), \tag{18.64}$$

where the coefficients $C_{12}{}^n(x)$ are c-number functions. This *operator product expansion* (OPE) will depend only on the operators $\mathcal{O}_1$, $\mathcal{O}_2$, and their separation and will be independent of the identity and location of the other fields appearing in the Green's function.

The expansion (18.64) implies that the Green's function (18.63) can be expanded for small x as follows:

$$G_{12}(x; y_1, \cdots, y_m) = \sum_n C_{12}{}^n(x)G_n(y_1, \cdots y_m), \tag{18.65}$$

where

$$G_n(y_1, \cdots, y_m) = \langle \mathcal{O}_n(0)\phi(y_1)\cdots\phi(y_m)\rangle\,, \tag{18.66}$$

and all of the dependence on x is now carried by the OPE coefficient functions. In the example of the previous section, the final amplitudes depended in a rather involved way on the small separation of the two operators, through the dependence of the coefficients in (18.61) on m_W. From the viewpoint of the operator product expansion, this dependence is carried by the coefficient functions and is determined for all matrix elements when these are computed.

In Sections 18.1 and 18.2, we used the renormalization group to compute the enhancement or suppression factors for operator matrix elements. Thus it is natural to expect that the form of the operator product coefficients is also determined by the renormalization group. We will now work out this relation. To begin, we rewrite the expansion (18.64) more precisely. The operators that appear in this relation must be defined at some renormalization scale M. Then the operator product expansion reads:

$$\big[\mathcal{O}_1(x)\big]_M\big[\mathcal{O}_2(0)\big]_M = \sum_n C_{12}{}^n(x; M)\big[\mathcal{O}_n(0)\big]_M. \tag{18.67}$$

Note that the coefficient functions can depend on M, since they must absorb the M-dependent operator rescalings. If we use the left-hand side of (18.67) to compute (18.63), this function obeys the Callan-Symanzik equation

$$\left[M\frac{\partial}{\partial M} + \beta\frac{\partial}{\partial g} + m\gamma + \gamma_1 + \gamma_2\right]G_{12}(x; y_1, \cdots, y_m; M) = 0. \tag{18.68}$$

Similarly, with the operator $\mathcal{O}_n$ normalized at M, the Green's function (18.66) obeys

$$\left[M\frac{\partial}{\partial M}+\beta\frac{\partial}{\partial g}+m\gamma+\gamma_n\right]G_n(y_1,\cdots,y_m;M)=0. \tag{18.69}$$

By applying (18.68) to the right-hand side of (18.65), we see that these relations are consistent only if the OPE coefficient functions obey the Callan-Symanzik equation

$$\left[M\frac{\partial}{\partial M}+\beta\frac{\partial}{\partial g}+\gamma_1+\gamma_2-\gamma_n\right]C_{12}{}^n(x;M)=0. \tag{18.70}$$

We now solve this equation by our standard methods. First, let us apply dimensional analysis. If the operators $\mathcal{O}_1$, $\mathcal{O}_2$, $\mathcal{O}_n$ have dimensions d_1, d_2, d_n, the coefficient function $C_{12}{}^n(x)$ must have the dimensions of $(\text{mass})^{d_1+d_2-d_n}$. Thus,

$$C_{12}{}^n(x)=\left(\frac{1}{|x|}\right)^{d_1+d_2-d_n}\widetilde{C}(xM), \tag{18.71}$$

where $\widetilde{C}(xM)$ is a dimensionless function. This function is determined from (18.70) according to the method of Section 12.3. Thus,

$$C_{12}{}^n(x)=\left(\frac{1}{|x|}\right)^{d_1+d_2-d_n}c(\bar{g}(1/x))\exp\left[\int_{1/x}^{M}d\log M'\big(\gamma_n-\gamma_1-\gamma_2\big)\right], \tag{18.72}$$

with $c(\bar{g})$ a dimensionless function of the running coupling constant at the separation scale $1/x$.

At a fixed point of the renormalization group, the γ functions would take definite values $\gamma_{j*}=\gamma_j(g_*)$. Then, the solution (18.72) can be evaluated as

$$C_{12}{}^n(x)=\left(\frac{1}{|x|}\right)^{d_1+d_2-d_n}c(g_*)\exp\Big[\log(xM)\big(\gamma_{n*}-\gamma_{1*}-\gamma_{2*}\big)\Big]. \tag{18.73}$$

Thus, in this case,

$$C_{12}{}^n(x)\sim\left(\frac{1}{|x|}\right)^{d_1^*+d_2^*-d_n^*}, \tag{18.74}$$

where

$$d_j^*=d_j+\gamma_j(g_*) \tag{18.75}$$

is the true scaling dimension of the operator $\mathcal{O}_j$ at the fixed point.

For the case of an asymptotically free theory, the scaling relation is complicated in the way that we worked out in Section 18.1. In the leading order of perturbation theory, the three γ functions take the form (18.24). Then the solution of (18.72) takes the form

$$C_{12}{}^n(x)\sim\left(\frac{1}{|x|}\right)^{d_1+d_2-d_n}\left(\frac{\log(1/|x|^2\Lambda^2)}{\log(M^2/\Lambda^2)}\right)^{(a_n-a_1-a_2)/2b_0} \tag{18.76}$$

In the example of Section 18.2, the original operators were currents with dimension 3 and $\gamma = 0$, at separation m_W^{-1}, and the final local operators had dimension 6. Thus, (18.76) does properly reproduce the dependence of (18.61) on m_W. Notice that the renormalization group dependence is less complicated for a product of currents, which have a fixed normalization independent of scale. This special case occurs often in applications of the operator product expansion.

We have written Eq. (18.70) without taking account of operator mixing. However, as we have already seen, operator mixing is often an essential part of the applications of the OPE. It is straightforward to include this effect by rewriting the analysis that leads to (18.70) using matrix-valued γ functions. For example, with operator mixing, the Callan-Symanzik equation for G_n will be modified to

$$\left[\delta_{np}\left(M\frac{\partial}{\partial M} + \beta\frac{\partial}{\partial g} + m\gamma\right) + \gamma_{np}\right] G_p(y_1, \cdots, y_m; M) = 0. \tag{18.77}$$

With these changes, (18.70) becomes

$$\begin{aligned}\left[M\frac{\partial}{\partial M} + \beta\frac{\partial}{\partial g}\right] &C_{12}{}^n(x;M) + \gamma_{1k} C_{k2}{}^n(x;M) \\ &+ \gamma_{2k} C_{1k}{}^n(x;M) - \gamma_{kn} C_{12}{}^k(x;M) = 0.\end{aligned} \tag{18.78}$$

Notice that the first two γ matrices act on the OPE coefficient from the left, while the third acts from the right. In the case of a product of currents, the first two γ matrices vanish and (18.78) simplifies to

$$\left[M\frac{\partial}{\partial M} + \beta\frac{\partial}{\partial g}\right] C_{12}{}^n(x;M) - C_{12}{}^k(x;M)\gamma_{kn} = 0. \tag{18.79}$$

This equation will play an important role in the analysis of Section 18.5.

18.4 Operator Analysis of e^+e^- Annihilation

It is not difficult to imagine that there is a connection between matrix elements in which currents are placed at short distances from one another and matrix elements in which currents deliver a hard momentum transfer. Thus we might expect that the idea of the operator product expansion will give us a new viewpoint from which to understand the theory of hard-scattering processes in QCD. In this section and the next, we will work out the relation of the operator product expansion to the perturbative QCD analysis of Chapter 17.

We begin by discussing the total cross section for e^+e^- annihilation to hadrons. Below Eq. (17.9), we argued that this total cross section could be computed in QCD perturbation theory, using a value of α_s corresponding to the scale of the total center of mass energy. However, this argument was a purely intuitive one, with many logical jumps. In this section, we will give a more rigorous argument to the same conclusion.

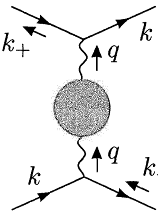


Figure 18.4. Diagrams whose imaginary part yields the total cross section for $e^+e^- \to$ hadrons.

In order to invoke the operator product expansion, we must write the total cross section for e^+e^- annihilation to hadrons as the matrix element of a product of currents. To do this, we use the optical theorem to relate the total e^+e^- scattering cross section to the forward scattering amplitude for $e^+e^- \to e^+e^-$. Ignoring the mass of the electron, we see from Eq. (7.49) that

$$\sigma(e^+e^-) = \frac{1}{2s}\,\mathrm{Im}\,\mathcal{M}(e^+e^- \to e^+e^-). \tag{18.80}$$

To compute the cross section for $e^+e^- \to$ hadrons, we consider in the computation of the imaginary part only the contributions from hadronic intermediate states. To leading order in α, but to all orders in the strong interactions, these contributions come from considering only diagrams of the form of Fig. 18.4, and taking the imaginary part of the hadronic contributions to the vacuum polarization.

The value of the diagrams shown in Fig. 18.4 is

$$i\mathcal{M} = (-ie)^2 \bar{u}(k)\gamma_\mu v(k_+)\frac{-i}{s}\big(i\Pi_h^{\mu\nu}(q)\big)\frac{-i}{s}\bar{v}(k_+)\gamma_\nu u(k), \tag{18.81}$$

where $s = q^2$ and $\Pi_h^{\mu\nu}(q)$ is the hadronic part of the vacuum polarization. By the Ward identity, this can be written

$$\Pi_h^{\mu\nu}(q) = (q^2 g^{\mu\nu} - q^\mu q^\nu)\Pi_h(q^2). \tag{18.82}$$

The $q^\mu q^\nu$ terms give zero when contracted with the external electron currents, so only the $g^{\mu\nu}$ term survives. To evaluate the electron spinor part of (18.81), we use the fact that, in this forward scattering amplitude, the initial and final momenta and spins are set equal. Then, averaging over the initial spin gives

$$\begin{aligned} \frac{1}{2}\cdot\frac{1}{2}\sum_{\text{spins}} \bar{u}(k)\gamma^\mu v(k_+)\bar{v}(k_+)\gamma_\mu u(k) &= \frac{1}{4}\,\mathrm{tr}\big[\not{k}\gamma^\mu \not{k}_+\gamma_\mu\big] \\ &= \frac{1}{4}\cdot(-2)\cdot 4(k\cdot \mathbf{k}_+) \\ &= -s. \end{aligned} \tag{18.83}$$

Thus, we find

$$\sigma(e^+e^- \to \text{hadrons}) = -\frac{4\pi\alpha}{s}\,\mathrm{Im}\,\Pi_h(s). \tag{18.84}$$

To check this result, we can look back to the one-loop value of Π in QED (7.91), or to the imaginary part of this expression given in Eq. (7.92):

$$\operatorname{Im}\Pi(s+i\epsilon) = -\frac{\alpha}{3}\sqrt{1-\frac{4m^2}{s}}\Big(1+\frac{2m^2}{s}\Big). \tag{18.85}$$

Combining (18.85) with (18.84), we obtain the correct leading-order cross section for production of a new heavy lepton in e^+e^- annihilation,

$$\sigma(e^+e^- \to L^+L^-) = \frac{4\pi\alpha^2}{3s}\sqrt{1-\frac{4m^2}{s}}\Big(1+\frac{2m^2}{s}\Big). \tag{18.86}$$

If we multiply (18.86) by a factor of 3 for color and sum over quark flavors with the squares of the quark charges, we obtain the leading-order prediction of QCD.

Now that we have relation (18.84), we complete the connection we wished to prove by noting that the hadronic vacuum polarization is simply a matrix element of a product of currents. Let J^μ be the electromagnetic current of quarks,

$$J^\mu = \sum_f Q_f \bar{q}_f \gamma^\mu q_f. \tag{18.87}$$

Then

$$i\Pi_h^{\mu\nu}(q) = -e^2\int d^4x\, e^{iq\cdot x}\,\langle 0|\, T\{J^\mu(x)J^\nu(0)\}\,|0\rangle\,. \tag{18.88}$$

In the limit in which the point x approaches 0, we can reduce the product of currents by applying the operator product expansion. Since we will be taking the vacuum expectation value of the product, we need only list the contribution from operators that are gauge-invariant Lorentz scalars. Thus,

$$J_\mu(x)J_\nu(0) \sim C_{\mu\nu}{}^{1}(x)\cdot 1 + C_{\mu\nu}{}^{\bar{q}q}(x)\bar{q}q(0) + C_{\mu\nu}{}^{F^2}(x)(F^a_{\alpha\beta})^2(0) + \cdots. \tag{18.89}$$

Note that we have included the operator 1 on the right-hand side, and the next possible operators in QCD have dimension 3 and 4, respectively. Since the operator $\bar{q}q$ violates chiral symmetry, its coefficient function must have an explicit factor of the quark mass. Thus, by dimensional analysis,

$$C_{\mu\nu}{}^{1} \sim x^{-6}, \qquad C_{\mu\nu}{}^{\bar{q}q} \sim mx^{-2}, \qquad C_{\mu\nu}{}^{F^2} \sim x^{-2}, \tag{18.90}$$

and the higher terms in the series are less singular as $x \to 0$.

To compute $\Pi_h^{\mu\nu}(q)$, we need the Fourier transform of the product of currents. Assuming that this Fourier transform is indeed dominated by the limit of short distances, we can compute it by Fourier-transforming the individual OPE coefficients. Since the currents are conserved, the individual terms in the OPE must give zero when dotted with q^μ. Thus the transformed OPE takes the form

$$\begin{aligned} -e^2 \int d^4x \, e^{iq\cdot x} J^\mu(x) J^\nu(0) \\ = -ie^2(q^2 g^{\mu\nu} - q^\mu q^\nu)\big[c^1(q^2)\cdot 1 + c^{\bar{q}q}(q^2)\cdot m\bar{q}q + c^{F^2}(q^2)\cdot (F^a_{\alpha\beta})^2 + \cdots\big], \end{aligned} \tag{18.91}$$

where the c^i are Lorentz-invariant c-number functions of q^2, and the factor of i at the beginning of the second line is inserted as a convenient convention. By dimensional analysis, we find

$$c^1 \sim (q^2)^0, \qquad c^{\bar{q}q} \sim (q^2)^{-2}, \qquad c^{F^2} \sim (q^2)^{-2}, \tag{18.92}$$

and the higher terms are more irrelevant for large q.

The OPE coefficients $c^i(q^2)$ can be computed from Feynman diagrams. As shown in Fig. 18.5, the coefficient of the operator 1 is the sum of diagrams with no external legs other than the current insertions. The leading QCD diagram is just the simple vacuum polarization diagram, multiplied by the color factor 3 and the sum of the squares of the quark charges. Combining these factors with Eq. (7.91), we have

$$c^1(q^2) = -\Big(3\sum_f Q_f^2\Big)\cdot \frac{\alpha}{3\pi}\log(-q^2). \tag{18.93}$$

The corrections to this result are of order $\alpha_s(q^2)$. The higher coefficient functions are extracted from diagrams with more external legs. For example, the coefficient function of $(F^a_{\alpha\beta})^2$ is determined by diagrams with two external gluon legs.

Still assuming that the Fourier transform of the product of currents can be computed from the OPE for the region of large timelike q^2, we can complete our evaluation of the cross section for $e^+e^- \to$ hadrons by taking the vacuum expectation value of (18.91), extracting the imaginary parts of the coefficient functions, and substituting the result into (18.84). We find

$$\begin{aligned} \sigma(e^+e^- \to \text{hadrons}) = \frac{4\pi\alpha^2}{s}\big[\operatorname{Im} c^1(q^2) + \operatorname{Im} c^{\bar{q}q}(q^2)\,\langle 0|\, m\bar{q}q\,|0\rangle \\ + \operatorname{Im} c^{F^2}(q^2)\,\langle 0|\,(F^a_{\alpha\beta})^2\,|0\rangle + \cdots\big]. \end{aligned} \tag{18.94}$$

The first term of this series is just the result of summing perturbative QCD diagrams for the e^+e^- total cross section. The additional terms give corrections to this result which depend on soft hadronic matrix elements, but these corrections are explicitly suppressed at high energy by factors $(q^2)^{-2}$. (Incidentally, this expansion, which applies equally well in the absence of QCD interactions, explains why (18.86) contains no term of order s^{-2} when expanded for large s.) If we insert the leading-order expression (18.93) into (18.94), we obtain the familiar result

$$\sigma(e^+e^- \to \text{hadrons}) = \frac{4\pi\alpha^2}{s}\sum_f Q_f^2. \tag{18.95}$$

(a)

(b)

(c)

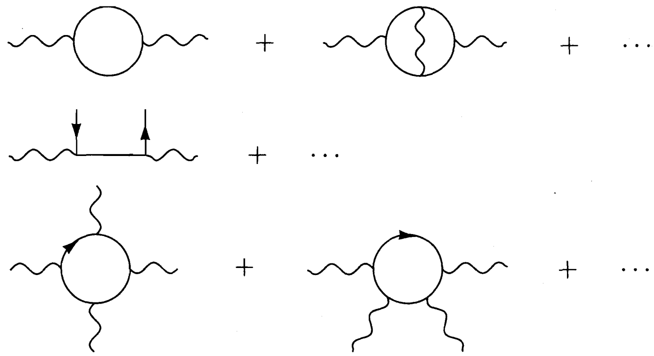


Figure 18.5. Feynman diagrams contributing the operator product coefficient, in the expansion of the product of currents, for the operator (a) 1; (b) $\bar{q}q$; (c) $(F^a_{\alpha\beta})^2$.

Our result (18.94) is pleasing, but the logic that led us to it was not correct. To compute the e^+e^- total cross section, we must compute $\Pi_h(q^2)$ in the region of large *timelike* momentum q, where the expectation value of the product of currents is dominated by intermediate states of high energy, involving large numbers of physical hadrons. Thus we need $\Pi_h(q^2)$ in precisely the region where it is not dominated by short-distance perturbations of the quark and gluon fields. To compute the product of currents from the short-distance expansion, we choose kinematic conditions such that the intermediate states that enter the computation of the product of currents are far off-shell, so that they cannot propagate far from the converging points x and 0. This condition is satisfied at large *spacelike* momentum, or, equivalently, at small spacelike separation. However, it seems at first sight that a computation in this region is useless for determination of the e^+e^- cross section.

Fortunately, there is a wonderful trick for relating the values of a quantum field theory amplitude in two well-separated kinematic regions. This trick, called the method of *dispersion relations*, makes use of the general analytic properties of the amplitude. Since (18.88) is the Fourier transform of a two-point correlation function, we know from the analysis of Section 7.1 that $\Pi_h(q^2)$ possesses a Källén-Lehmann spectral representation. Thus, $\Pi_h(q^2)$ is an analytic function of q^2 with a branch cut on the positive q^2 axis and no other singularities in the complex q^2 plane. This analytic structure is shown in Fig. 18.6. The discontinuity of $\Pi_h(q^2)$ across the branch cut is $(2i)$ times the imaginary part of Π_h and so is directly related to the total e^+e^- annihilation cross section.

With this additional knowledge about $\Pi_h(q^2)$, we can argue as follows. Let $q^2 = -Q_0^2$ be a value sufficiently far into the spacelike region of q that the Fourier transform of the product of currents can be computed from the

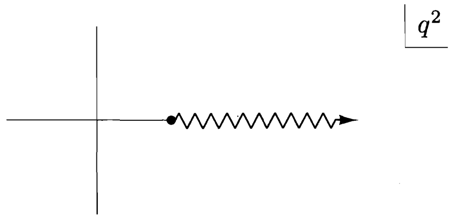


Figure 18.6. Analytic singularities of $\Pi_h(q^2)$ in the complex q^2 plane.

operator product expansion. Now consider the integral

$$I_n = -4\pi\alpha \oint \frac{dq^2}{2\pi i} \frac{1}{(q^2+Q_0^2)^{n+1}} \Pi_h(q^2), \tag{18.96}$$

for $n \geq 1$, evaluated on a contour encircling $q^2 = -Q_0^2$. If we contract the contour onto the pole, we find

$$I_n = -4\pi\alpha \frac{1}{n!} \frac{d^n}{d(q^2)^n} \Pi_h \bigg|_{q^2=-Q_0^2}, \tag{18.97}$$

which can be computed by evaluating Π_h from the the operator product relation (18.91),

$$\Pi_h(q^2) = -e^2\big[c^1(q^2) + c^{\bar{q}q}(q^2)\langle 0|\, m\bar{q}q\,|0\rangle + c^{F^2}(q^2)\langle 0|\,(F^a_{\alpha\beta})^2\,|0\rangle + \cdots\big]. \tag{18.98}$$

On the other hand, we can evaluate the integral by distorting the contour to the form of Fig. 18.7. Since none of the coefficient functions grow faster than $(q^2)^0$ times logarithms as $q^2 \to \infty$, the contour at infinity can be neglected for $n \geq 1$. The piece of the contour that wraps around the branch cut gives

$$\begin{aligned} I_n &= -4\pi\alpha \int \frac{dq^2}{2\pi i} \frac{1}{(q^2+Q_0^2)^{n+1}} \operatorname{Disc}\Pi_h(q^2) \\ &= -4\pi\alpha \int \frac{dq^2}{2\pi} \frac{1}{(q^2+Q_0^2)^{n+1}} \frac{1}{i} 2i \operatorname{Im}\Pi_h(q^2) \\ &= \frac{1}{\pi}\int_0^\infty ds\, \frac{s}{(s+Q_0^2)^{n+1}}\sigma(s). \end{aligned} \tag{18.99}$$

This is an integral over the total cross section for $e^+e^- \to$ hadrons. By equating (18.97) and (18.99), we obtain a series of integral relations between the OPE coefficients, evaluated in QCD perturbation theory, and the observable cross section. These relations, which were first constructed by Novikov, Shifman, Voloshin, Vainshtein, and Zakharov, are known as the *ITEP sum rules.*†

†The theory of these sum rules is reviewed in V. A. Novikov, L. B. Okun, M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, *Phys. Repts.* **41**, 1 (1978).

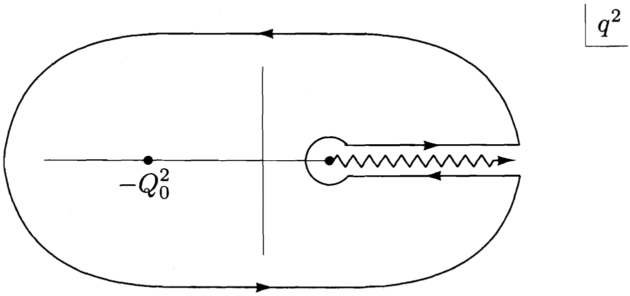


Figure 18.7. Contour of integration involved in the derivation of the ITEP sum rules for $\sigma(e^+e^- \to \text{hadrons})$.

Evaluating the sum rules with only the leading QCD expression for $c^1(q^2)$, we find

$$\int_0^\infty ds \frac{s}{(s+Q_0^2)^{n+1}}\sigma(s) = \frac{4\pi\alpha^2}{n(Q_0^2)^n}\sum_f Q_f^2 + \mathcal{O}(\alpha_s(Q_0^2)) + \mathcal{O}((Q_0^2)^{-2}). \quad (18.100)$$

The leading-order relation is consistent with the lowest-order cross section given in Eq. (18.95). The corrections come from higher orders of QCD perturbation theory, with α_s taken at the scale Q_0^2, and from the higher operator terms in the OPE.

If the correction terms in (18.100) converged to zero uniformly in n, we could invert the sum rules and derive from them our result (18.94). However, the true situation is more subtle. Because the derivatives in (18.97) emphasize terms with stronger q^2 variation, the correction terms in the ITEP sum rules are more and more important as n increases. Thus the most important deviations of the cross section from the prediction of QCD perturbation theory are oscillations about this prediction, which average out in the sum rules for low n. The comparison of theory and experiment is shown in Fig. 18.8. At large s, (18.94) is quite accurate. As s becomes smaller, however, the oscillations grow in size. Eventually, they come to dominate the total cross section as the resonances associated with quark-antiquark bound states.

18.5 Operator Analysis of Deep Inelastic Scattering

We now apply the operator product expansion to another example of a QCD hard-scattering process, deep inelastic electron scattering. In Chapter 17 we found that the predictions of QCD for deep inelastic scattering are precise but also intricate in structure. At a first level, QCD implies that deep inelastic scattering is described by the parton model, in which the incident electron scatters from quarks and antiquarks that carry fractions of the total momentum of the proton. These fractions are determined by parton distribution

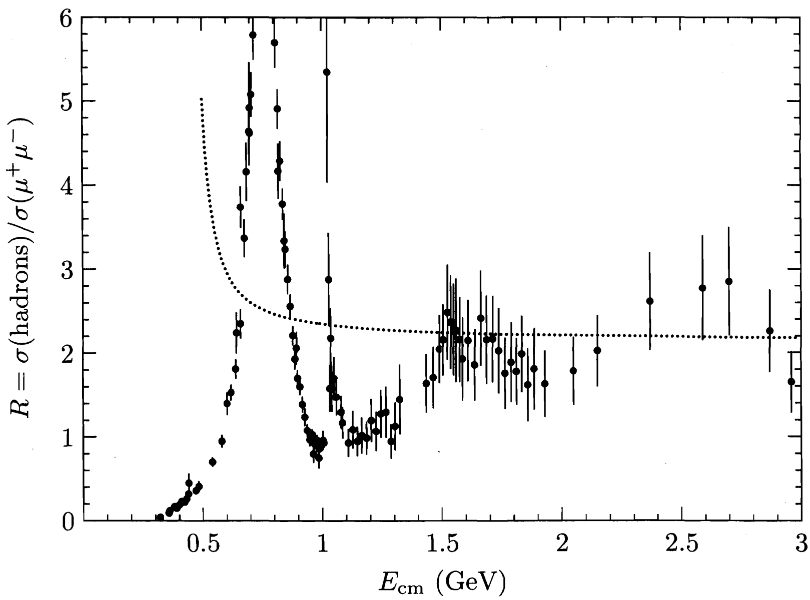


Figure 18.8. Experimental measurements of the total cross section for the reaction $e^+e^- \rightarrow$ hadrons at energies below 3 GeV, compared to the prediction of perturbative QCD for 3 quark flavors. The data are taken from the compilation of M. Swartz, *Phys. Rev.* **D**53, 5268 (1996). Complete references to the various results are given there.

functions, which reflect the form of the proton wavefunction and are determined by soft QCD dynamics. However, we saw in Section 17.5 that effects of QCD perturbation theory cause the parton distributions to change their form as a function of the momentum transfer Q^2. We will now show that much of this picture can be reconstructed from our new viewpoint, using the operator product expansion.

In the previous section, we derived the OPE relations for the e^+e^- annihilation cross section in three steps. First, we used the optical theorem to relate this cross section to a matrix element of a product of currents. Second, we applied the operator product expansion to the product of currents. Unfortunately, this expansion could be used only in an unphysical kinematic region. However, in the third step, we used the method of dispersion relations to connect this unphysical result to an integral over the cross section we wished to predict. In our discussion of deep inelastic scattering, we will go through these same three steps. To obtain our final result, we will need to add a fourth step, involving QCD operator rescaling.

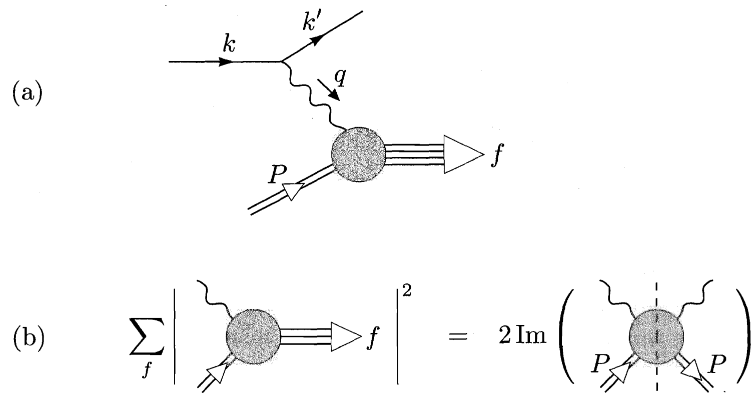


Figure 18.9. Computation of the cross section for deep inelastic electron scattering: (a) general structure of the amplitudes; (b) application of the optical theorem.

Kinematics of Deep Inelastic Scattering

We begin by writing a general expression for the deep inelastic scattering cross section. The matrix element for deep inelastic electron scattering to a final state f is computed as shown in Fig. 18.9(a):

$$i\mathcal{M}(ep \to ef) = (-ie)\bar{u}(k')\gamma_\mu u(k)\frac{-i}{q^2}(ie)\int d^4x\, e^{iq\cdot x}\langle f|\, J^\mu(x)\, |P\rangle\,, \quad (18.101)$$

where $J^\mu(x)$ is the quark electromagnetic current (18.87). The core of this expression is the hadronic matrix element of the current between the proton and some high-energy hadronic state. This matrix element must be squared and summed over possible final states. That sum can be computed, using the optical theorem, by relating it to the forward matrix element of two currents in the proton state, as shown in Fig. 18.9(b). Define

$$W^{\mu\nu} = i\int d^4x\, e^{iq\cdot x}\langle P|\, T\{J^\mu(x)J^\nu(0)\}\, |P\rangle\,, \quad (18.102)$$

averaged over the spin of the proton. This object is known as the *forward Compton amplitude*, since if it is evaluated at $q^2 = 0$ and contracted with physical polarization vectors, it gives the forward amplitude for photon-proton scattering:

$$i\mathcal{M}(\gamma p \to \gamma p) = (ie)^2\epsilon^*_\mu(q)\epsilon_\nu(q)\big(-iW^{\mu\nu}(P,q)\big). \quad (18.103)$$

However, in the following discussion we will need to analyze (18.102) for general spacelike q and for general polarization states.

The optical theorem for Compton scattering from a proton is

$$2\,\mathrm{Im}\,\mathcal{M}(\gamma p \to \gamma p) = \sum_f \int d\Pi_f \big|\mathcal{M}(\gamma p \to f)\big|^2. \quad (18.104)$$

In the generalization given in (7.49), this result extends to the more general situation in which the initial and final photon polarizations can differ arbitrarily. Transcribing (18.104) to $W^{\mu\nu}$, we find

$$2\,\mathrm{Im}\, W^{\mu\nu}(P,q) = \sum_f \int d\Pi_f \,\langle P|\, J^\mu(-q)\,|f\rangle \,\langle f|\, J^\nu(q)\,|P\rangle\,, \tag{18.105}$$

where $J^\mu(q)$ is the Fourier transform of the current.

We can now compute the deep inelastic cross section in terms of $W^{\mu\nu}$, using (18.105) to represent the square of the last factor. The cross section should be averaged over initial and summed over final electron spins. Thus,

$$\begin{aligned}\sigma(ep\to eX) = \frac{1}{2s}\int \frac{d^3k'}{(2\pi)^3}\frac{1}{2k'}e^4\frac{1}{2}\sum_{\text{spins}}\big[\bar u(k)\gamma_\mu u(k')\bar u(k')\gamma_\nu u(k)\big]\\ \cdot\Big(\frac{1}{Q^2}\Big)^2\cdot 2\,\mathrm{Im}\,W^{\mu\nu}(P,q).\end{aligned} \tag{18.106}$$

The electron spinor product can be evaluated as

$$\begin{aligned}\tfrac{1}{2}\sum_{\text{spins}}\big[\bar u(k)\gamma_\mu u(k')\bar u(k')\gamma_\nu u(k)\big] &= \tfrac{1}{2}\,\mathrm{tr}\big[\not k\gamma_\mu \not k'\gamma_\nu\big]\\ &= 2\big(k_\mu k'_\nu + k_\nu k'_\mu - g_{\mu\nu}k\cdot k'\big).\end{aligned} \tag{18.107}$$

It is useful to convert the integral over the final electron momentum k' and scattering angle θ to an integral over the dimensionless variables x and y that we introduced in Section 17.3. These variables are given in terms of the initial and final electron energies k and k' by

$$x = \frac{Q^2}{2P\cdot q} = \frac{2kk'(1-\cos\theta)}{2m(k-k')}, \qquad y = \frac{2P\cdot q}{2P\cdot k} = \frac{k-k'}{k}. \tag{18.108}$$

Then

$$\left|\frac{\partial(x,y)}{\partial(k',\cos\theta)}\right| = \frac{2k'}{2m(k-k')} = \frac{2k'}{ys}, \tag{18.109}$$

and so

$$\int \frac{d^3k'}{(2\pi)^3}\frac{1}{2k'} = \int \frac{2\pi dk' k' d\cos\theta}{(2\pi)^3\cdot 2} = \int dx dy \frac{ys}{(4\pi)^2}. \tag{18.110}$$

Using (18.107) and (18.110) to simplify (18.106), we find

$$\frac{d^2\sigma}{dxdy}(ep\to eX) = \frac{2\alpha^2 y}{(Q^2)^2}\big(k_\mu k'_\nu + k_\nu k'_\mu - g_{\mu\nu}k\cdot k'\big)\,\mathrm{Im}\,W^{\mu\nu}(P,q). \tag{18.111}$$

To go further, we need to know something about the structure of $W^{\mu\nu}$. In the previous section, we used current conservation to write the matrix element of currents in terms of a single scalar function $\Pi_h(q^2)$, as in Eq. (18.82). In the case of the forward Compton amplitude, the Ward identity again requires

$$q_\mu W^{\mu\nu} = q_\nu W^{\mu\nu} = 0, \tag{18.112}$$

but now there are two possible tensors built from P and q that satisfy these constraints. Thus the forward Compton amplitude is written as an expression involving two scalar form factors:

$$W^{\mu\nu} = \Big(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\Big) W_1 + \Big(P^\mu - q^\mu \frac{P\cdot q}{q^2}\Big)\Big(P^\nu - q^\nu \frac{P\cdot q}{q^2}\Big) W_2. \tag{18.113}$$

The scalar functions W_1, W_2 depend on the two invariants of the problem, $(P\cdot q)$ and q^2, or, alternatively, x and Q^2. If we insert (18.113) into (18.111) and use the fact that dotting q^μ with the lepton tensor gives zero, we find

$$\begin{aligned} \frac{d^2\sigma}{dx dy}(ep \to eX) &= \frac{2\alpha^2 y}{(Q^2)^2}\big[2k\cdot P k'\cdot P \operatorname{Im} W_2 + 2k\cdot k' \operatorname{Im} W_1\big] \\ &= \frac{\alpha^2 y}{Q^4}\big[s^2(1-y)\operatorname{Im} W_2 + 2xys \operatorname{Im} W_1\big]. \end{aligned} \tag{18.114}$$

Expression (18.114) is completely general and makes no assumptions about the nature of the strong interactions. It is also rather formal. However, we can easily get an idea of the relation of this formula to our earlier analysis by evaluating $W^{\mu\nu}$ in the parton model and working out the parton expressions for W_1 and W_2. In the parton model, we replace the proton matrix element in (18.102) by a sum of quark matrix elements, weighted with the parton distribution functions. Thus,

$$W^{\mu\nu} \approx i\int d^4x\, e^{iq\cdot x} \int_0^1 d\xi \sum_f f_f(\xi)\cdot \frac{1}{\xi}\langle q_f(p)|\, T\{J^\mu(x) J^\nu(0)\}\, |q_f(p)\rangle\Big|_{p=\xi P}. \tag{18.115}$$

The factor $(1/\xi)$ in front of the matrix element gives the proper normalization of the proton state in terms of the quark states. The simplest way to understand this factor is to note that the kinematic prefactor $(1/2s)$ in (18.106) and in other expressions involving an initial-state proton becomes $(1/2\xi s)$, under the ξ integral, in the parton model.

We now evaluate the matrix element in (18.115) using noninteracting fermions. There are two Feynman diagrams, shown in Fig. 18.10. The first diagram on the right in Fig. 18.10 has the value

$$i\int_0^1 d\xi \sum_f f_f(\xi)\frac{1}{\xi} Q_f^2 \bar{u}(p)\gamma^\mu \frac{i(\not{p}+\not{q})}{(p+q)^2 + i\epsilon}\gamma^\nu u(p); \tag{18.116}$$

the second diagram gives a contribution identical to this one after the interchange of q, μ with $(-q)$, ν. To evaluate (18.116), we average over the quark

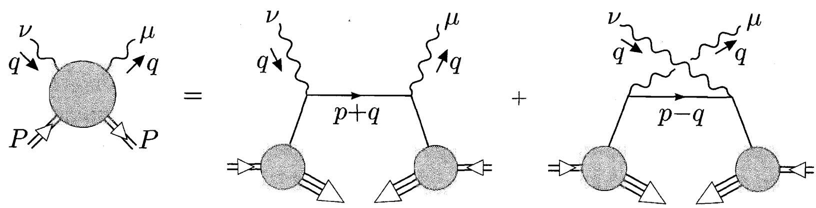


Figure 18.10. Evaluation of $W^{\mu\nu}$ in the parton model.

spin to find

$$\int_0^1 d\xi \sum_f f_f(\xi)\frac{1}{\xi}\cdot\frac{1}{2}\,\mathrm{tr}\big[\not{p}\gamma^\mu(\not{p}+\not{q})\gamma^\nu\big]\frac{-1}{2p\cdot q+q^2+i\epsilon}$$

$$= \int_0^1 d\xi \sum_f f_f(\xi)\frac{1}{\xi}\cdot 2\big(p^\mu(p+q)^\nu + p^\nu(p+q)^\mu - g^{\mu\nu}p\cdot(p+q)\big) \cdot \frac{-1}{2\xi P\cdot q - Q^2 + i\epsilon}. \tag{18.117}$$

The imaginary part of this expression, which we need to evaluate (18.114), comes from the last factor in (18.117):

$$\mathrm{Im}\Big(\frac{-1}{2\xi P\cdot q - Q^2 + i\epsilon}\Big) = \pi\delta(2\xi P\cdot q - Q^2) = \frac{\pi}{ys}\delta(\xi - x). \tag{18.118}$$

In the second diagram of Fig. 18.10, the two factors in the denominator have a relative + sign, so this diagram has no imaginary part in the physical region for deep inelastic scattering. Thus, we find that in the parton model,

$$\mathrm{Im}\,W^{\mu\nu} = \sum Q_f^2 f_f(x)\frac{1}{x}\frac{\pi}{ys}\big(4x^2P^\mu P^\nu + 2x(P^\mu q^\nu + P^\nu q^\mu) - g^{\mu\nu}xys\big). \tag{18.119}$$

By adding and subtracting terms proportional to $q^\mu q^\nu$, we can see that this expression is of the form (18.113), with

$$\mathrm{Im}\,W_1 = \pi\sum_f Q_f^2 f_f(x), \qquad \mathrm{Im}\,W_2 = \frac{4\pi}{ys}\sum_f Q_f^2 x f_f(x). \tag{18.120}$$

The parton model expressions for W_1 and W_2 obey the relation

$$\mathrm{Im}\,W_1 = \frac{ys}{4x}\,\mathrm{Im}\,W_2. \tag{18.121}$$

This is another form of the Callan-Gross relation, since the substitution of (18.121) into (18.114) gives

$$\frac{d^2\sigma}{dxdy}(ep\to eX) = \frac{\alpha^2 ys^2}{2Q^4}\big[1+(1-y)^2\big]\,\mathrm{Im}\,W_2, \tag{18.122}$$

with the y dependence characteristic of free fermions, as in Eq. (17.125). Finally, substituting from (18.120) for the imaginary part of W_2, we recover this parton model expression precisely:

$$\frac{d^2\sigma}{dxdy}(ep \to eX) = \frac{2\pi\alpha^2 s}{Q^4}\Big(\sum_f Q_f^2 x f_f(x)\Big)\big[1+(1-y)^2\big]. \tag{18.123}$$

This equation will give us a reference point for comparison with more general expressions that we will derive as we continue our analysis.

Expansion of the Operator Product

Since the forward Compton amplitude is a matrix element of a product of currents, an alternative strategy for calculating $W^{\mu\nu}$ is to expand this product as a series of local operators. Like the parton model evaluation, this method makes use of asymptotic freedom. However, in this case, the assumption is applied more directly. The computation of the operator product coefficients will take place explicitly at a small distance of order $1/Q$, and so we can calculate these coefficients in a perturbation theory whose coupling constant is $\alpha_s(Q^2)$.

In the previous section, we computed the coefficients of operators that contribute to the vacuum expectation value of the product of currents by considering the various ways of contracting the quark fields in the product. Here, we should note that the operator 1 does not contribute to the Compton scattering amplitude. The leading contributions come from operators that can create and annihilate quarks in the proton wavefunction.

The most important terms in the operator product of two currents J^μ come from products of two quark currents $\bar{q}_f\gamma^\mu q_f$ with quarks of the same flavor. Therefore we will begin by studying the OPE of the individual quark currents. To zeroth order in α_s, the leading terms of the operator product of quark currents are given by

$$\begin{aligned}&\bar{q}\gamma^\mu q(x)\,\bar{q}\gamma^\nu q(0)\\ &\quad= \bar{q}(x)\gamma^\mu \overline{q(x)\bar{q}(0)}\gamma^\nu q(0) + \bar{q}(x)\gamma^\mu \overline{q(x)\bar{q}(0)\gamma^\nu q(0)} + \cdots,\end{aligned} \tag{18.124}$$

where the contractions should be evaluated as Feynman propagators for the quark fields. The terms with explicit contractions are singular as $x \to 0$; the remaining terms are nonsingular and thus less important in the short-distance limit. In the OPE of currents with quarks of different flavor, there are no corresponding singular terms; we will argue below that this conclusion is valid even beyond the leading order in α_s.

To evaluate $W^{\mu\nu}$, we must take the Fourier transform of the terms in (18.124), as indicated in (18.102). When we do this, we should remember that the propagators carry not only the Fourier transform momentum q but also whatever momentum is carried in through the quark fields. To take account of this, it is convenient to represent

$$\int d^4x\, e^{iq\cdot x}\, \bar{q}(x)\gamma^\mu \overline{q(x)\bar{q}(0)}\gamma^\nu q(0) = \bar{q}\gamma^\mu \frac{i(i\partial\!\!\!/ + q\!\!\!/)}{(i\partial + q)^2}\gamma^\nu q(0), \tag{18.125}$$

where the derivatives ∂ act to the right on the quark field. Notice that this contribution has the structure of the first diagram on the right in Fig. 18.10. Similarly, the second contraction indicated in (18.124) has the form of the second diagram in Fig. 18.10.

In the short-distance limit, the momentum q will be larger than any external momentum entering the quark fields. Thus we should expand

$$\frac{1}{(i\partial + q)^2} = \frac{-1}{Q^2 - 2iq\cdot\partial + \partial^2} = -\frac{1}{Q^2}\sum_{n=0}^{\infty}\left(\frac{2iq\cdot\partial - \partial^2}{Q^2}\right)^n. \tag{18.126}$$

We will argue below that the terms with ∂^2 in the numerator are unimportant and may be dropped. However, we should retain all powers of the ratio $(2iq\cdot\partial/Q^2)$. This ratio has Q^2 in the denominator and so is formally suppressed in the short-distance limit. However, in the parton model

$$\frac{2iq\cdot\partial}{Q^2} \to \frac{2q\cdot\xi P}{Q^2} = 1, \tag{18.127}$$

so, eventually, all of these terms must be equally important. We will see how this works in a moment.

The last step required to reduce the operator product (18.124) to a useful form is to reduce the product of Dirac matrices. We know from (18.113) that, after we average over the proton spin, $W^{\mu\nu}$ will be symmetric under the interchange of μ and ν. Thus, it does no harm to symmetrize the OPE. We can then reduce the product of three Dirac matrices to one by using the identity

$$\tfrac{1}{2}\left(\gamma^\mu\gamma^\alpha\gamma^\nu + \gamma^\nu\gamma^\alpha\gamma^\mu\right) = g^{\mu\alpha}\gamma^\nu + \gamma^\mu g^{\alpha\nu} - g^{\mu\nu}\gamma^\alpha, \tag{18.128}$$

which is easily proved from the anticommutation relations. By the use of (18.126) and (18.128), we can rewrite (18.125) as

$$-i\bar{q}\Big(\gamma^\mu(i\partial^\nu) + \gamma^\nu(i\partial^\mu) - ig^{\mu\nu}\partial\!\!\!/ + \gamma^\mu q^\nu + \gamma^\nu q^\mu - g^{\mu\nu}q\!\!\!/\Big)\frac{1}{Q^2}\sum_{n=0}^{\infty}\left(\frac{2iq\cdot\partial}{Q^2}\right)^n q. \tag{18.129}$$

We can remove the term $(i\partial\!\!\!/)q$, which vanishes to leading order in α_s, since the quark field obeys the Dirac equation. To compute W_1 and W_2, we can also drop the terms with explicit factors of q^μ, since these will eventually be organized into the general form (18.113). Then, finally, (18.125) takes the form

$$\begin{aligned}\int d^4x\, e^{iq\cdot x}&\bar{q}(x)\gamma^\mu \overline{q(x)\bar{q}(0)}\gamma^\nu q(0)\\ &= -i\bar{q}\left(2\gamma^\mu(i\partial^\nu) - g^{\mu\nu}q\!\!\!/\right)\frac{1}{Q^2}\sum_{n=0}^{\infty}\left(\frac{2iq\cdot\partial}{Q^2}\right)^n q,\end{aligned} \tag{18.130}$$

symmetrized under $\mu \leftrightarrow \nu$.

The second term in (18.124) differs from the first by the interchange of the points x and 0 and the interchange of indices μ and ν. Its Fourier transform is thus given by (18.130) with the replacement $q \to -q$. The complete operator product therefore contains only terms even in q. All remaining contributions from the singular terms of the operator product contain the operator

$$\bar{q}\gamma^{\mu_1}(i\partial^{\mu_2})\cdots(i\partial^{\mu_k})q, \tag{18.131}$$

with an even number of indices, with these indices either identified with μ or ν or contracted with powers of q. To write the relevant terms of the operator product expansion, we will modify this operator in two ways. First, since the operator in (18.131) has n vector indices, it contains components that transform under many different irrreducible representations of the Lorentz group. Each component has a different rescaling law under renormalization. However, we will see below that only the component of (18.131) with the highest spin is relevant to our analysis. This component is obtained by totally symmetrizing the indices $\mu_1, \ldots, \mu_n$ and then subtracting terms proportional to $g^{\mu_i\mu_j}$ so that the operator is traceless on all pairs of indices. We will retain only this component when we write out the operator product of currents. Second, the operator (18.131) does not transform simply under gauge transformations. Since the original currents J^μ were invariant to color gauge transformations, the operator product of two currents must be a sum of gauge-invariant operators. We can make (18.131) gauge-invariant by replacing each factor of $(i\partial^\mu)$ with a covariant derivative (iD^μ). This modification adds only terms proportional to the strong coupling constant g, so it has no effect on our derivation of the operator product coefficients.

Incorporating these changes, let us define a spin-n operator with quarks of flavor f as follows:

$$\mathcal{O}_f^{(n)\mu_1\cdots\mu_n} = \bar{q}_f\gamma^{\{\mu_1}(iD^{\mu_2})\cdots(iD^{\mu_n\}})q_f - \text{traces}, \tag{18.132}$$

with indices symmetrized and with appropriate subtractions. We can use these operators to write a final expression for the most singular part of the OPE of two currents J^μ. The leading terms in this operator product come from (18.130) and the corresponding contraction with $q \leftrightarrow -q$. Extracting the pieces of these expressions that contain the highest spin operators (18.132), we find

$$\begin{aligned} i\int d^4x\, e^{iq\cdot x} J^\mu(x)J^\nu(0) \\ = \sum_f Q_f^2\Bigg[4\sum_{n=2}^{\infty}\frac{(2q^{\mu_1})\cdots(2q^{\mu_{n-2}})}{(Q^2)^{n-1}}\mathcal{O}_f^{(n)\mu\nu\mu_1\cdots\mu_{n-2}} \\ - g^{\mu\nu}\sum_{n=2}^{\infty}\frac{(2q^{\mu_1})\cdots(2q^{\mu_n})}{(Q^2)^n}\mathcal{O}_f^{(n)\mu_1\cdots\mu_n}\Bigg] + \cdots, \end{aligned} \tag{18.133}$$

where the sums over n run over even integers only.

Expression (18.133) has been derived in the leading order in α_s. Higher-order Feynman diagrams will contribute corrections to the coefficient functions of order $\alpha_s(Q^2)$. These corrections will be important only if they are multiplied by large logarithms. If we consider the operators $\mathcal{O}_f^{(n)}$ appearing on the right-hand side to be normalized at the renormalization scale Q, there is no large ratio of momenta available to enhance the QCD corrections to the coefficient functions. Large logarithmic corrections may still arise at a later stage of the calculation, when we compute the matrix elements of the operators $\mathcal{O}_f^{(n)}$.

From the expansion (18.133), it is straightforward to compute an expansion for $W^{\mu\nu}$ by taking its expectation value in the proton state. To carry out this computation, we need to know the proton matrix elements of the operators $\mathcal{O}_f^{(n)}$. Notice that these matrix elements cannot depend on the direction of the momentum q^μ, since that dependence has been isolated in the coefficient functions. This means that only the proton momentum P^μ is available to carry the vector indices of the matrix element. We can therefore write the spin-averaged matrix element of $\mathcal{O}_f^{(n)}$ as

$$\langle P|\, \mathcal{O}_f^{(n)\mu_1\cdots\mu_n} \,|P\rangle = A_f^n \cdot 2P^{\mu_1}\cdots P^{\mu_n} - \text{traces}. \tag{18.134}$$

The coefficients A_f^n are dimensionless. They are not quite pure numbers, because they depend on the renormalization scale of the operators, but we will treat them as constants in the next few paragraphs.

For the case $n = 1$, the operators $\mathcal{O}_f^{(1)}$ reduce simply to the quark flavor currents $\bar{q}\gamma^\mu q$; in this case the operators are normalized independently of any scale and the coefficients A_f^1 are truly constants. From our general discussion of form factors in Section 6.2, we know that the proton matrix element of a conserved flavor current at zero momentum transfer is given by

$$\langle P|\, \bar{q}_f\gamma^\mu q_f \,|P\rangle = \bar{u}(P)\gamma^\mu u(P) F_{f1}(0), \tag{18.135}$$

where $F_{f1}(0)$ is equal to the value of the corresponding conserved charge in the proton state. For the quark currents, this charge is just the number of quarks (minus antiquarks) of flavor f in the state $|P\rangle$, which we will call N_f. Averaging (18.135) over the proton spin, we find

$$\langle P|\, \bar{q}_f\gamma^\mu q_f \,|P\rangle = 2P^\mu \cdot N_f. \tag{18.136}$$

Thus, for $n = 1$,

$$A_f^1 = N_f = \begin{cases} 2 & f = u, \\ 1 & f = d. \end{cases} \tag{18.137}$$

Similarly, $\mathcal{O}_f^{(2)}$ is the contribution of the quark flavor f to the energy-momentum tensor of QCD:

$$\left(T^{\mu\nu}\right)_f = \bar{q}_f\gamma^{\{\mu}(iD^{\nu\}})q_f. \tag{18.138}$$

Thus, A_f^2 is the fraction of the total energy-momentum of the proton that is carried by the quark flavor f.

When we evaluate the series for $W^{\mu\nu}$ using (18.133) and the expression (18.134) for the operator matrix elements, we find

$$W^{\mu\nu} = \sum_f Q_f^2 \left[8 \sum_n P^\mu P^\nu \frac{(2q\cdot P)^{n-2}}{(Q^2)^{n-1}} A_f^n - 2g^{\mu\nu} \sum_n \frac{(2q\cdot P)^n}{(Q^2)^n} A_f^n \right] + \cdots, \tag{18.139}$$

where the sums over n run over even integers from 2 to infinity. In addition to the corrections to the OPE omitted in (18.133), we have also dropped contributions from the trace terms in (18.134). This is quite appropriate: In each of these terms, two factors of the proton momentum $P^\alpha P^\beta$ are replaced by $g^{\alpha\beta} m_p^2$, were $m_p^2 = P^2$ is the proton mass. When the indices are contracted with powers of q, we obtain a term of order

$$m_p^2 Q^2 \ll (2q\cdot P)^2. \tag{18.140}$$

Since $(Q^2/2P\cdot q) = x$, which is held fixed in deep inelastic scattering as Q^2 becomes large, the contribution from the trace terms is suppressed by a factor m_p^2/Q^2, times powers of x.

In general, an operator of dimension d has a coefficient function in the operator product expansion of currents that has dimension $(\text{mass})^{6-d}$; in the Fourier transform of the OPE, this coefficient function will carry a suppression factor

$$\left(\frac{1}{Q}\right)^{d-2}. \tag{18.141}$$

However, if the operator has spin s, the operator matrix element will contribute s factors of the vector P^μ, so that, in the kinematic region of deep inelastic scattering, the contribution will be of order

$$\left(\frac{2P\cdot q}{Q^2}\right)^s \left(\frac{1}{Q}\right)^{d-s-2}. \tag{18.142}$$

Thus, the relative size of contributions from the OPE to deep inelastic scattering is controlled, not exactly by the dimension of the operator, but rather by the *twist*, defined as

$$t = d - s. \tag{18.143}$$

In our selection of the leading terms in the operator product expansion of currents, we have consistently kept the contribution of leading spin for each dimension or for each power of Q^{-1} in the coefficient. The operators $\mathcal{O}_f^{(n)}$ all have twist $t = 2$, which is the smallest possible value for QCD operators other than the operator 1.

In the operator product of two different flavor currents—for example, $\bar{u}\gamma^\mu u$ and $\bar{d}\gamma^\nu d$—the leading terms in the OPE have the quark structure $(\bar{u}\Gamma u \bar{d}\Gamma d)$ and thus have twist $t \geq 4$. Thus, to all orders in α_s, the cross terms in the operator product of currents J^μ are suppressed by at least a factor

$(1/Q^2)$ relative to the leading-twist terms presented in (18.133). If we neglect these suppressed terms, the expression for $W^{\mu\nu}$ separates, to all orders, into a sum of contributions

$$W^{\mu\nu} = \sum_f Q_f^2 W_f^{\mu\nu}, \tag{18.144}$$

where $W_f^{\mu\nu}$ is the matrix element of two quark flavor currents $\bar{q}_f \gamma^\mu q_f$.

We can read from (18.139) the following expressions for W_1 and W_2:

$$\begin{aligned} W_1 &= \sum_f Q_f^2 \sum_n 2 \frac{(2q\cdot P)^n}{(Q^2)^n} A_f^n, \\ W_2 &= \sum_f Q_f^2 \sum_n \frac{8}{Q^2} \frac{(2q\cdot P)^{n-2}}{(Q^2)^{n-2}} A_f^n, \end{aligned} \tag{18.145}$$

where the sum over n in each line runs over even integers from 2 to infinity. Like (18.139), these expressions explicitly separate according to (18.144). It is noteworthy that the series (18.145) satisfy the Callan-Gross relation in the form (18.121), without further parton model input. However, this relation is corrected in order α_s due to the next-order contributions to the operator product coefficients.

Because the leading contributions to the deep inelastic form factors can be written as sums over quark flavors, it is tempting to reverse the logic of Eq. (18.120) and use these equations to define the parton distribution functions. In particular, let us define

$$x f_f^+(x, Q^2) = \frac{ys}{4\pi} \operatorname{Im} W_{2f}(x, Q^2), \tag{18.146}$$

where W_{2f} is the second form factor of $W_f^{\mu\nu}$, defined in (18.144), neglecting terms suppressed by powers of Q^2. In the parton model evaluation,

$$f_f^+(x) = f_f(x) + f_{\bar{f}}(x). \tag{18.147}$$

From (18.123) and the definition (18.146), we know that $f_f^+(x)$ enters in the correct way into the formula for the deep inelastic scattering cross section. However, parton distribution functions have other important properties, including the normalization conditions (17.36) and (17.39) and the evolution with Q^2 discussed in Section 17.6. We must now see whether we can derive these properties from (18.146) using the operator product expansion.

The Dispersion Integral

The operator product analysis has given us explicit expressions for W_1 and W_2 as a series in inverse powers of Q^2. In the following discussion, we will concentrate on the analysis of W_2. We must work out the relation of its series expansion to the observable deep inelastic scattering cross section. As in the discussion of Section 18.4, the OPE analysis naturally takes place in an unphysical kinematic region. To make the operator product expansion, we

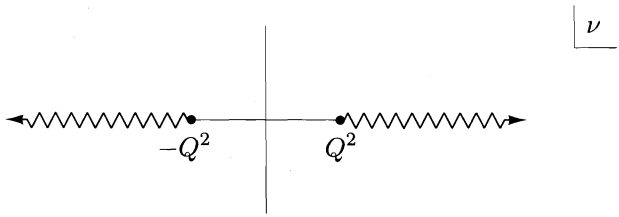


Figure 18.11. Analytic singularities of $W_2(\nu, Q^2)$ in the complex ν plane, for fixed Q^2.

needed to consider Q^2 to be larger than any other kinematic invariant. However, in the physical region for deep inelastic scattering, $2P \cdot q \geq Q^2$. We need a formula that connects these two distinct regions.

To state this problem more precisely, define

$$\nu = 2P \cdot q = ys; \tag{18.148}$$

in the frame in which the proton is at rest, $\nu = 2m_p q^0$. The form factor W_2 can be viewed as a function of ν and Q^2. Then, for fixed Q^2, the OPE gives a series expansion about the point $\nu = 0$, while the physical region for deep inelastic scattering is $\nu \geq Q^2$. Because this region is associated with a physical scattering process, $W_2(\nu, Q^2)$, viewed as an analytic function of ν for fixed Q^2, will have a branch cut along the real ν axis in this region. The discontinuity across this branch cut will be $(2i)$ times the imaginary part of W_2, which appears in the expression (18.123) for the deep inelastic cross section. Because expression (18.102) is symmetric under the interchange of (q, μ) and $(-q, \nu)$, W_2 must obey

$$W_2(-\nu, Q^2) = W_2(\nu, Q^2). \tag{18.149}$$

Thus, W_2 must also have a branch cut along the negative real axis, from $\nu = -Q^2$ to $-\infty$. The discontinuity across this cut gives the cross section for the u-channel process in which positive energy comes in through the second current and out through the first. Since $q^2 = -Q^2 < 0$, there is no possible physical t-channel process; thus W_2 has no further singularities in the complex ν plane. The analytic structure of $W_2(\nu, Q^2)$ is shown in Fig. 18.11.

Now consider the contour integral

$$I_n = \int \frac{d\nu}{2\pi i} \frac{1}{\nu^{n-1}} W_2(\nu, Q^2), \tag{18.150}$$

for n even, taken on a small circle surrounding the origin. This integral picks out the coefficient of ν^{n-2} in the series expansion for W_2. The OPE formula (18.145) gives us the leading contribution to this coefficient for large Q^2:

$$I_n = \sum_f Q_f^2 \frac{8}{(Q^2)^{n-1}} A_f^n. \tag{18.151}$$

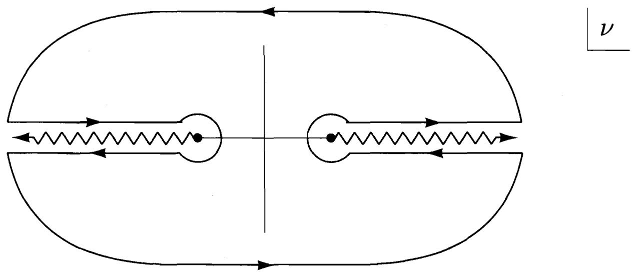


Figure 18.12. Contour of integration involved in the derivation of the moment sum rules for W_2.

The corrections to this formula are of order $\alpha_s(Q^2)$, from the evaluation of the OPE coefficient functions.

On the other hand, we can also distort the contour as shown in Fig. 18.12 and evaluate it as an integral over the discontinuities of W_2. By the symmetry (18.149), the two branch cuts give equal contributions. Thus,

$$I_n = 2\int_{Q^2}^{\infty} \frac{d\nu}{2\pi i}\frac{1}{\nu^{n-1}}(2i)\,\mathrm{Im}\,W_2(\nu, Q^2). \tag{18.152}$$

Now change variables to $x = Q^2/\nu$. The integral becomes

$$I_n = \frac{8}{(Q^2)^{n-1}}\int_0^1 dx\, x^{n-2}\frac{\nu}{4\pi}\,\mathrm{Im}\,W_2. \tag{18.153}$$

When we equate (18.151) and (18.153) and relate $\mathrm{Im}\,W_2$ to the parton distributions $f_f^+(x)$ using (18.146), the relation we have derived splits into a series of sum rules,

$$\int_0^1 dx\, x^{n-1} f_f^+(x, Q^2) = A_f^n, \tag{18.154}$$

for n even. These relations are known as the *moment sum rules* for the deep inelastic form factors. They relate the x moments of the parton distribution functions, as defined by Eq. (18.146), to the proton matrix elements of twist-2 operators.

Because W_2 is a symmetric function of ν, the moment sum rules apply only for even n. However, in deep inelastic neutrino scattering, there is a third form factor in $W^{\mu\nu}$, associated with the interference term between the vector and axial vector parts of the weak interaction current. In Problem 18.2, we show that this form factor can be used to derive a set of sum rules for odd n:

$$\int_0^1 dx\, x^{n-1} f_f^-(x, Q^2) = A_f^n, \tag{18.155}$$

where A_f^n is the coefficient of the proton matrix element (18.134) for odd n, and $f_f^-(x)$ is a form factor which, in the parton model, evaluates to

$$f_f^-(x) = f_f(x) - f_{\bar{f}}(x). \tag{18.156}$$

Combining this information with the argument given below (18.136), we can see that the definition of the parton distribution functions from the deep inelastic form factors has the correct normalization. Using (18.137), we find

$$\int_0^1 dx\, f^-(x) = N_f, \tag{18.157}$$

the (net) number of quarks of flavor f in the proton. Similarly, (18.154) and (18.138) imply

$$\int_0^1 dx\, x f^+(x) = \langle x \rangle_f\,, \tag{18.158}$$

where $\langle x \rangle_f$ is the fraction of the total energy-momentum of the proton carried by quarks and antiquarks of flavor f.

Operator Rescaling

If the coefficients A_f^n were truly constants, relations (18.154) and (18.155) would be consistent with parton distribution functions that satisfy Bjorken scaling. However, as we remarked below (18.134), these factors actually depend on Q^2, since this is the normalization point of the operators in the operator product expansion (18.133). Since this dependence comes only through operator rescaling, it involves only logarithms of Q^2, and so contributes only a slow violation of Bjorken scaling. We can work out the Q^2 dependence of the parton distribution functions quantitatively by summing the leading logarithmic corrections to the matrix elements of the twist-2 operators.

To account for these corrections, let us first assume (incorrectly, as we will see below) that the twist-2 operators (18.132) are renormalized without operator mixing. Then the leading logarithmic corrections to the matrix element of the operator $\mathcal{O}_f^{(n)}$ would be summed by rescaling the operator normalized at Q to operators normalized at a standard reference point μ, of order 1 GeV. The relation between these conventions would be

$$\left[\mathcal{O}_f^{(n)}\right]_Q = \left(\frac{\log(Q^2/\Lambda^2)}{\log(\mu^2/\Lambda^2)}\right)^{a_f^n/2b_0} \left[\mathcal{O}_f^{(n)}\right]_\mu, \tag{18.159}$$

where a_f^n is the first coefficient of the γ function of $\mathcal{O}_f^{(n)}$. Then the factors A_f^n would depend on Q^2 according to

$$A_f^n(Q^2) = \left(\frac{\log(Q^2/\Lambda^2)}{\log(\mu^2/\Lambda^2)}\right)^{a_f^n/2b_0} A_f^n(\mu^2). \tag{18.160}$$

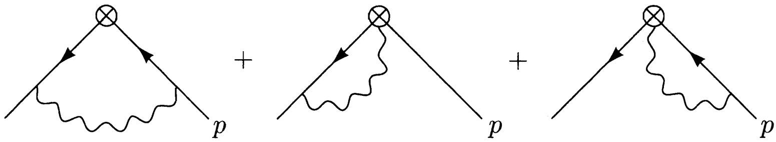
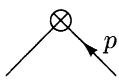


Figure 18.13. Diagrams contributing the anomalous dimension of the quark twist-2 operators.

This equation agrees with the scale dependence of operator product coefficients written in (18.76), for the special case of an operator product of currents, $a_1 = a_2 = 0$. To find the explicit form of the rescaling factor, we must compute a_f^n.

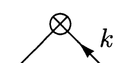
To compute the γ functions of the quark twist-2 operators, we must compute their counterterms for operator rescaling. These are determined by the diagrams shown in Fig. 18.13. It suffices to compute these diagrams with external momentum p entering through the quark line and zero external momentum injected into the operator. Under these conditions, the matrix element of the operator $\mathcal{O}_f^{(n)}$, in leading order, equals



$$= \gamma^{\mu_1} p^{\mu_2} \cdots p^{\mu_n}. \tag{18.161}$$

Here and at all later points in the discussion, we will treat the matrix elements of $\mathcal{O}_f^{(n)}$ as though they are symmetrized in the n indices and have all possible traces subtracted. We must now evaluate the diagrams of Fig. 18.13 and collect all terms that rescale this structure.

The first diagram of Fig. 18.13 is quite straightforward to evaluate:



$$\begin{aligned} &= \int \frac{d^4k}{(2\pi)^4} (ig)^2 \gamma^\lambda t^a \frac{i\not{k}}{k^2} \gamma^{\mu_1} k^{\mu_2} \cdots k^{\mu_n} \frac{i\not{k}}{k^2} \gamma_\lambda t^a \frac{-i}{(k-p)^2} \\ &= -ig^2 C_2(r) \cdot (-2) \int \frac{d^4k}{(2\pi)^4} \frac{1}{(k^2)^2 (k-p)^2} \not{k} \gamma^{\mu_1} k^{\mu_2} \cdots k^{\mu_n} \not{k}. \end{aligned} \tag{18.162}$$

We combine denominators using identity (6.40):

$$\frac{1}{(k^2)^2 (k-p)^2} = \int_0^1 dx \, \frac{2(1-x)}{(\mathbf{k}^2 - \Delta)^3}; \tag{18.163}$$

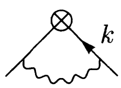
the quantities in the denominator on the right are $\mathbf{k} = k - xp$ and $\Delta = -x(1-x)p^2$. We must now shift the integral, substitute $k = \mathbf{k} + xp$ in the numerator, and pick out a term proportional to $(n-1)$ powers of p. If this term contains the factor $g^{\mu_i \mu_j}$, we may drop it, since it contributes to the coefficient of an operator of higher twist and since, in any event, it will be removed when we subtract traces. Thus, we must choose carefully which two factors of k we replace with $\mathbf{k}$ when we replace the others with (xp). The following choices, simplified using the rotational symmetry of the **k** integral, do not give useful contributions:

$$\begin{aligned} \mathbf{k}^{\mu_i}\mathbf{k}^{\mu_j} &= \frac{1}{4}\mathbf{k}^2 g^{\mu_i\mu_j}, \\ \not{\mathbf{k}}\gamma^{\mu_1}\mathbf{k}^{\mu_j} &\rightarrow \frac{1}{4}\mathbf{k}^2\gamma^{\mu_j}\gamma^{\mu_1} = \frac{1}{4}\mathbf{k}^2 g^{\mu_1\mu_j}. \end{aligned} \tag{18.164}$$

In the second line, we have used the symmetry under $\mu_1 \leftrightarrow \mu_j$. The one remaining placement of the factors of **k** is

$$\not{\mathbf{k}}\gamma^{\mu_1}\not{\mathbf{k}} \rightarrow \frac{1}{4}\mathbf{k}^2\gamma^\nu\gamma^{\mu_1}\gamma_\nu = -\frac{1}{2}k^2\gamma^{\mu_1}. \tag{18.165}$$

Thus (18.162) has the value



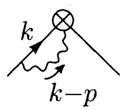
$$\begin{aligned} &= -ig^2\frac{4}{3}\int_0^1 dx\cdot 2(1-x)\int\frac{d^4\mathbf{k}}{(2\pi)^4}\frac{\mathbf{k}^2}{(\mathbf{k}^2-\Delta)^3}\gamma^{\mu_1}(xp^{\mu_2})\cdots(xp^{\mu_n}) \\ &= -i\frac{8}{3}g^2\int_0^1 dx(1-x)x^{n-1}\frac{i}{(4\pi)^2}\Gamma(2-\tfrac{d}{2})\gamma^{\mu_1}p^{\mu_2}\cdots p^{\mu_n} \\ &= \frac{4}{3}\frac{2}{n(n+1)}\frac{g^2}{(4\pi)^2}\Gamma(2-\tfrac{d}{2})\gamma^{\mu_1}p^{\mu_2}\cdots p^{\mu_n}. \end{aligned} \tag{18.166}$$

It is not so obvious that there are additional contributions to the rescaling of the operators $\mathcal{O}_f^{(n)}$. Note, however, that the covariant derivatives in (18.132) contain explicit factors of the gauge field,

$$iD^{\mu_j} = i\partial^{\mu_j} - gA^{a\mu_j}t^a, \tag{18.167}$$

and these may be contracted with gauge field vertices on the external legs. These contributions give rise to the second and third diagrams in Figure 18.13. The term in which two factors of $A^{a\mu}$ from (18.167) are contracted with one another is proportional to $G^{\mu_i\mu_j}$ and thus does not contribute to the rescaling of the leading-twist operators.

The contributions we have just described have the form of sums over j, where μ_j is the index of the derivative that includes the contraction. Then the second diagram of Fig. 18.3 is the sum over j of the following integral:

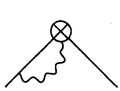


$$\begin{aligned} &= \int\frac{d^4k}{(2\pi)^4}(ig)\gamma_\lambda t^a\frac{i\not{k}}{k^2}\gamma^{\mu_1}k^{\mu_2}\cdots k^{\mu_{j-1}} \\ &\qquad\cdot(-gt^a g^{\lambda\mu_j})p^{\mu_{j+1}}\cdots p^{\mu_n}\frac{-i}{(k-p)^2} \\ &= ig^2C_2(r)\int\frac{d^4k}{(2\pi)^4}\frac{1}{k^2(k-p)^2}\gamma^{\mu_j}\not{k}\gamma^{\mu_1}k^{\mu_2}\cdots k^{\mu_{j-1}}p^{\mu_{j+1}}\cdots p^{\mu_n}. \end{aligned} \tag{18.168}$$

Since μ_j and μ_1 are symmetrized, we can use (18.128) to rewrite

$$\begin{aligned}\gamma^{\mu_j}\not{k}\gamma^{\mu_1} &\to k^{\mu_j}\gamma^{\mu_1}+\gamma^{\mu_j}k^{\mu_1}-g^{\mu_j\mu_1}\not{k}\\ &\to 2\gamma^{\mu_1}k^{\mu_j},\end{aligned} \tag{18.169}$$

where, in the second line, the symmetrization of indices and subtraction of traces is understood. Now combine denominators. To obtain a term with $(n-1)$ factors of p, we must replace every factor k in the numerator of (18.168) with (xp). This gives



$$\begin{aligned} &= ig^2\frac{4}{3}\int_0^1 dx\int\frac{d^4\mathbf{k}}{(2\pi)^4}\frac{1}{(\mathbf{k}^2-\Delta)^2}2\gamma^{\mu_1}(xp^{\mu_2})\cdots(xp^{\mu_j})p^{\mu_{j+1}}\cdots p^{\mu_n}\\ &= ig^2\frac{8}{3}\int_0^1 dx\,x^{j-1}\frac{i}{(4\pi)^2}\Gamma(2-\tfrac{d}{2})\gamma^{\mu_1}p^{\mu_2}\cdots p^{\mu_n}\\ &= -\frac{4}{3}\frac{2}{j}\frac{g^2}{(4\pi)^2}\Gamma(2-\tfrac{d}{2})\gamma^{\mu_1}p^{\mu_2}\cdots p^{\mu_n}.\end{aligned} \tag{18.170}$$

This contribution must be summed over j from 2 to n. The third diagram of Fig. 18.13 makes an equal contribution.

Summing the rescaling factors from the three diagrams of Fig. 18.13, we find for the operator rescaling counterterm of $\mathcal{O}_f^{(n)}$

$$\delta_f=\frac{g^2}{(4\pi)^2}\frac{4}{3}\left[4\sum_{j=2}^{n}\frac{1}{j}-\frac{2}{n(n+1)}\right]\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}. \tag{18.171}$$

From this result, we can derive the Callan-Symanzik γ function by the use of (18.23) and the field strength renormalization counterterm (18.9). We find

$$\gamma_f^n=\frac{8}{3}\frac{g^2}{(4\pi)^2}\left[1+4\sum_{2}^{n}\frac{1}{j}-\frac{2}{n(n+1)}\right]. \tag{18.172}$$

Notice that this expression vanishes for $n=1$, so that there is no rescaling of A_f^1, as required by (18.157). For $n>1$, γ_f^n is positive and so its coefficient a_f^n is negative. This implies that the higher moments of the quark distribution functions are suppressed as Q^2 becomes large.

Operator Mixing

The QCD rescaling of the operators $\mathcal{O}_f^{(n)}$ is still more complicated because QCD contains additional twist-2 operators which can be built from gluon fields. These new operators are mixed with the quark twist-2 operators by the diagrams of Fig. 18.14.

For n even, the diagrams of Fig. 18.14 give the operators $\mathcal{O}_f^{(n)}$ matrix elements in the state of a gluon with momentum p. The tensor structure of

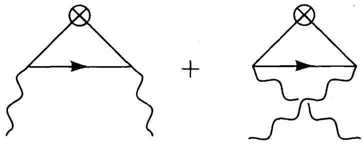


Figure 18.14. Diagrams that produce operator mixing between twist-2 quark and gluon operators.

this matrix element contains the term

$$g^{\alpha\beta}p^{\mu_1}\cdots p^{\mu_n}, \tag{18.173}$$

where α, β are the polarization indices of the external gluons. This structure arises from the operator

$$\mathcal{O}_g^{(n)\mu_1\cdots\mu_n} = -F^{\{\mu_1\nu}(iD^{\mu_2})\cdots(iD^{\mu_{n-1}})F^{\mu_n\}}{}_\nu - \text{traces}, \tag{18.174}$$

symmetrized on $\mu_1, \cdots \mu_n$, with traces subtracted. These operators have dimension $(n+2)$ and spin n, and thus have twist 2.

The gluon operators (18.174) are relevant only for n even. Using the manipulation

$$F^{\mu_1\nu}(iD^{\mu_2})\cdots F^{\mu_n}{}_\nu = i\partial^{\mu_2}\big(F^{\mu_1\nu}\cdots F^{\mu_n}{}_\nu\big) - (iD^{\mu_2})F^{\mu_1\nu}\cdots F^{\mu_n}{}_\nu, \tag{18.175}$$

we can transfer the covariant derivatives from one factor of $F^{\mu\nu}$ to the other, giving

$$\mathcal{O}_g^{(n)} = (-1)^n\mathcal{O}_g^{(n)} + \partial^{\mu_1}\big(\mathcal{O}'\big). \tag{18.176}$$

Thus, for n odd, the operator $\mathcal{O}_g^{(n)}$ is equal to a total derivative. The matrix elements of a total derivative are proportional to the momentum injected into this operator. Since zero momentum is injected in the calculation of the proton matrix elements of the OPE of currents, the operators $\mathcal{O}_g^{(n)}$ have no effect on the deep inelastic scattering cross section for n odd.

For n even, however, we must take account of the mixing of $\mathcal{O}_g^{(n)}$ with $\mathcal{O}_f^{(n)}$. The computation of the diagrams of Fig. 18.14 is quite similar to the other operator rescaling calculations we have done in this chapter, and so we reserve working out the details for Problem 18.3. We find that the diagrams of Fig. 18.14 contain a structure proportional to (18.173) with the coefficient

$$\frac{2(n^2+n+2)}{n(n+1)(n+2)}\frac{g^2}{(4\pi)^2}\Gamma(2-\tfrac{d}{2}). \tag{18.177}$$

From this computation, we find that the renormalized twist-2 quark operator, properly normalized at the scale M, is given in terms of bare operators by

$$\big[\mathcal{O}_f^{(n)}\big]_M = (1+\delta_f)\big[\mathcal{O}_f^{(n)}\big]_0 + (\delta_g)\big[\mathcal{O}_g^{(n)}\big]_0, \tag{18.178}$$

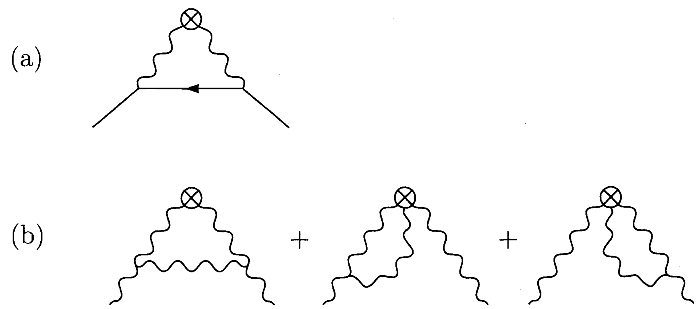


Figure 18.15. Diagrams contributing to the operator rescaling of twist-2 gluon operators: (a) contributions to gluon-quark mixing; (b) contributions to diagonal gluon operator renormalization.

where δ_f is given by (18.171) and

$$\delta_g = -\frac{g^2}{(4\pi)^2}\frac{2(n^2+n+2)}{n(n+1)(n+2)}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}. \tag{18.179}$$

This equation gives us two elements of the anomalous dimension matrix of twist-2 operators.

The remaining elements of the γ matrix for twist-2 operators are generated by the diagrams shown in Fig. 18.15. The diagram of Fig. 18.15(a) gives the mixing of $\mathcal{O}_g^{(n)}$ back into $\mathcal{O}_f^{(n)}$. The diagrams of Fig. 18.15(b), combined with the counterterm δ_3 for gluon field strength rescaling, gives the diagonal anomalous dimension. The counterterm δ_3 is given explicitly, in Feynman-'t Hooft gauge, in (16.74). The remainder of this anomalous dimension computation is discussed in Problem 18.3.

To describe the complete anomalous dimension matrix, we begin by considering a strong interaction model with one quark flavor. In this case, there is one twist-two operator $\mathcal{O}_f^{(n)}$ which mixes with $\mathcal{O}_g^{(n)}$. These two operators mix through a 2×2 matrix

$$\gamma^n = -\frac{g^2}{(4\pi)^2}\begin{pmatrix} a_{ff}^n & a_{fg}^n \\ a_{gf}^n & a_{gg}^n \end{pmatrix}, \tag{18.180}$$

where

$$\begin{aligned} a_{ff}^n &= -\frac{8}{3}\Big[1 + 4\sum_2^n \frac{1}{j} - \frac{2}{n(n+1)}\Big], \\ a_{fg}^n &= 4\frac{n^2+n+2}{n(n+1)(n+2)}, \\ a_{gf}^n &= \frac{16}{3}\frac{n^2+n+2}{n(n^2-1)}, \\ a_{gg}^n &= -6\Big[\frac{1}{3} + \frac{2}{9}n_f + 4\sum_2^n \frac{1}{j} - \frac{4}{n(n-1)} - \frac{4}{(n+1)(n+2)}\Big]. \end{aligned} \tag{18.181}$$

Notice that this matrix is not symmetric. In the last line, n_f is the number of quark flavors, equal to 1 in this case; this term comes from (16.74).

In the realistic case, QCD contains several quark flavors—u, d, s, and also c and b when we work at momenta sufficiently large that we can ignore the masses of these particles. Then the anomalous dimension matrix γ^n has size $(n_f+1)\times(n_f+1)$. The submatrix acting on quark operators is diagonal, with all of the diagonal entries being given by a^n_{ff} in (18.181). The quark-gluon and gluon-quark entries are all given by a^n_{fg} and a^n_{gf}, respectively, and are independent of the flavor. The gluon diagonal entry is given by a^n_{gg} in (18.181) with the realistic value of n_f. This means that the gluon operator mixes with only one linear combination of quark operators:

$$\sum_f \mathcal{O}_f^{(n)}; \tag{18.182}$$

the orthogonal linear combinations are simply rescaled, with the exponent given by a^n_{ff} or (18.172).

Let us now apply this analysis of operator mixing to the evaluation of the moment sum rules. For odd n, there is no operator mixing, and so the Q^2 dependence of the right-hand side of (18.155) is correctly given by the simple rescaling (18.160).

For even n, we must take operator mixing into account. The right-hand side of the sum rule (18.154) is the proton matrix element of a twist-2 operator normalized at the scale Q. Let us write an arbitrary linear combination of these operators as

$$c_i^n\left[\mathcal{O}_i^{(n)}\right]_Q, \tag{18.183}$$

where the index i runs over g and the various flavors f. To rescale this operator to a fixed reference momentum μ, we rewrite the coefficients in a basis of *left* eigenvectors of γ^n and rescale each eigenvector acccording to (18.159). In terms of the matrix a^n_{ij} of rescaling coefficients, we can write the resçaling abstractly as

$$c_i^n\left[\mathcal{O}_i^{(n)}\right]_Q, = c_i^n\left\{\left(\frac{\log(Q^2/\Lambda^2)}{\log(\mu^2/\Lambda^2)}\right)^{a^n/2b_0}\right\}_{ij}\left[\mathcal{O}_j^{(n)}\right]_\mu. \tag{18.184}$$

This rescaling, acting with c_i^n to the left of the matrix (a^n), is precisely the prescription required by Eq. (18.79).

Let us work this out explicitly for the case $n=2$. The right-hand side of the moment sum rule (18.154) is given by the matrix element of $\mathcal{O}_f^{(2)}$. We rewrite this as

$$\mathcal{O}_f^{(2)} = \left(\mathcal{O}_f^{(2)} - \frac{1}{n_f}\sum_{f'}\mathcal{O}_{f'}^{(2)}\right) + \frac{1}{n_f}\sum_{f'}\mathcal{O}_{f'}^{(2)}. \tag{18.185}$$

The first term is simply rescaled; the second term mixes with the gluon operator $\mathcal{O}_g^{(2)}$. The anomalous dimension matrix acting on $(\sum_f \mathcal{O}_f, \mathcal{O}_g)$ for $n=2$

has coefficients

$$\begin{pmatrix} a^2_{ff} & a^2_{fg}n_f \\ a^2_{gf} & a^2_{gg} \end{pmatrix} = \begin{pmatrix} -\frac{64}{9} & \frac{4}{3}n_f \\ \frac{64}{9} & -\frac{4}{3}n_f \end{pmatrix}. \tag{18.186}$$

The left eigenvectors of this matrix, and their corresponding eigenvalues, are

$$\begin{aligned} (1,1) \quad &\rightarrow \quad a^2 = 0 \\ (\frac{16}{3}, -n_f) \quad &\rightarrow \quad a^2 = -\frac{4}{3}\Big(\frac{16}{3} + n_f\Big). \end{aligned} \tag{18.187}$$

Notice that the first eigenvector gives a linear combination of operators $c_i^2 \mathcal{O}_i^{(2)}$ with zero anomalous dimension. This operator is in fact the total energy momentum tensor of QCD,

$$T^{\mu\nu} = \sum_f \mathcal{O}_f^{(2)\mu\nu} + \mathcal{O}_g^{(2)\mu\nu}, \tag{18.188}$$

which must have $\gamma = 0$. If we expand the second term in (18.185) in terms of the components (18.187), we can compute the full form of the operator rescaling. We find

$$\begin{aligned} \big[\mathcal{O}_f^{(2)}\big]_Q &= \frac{1}{16/3 + n_f} T \\ &+ \frac{1}{n_f(\frac{16}{3} + n_f)} \left(\frac{\log(Q^2/\Lambda^2)}{\log(\mu^2/\Lambda^2)}\right)^{-\frac{4}{3}(\frac{16}{3}+n_f)/2b_0} \Big[\frac{16}{3}\sum_f \mathcal{O}_f^{(2)} - n_f \mathcal{O}_g^{(2)}\Big]_\mu \\ &+ \left(\frac{\log(Q^2/\Lambda^2)}{\log(\mu^2/\Lambda^2)}\right)^{-32/9b_0} \Big[\mathcal{O}_f^{(2)} - \frac{1}{n_f}\sum_{f'} \mathcal{O}_{f'}^{(2)}\Big]_\mu, \end{aligned} \tag{18.189}$$

where T is the energy-momentum tensor (18.188). The right-hand side of the $n = 2$ moment sum rule is given by the coefficient of the proton matrix element of this operator. To evaluate this coefficient, we need to define gluon analogues of the A_f^n, by writing, analogously to (18.134),

$$\langle P|\, \mathcal{O}_g^{(n)\mu_1\cdots\mu_n} \,|P\rangle = A_g^n \cdot 2P^{\mu_1}\cdots P^{\mu_n} - \text{traces}. \tag{18.190}$$

For the case $n = 2$, we note in particular that

$$\langle P|\, T^{\mu\nu} \,|P\rangle = 2P^\mu P^\nu; \tag{18.191}$$

thus, (18.188) implies

$$\sum_f A_f^2 + A_g^2 = 1. \tag{18.192}$$

If we replace each operator in (18.189) by the corresponding coefficient $A_i^2(\mu)$, we will have an expression for the right-hand side of the $n = 2$ moment sum rule which makes its Q^2 dependence explicit.

Although expression (18.189) is rather complicated, it has a simple form in the extreme limit $Q^2 \to \infty$. At asymptotic Q^2, the last two terms of (18.189)

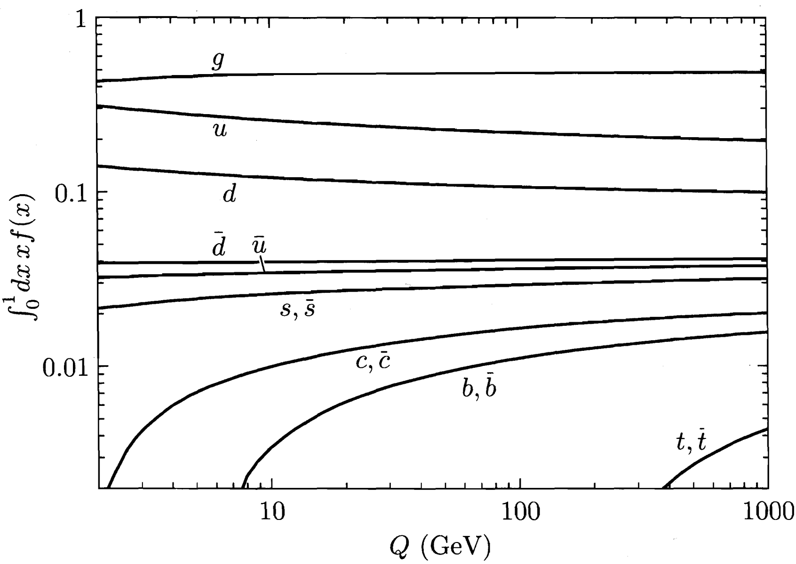


Figure 18.16. Fractions of the total energy-momentum of the proton carried by various parton species, as a function of Q, according to the CTEQ fit to deep inelastic scattering data described in Fig. 17.6. The Q dependence of the curves is calculated from the QCD evolution equations.

tend to zero, and the right-hand side of (18.189) becomes a fixed number times the energy-momentum tensor. Then, using (18.191), we can evaluate the $n = 2$ moment sum rule completely:

$$\int_0^1 dx\, x f_f^+(x) \to \frac{1}{16/3 + n_f}. \tag{18.193}$$

In this extreme limit, we find that each quark flavor carries the same fixed fraction of the energy-momentum of the proton. By (18.192), the remainder is carried by the gluons. To illustrate, in a theory with $n_f = 4$, each quark flavor carries 3/28 of the total momentum of the proton, and the gluons carry the remaining 4/7. Figure 18.16 shows how slowly these asymptotic results are approached starting from realistic parton distributions.

Relation to the Altarelli-Parisi Equations

The operator mixing analysis just described gives predictions for the moments of parton distributions which imply that these integrals are Q^2 dependent. Of the various moment integrals that do not involve operator mixing, only the $n = 1$ integrals which give the flavor quantum numbers of the proton are constant as a function of Q^2. The rest decrease as powers of $\log Q^2$. Similarly, one linear combination of the matrix elements of $n = 2$ twist-2 operators remains constant with Q^2. This relation is expression by the sum rule (18.192). To write this relation more clearly, let us introduce the parton distribution of gluons as a smooth function satisfying the relations

$$\int_0^1 dx\, x^{n-1} f_g(x, Q^2) = A_g^{(n)}(Q^2). \tag{18.194}$$

Then (18.192) becomes just the total momentum sum rule for parton distributions (17.39):

$$\int_0^1 dx\, x\Big[\sum_f f_f^+(x) + f_g(x)\Big] = 1. \tag{18.195}$$

It is not difficult to verify that, for $n > 2$, all of the eigenvalues of the matrix a_{ij}^n of anomalous dimension coefficients are negative. Thus, all of the higher moment sum rules decrease, subject to the flavor charge and momentum conservation laws. In other words, the operator renormalization analysis predicts that parton distributions shift down to smaller values of x as $\log Q^2$ increases. It is pleasing that this is the same conclusion that we reached in Section 17.5, where we derived the Altarelli-Parisi equations to describe this evolution of the parton distributions.

Given that the operator analysis and the Altarelli-Parisi equations imply the same qualitative behavior for the parton distributions, how do these analyses compare quantitatively? To compare them directly, we should work out what predictions the Altarelli-Parisi equations make for the moments of the parton distribution functions. Let us begin with the simpler case of $f_f^-(x) = f_f(x) - f_{\bar{f}}(x)$.

To find the Altarelli-Parisi equation for this quantity, subtract the last two equations of (17.128). The term involving the gluon distribution cancels, and we find

$$\frac{d}{d\log Q^2} f_f^-(x) = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dz}{z} P_{q\leftarrow q}(z) f_f^-\big(\frac{x}{z}\big). \tag{18.196}$$

Now define

$$M_{fn}^- = \int_0^1 dx\, x^{n-1} f_f^-(x). \tag{18.197}$$

This quantity obeys the differential equation

$$\frac{d}{d\log Q^2} M_{fn}^- = \frac{\alpha_s(Q^2)}{2\pi} \int_0^1 dx\, x^{n-1} \int_x^1 \frac{dz}{z} P_{q\leftarrow q}(z) f_f^-\big(\frac{x}{z}\big). \tag{18.198}$$

Interchange the order of integration on the right-hand side, and change variables to $y = x/z$:

$$\begin{aligned}\int_0^1 dx\, x^{n-1} \int_x^1 \frac{dz}{z} &= \int_0^1 \frac{dz}{z} \int_0^z dx\, x^{n-1} \\ &= \int_0^1 \frac{dz}{z} \int_0^1 dy\, y^{n-1} z^n \\ &= \int_0^1 dz\, z^{n-1} \int_0^1 dy\, y^{n-1}. \end{aligned} \tag{18.199}$$

Then the right-hand side of the differential equation neatly factorizes:

$$\frac{d}{d\log Q^2} M_{nf}^{-} = \frac{\alpha_s(Q^2)}{2\pi} \left[\int_0^1 dz\, z^{n-1} P_{q\leftarrow q}(z) \right] \cdot \int_0^1 dy\, y^{n-1} f_f^{-}(y); \tag{18.200}$$

the last factor is again M_{fn}^{-}. The coefficient in this relation is the nth moment of the splitting function $P_{q\leftarrow q}(z)$. We can compute this from the explicit form of this function given in (17.129):

$$\int_0^1 dz\, z^{n-1} P_{q\leftarrow q}(z) = \int_0^1 dz\, z^{n-1} \frac{4}{3} \left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z) \right]. \tag{18.201}$$

The integral over the distribution is done by using the definition (17.105):

$$\begin{aligned}\int_0^1 dz\, z^{n-1} \frac{1}{(1-z)_+} &= \int_0^1 dz\, \frac{z^{n-1}-1}{(1-z)} \\ &= \int_0^1 dz (-1 - z - \cdots - z^{n-2}) \\ &= -\sum_1^{n-1} \frac{1}{j}. \end{aligned} \tag{18.202}$$

Then

$$\begin{aligned}\int_0^1 dz\, z^{n-1} P_{q\leftarrow q}(z) &= -\frac{4}{3} \left[\sum_1^{n-1} \frac{1}{j} + \sum_1^{n+1} \frac{1}{j} - \frac{3}{2} \right] \\ &= -\frac{2}{3} \left[1 + 4 \sum_2^{n} \frac{1}{j} - \frac{2}{n(n+1)} \right]. \end{aligned} \tag{18.203}$$

Remarkably, this is just $a_f^n/4$, as the anomalous dimension coefficient is given in (18.172) or (18.181). Thus, according to the Altarelli-Parisi equations, the nth moment of $f_f^-(x)$ obeys

$$\frac{d}{d\log Q^2} M_{fn}^- = \frac{\alpha_s(Q^2)}{8\pi} a_f^n \cdot M_{fn}^-. \tag{18.204}$$

To integrate this equation, we need the explicit form of $\alpha_s(Q^2)$. Inserting expression (17.17), we find

$$\frac{d}{d\log Q^2} M_{fn}^- = \frac{a_f^n}{2b_0} \frac{1}{\log(Q^2/\Lambda^2)} M_{fn}^-. \tag{18.205}$$

The solution of this equation, derived from the Altarelli-Parisi equations, is precisely the function (18.160) that we derived from the operator analysis of the moment sum rules for f_f^-.

It is not difficult to check that this conclusion is more general. By taking the nth moment of the full Altarelli-Parisi equations (17.128), we convert these equations to a set of ordinary differential equations for the moments. The linear combination of quark distribution functions

$$\sum_f \left(f_f(x) + f_{\bar{f}}(x)\right) \tag{18.206}$$

couples to the gluon distribution and leads to a 2×2 set of equations. All orthogonal linear combinations separate from the gluon distribution and thus have moments that obey equations identical to (18.205). To analyze the coupled equations, define

$$M_n^+ = \int_0^1 dx\, x^{n-1} \sum_f \left(f_f(x) + f_{\bar{f}}(x)\right); \qquad M_{gn} = \int_0^1 dx\, x^{n-1} f_g(x). \tag{18.207}$$

Then one can show, by the manipulations that led to (18.205), that the Altarelli-Parisi equations predict for these moments the set of coupled equations

$$\begin{aligned} \frac{d}{d\log Q^2} M_n^+ &= \frac{1}{2b_0} \frac{1}{\log(Q^2/\Lambda^2)} \left[a_{ff}^n M_n^+ + a_{fg}^n M_{gn}\right], \\ \frac{d}{d\log Q^2} M_{gn} &= \frac{1}{2b_0} \frac{1}{\log(Q^2/\Lambda^2)} \left[n_f a_{gf}^n M_n^+ + a_{gg}^n M_{gn}\right], \end{aligned} \tag{18.208}$$

where the coefficients a_{ij}^n are proportional to the nth moments of the splitting functions given in (17.129) and (17.130). In all cases, one can see that these coefficients agree precisely with the corresponding coefficients in (18.181). Thus, the solution of these equations gives the same Q^2 dependence for the moments of parton distribution functions that we found from the operator analysis.

Remarkably, the analysis of parton splitting functions given in Chapter 17 and the analysis of operator renormalization factors given above have turned out to be two views of the same basic phenomenon. Both sets of equations express the manner in which the constituents of hadrons in QCD are resolved, layer by layer, by hard-scattering processes at successively higher values of the momentum transfer. Our understanding that a quark, when studied on a fine scale, is resolved into a set of quarks, antiquarks, and gluons indicates that we have gone far beyond the simple notions of one-particle relativistic mechanics. Our two complementary derivations of this idea reinforce its fundamental character as a prediction of quantum field theory. It is especially pleasing that, as we saw at the end of Section 17.5, Nature apparently accepts this prediction and makes this consequence of quantum field theory an essential part of the structure of hadrons.

Problems

18.1 Matrix element for proton decay. Some advanced theories of particle interactions include heavy particles X whose couplings violate the conservation of baryon number. Integrating out these particles produces an effective interaction that allows the proton to decay to a positron and a photon or a pion. This effective interaction is most easily written using the definite-helicity components of the quark and electron fields: If u_L, d_L, u_R, e_R are two-component spinors, then this effective interaction is

$$\Delta\mathcal{L} = \frac{2}{m_X^2}\epsilon_{abc}\epsilon^{\alpha\beta}\epsilon^{\gamma\delta}e_{R\alpha}u_{Ra\beta}u_{Lb\gamma}d_{Lc\delta}.$$

A typical value for the mass of the X boson is $m_X = 10^{16}$ GeV.

(a) Estimate, in order of magnitude, the value of the proton lifetime if the proton is allowed to decay through this interaction.

(b) Show that the three-quark operator in $\Delta\mathcal{L}$ has an anomalous dimension

$$\gamma = -4\frac{g^2}{(4\pi)^2}.$$

Estimate the enhancement of the proton decay rate due to the leading QCD corrections.

18.2 Parity-violating deep inelastic form factor. In this problem, we first motivate the presence of additional deep inelastic form factors that are proportional to *differences* of quark and antiquark distribution functions. Then we define these functions formally and work out their properties.

(a) Analyze neutrino-proton scattering following the method used at the beginning of Section 18.5. Define

$$J_+^\mu = \bar{u}\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)d, \qquad J_-^\mu = \bar{d}\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)u.$$

Let

$$W^{\mu\nu(\nu)} = 2i\int d^4x\, e^{iq\cdot x}\,\langle P|\,T\{J_-^\mu(x)J_+^\nu(0)\}\,|P\rangle\,,$$

averaged over the proton spin. Show that the cross section for deep inelastic neutrino scattering can be computed from $W^{\mu\nu(\nu)}$ according to

$$\frac{d^2\sigma}{dxdy}(\nu p \to \mu^- X) = \frac{G_F^2 y}{2\pi^2} \cdot \text{Im}\big[(k_\mu k'_\nu + k_\nu k'_\mu - g_{\mu\nu} k\cdot k' - i\epsilon_{\mu\nu}{}^{\alpha\beta} k'_\alpha k_\beta) W^{\mu\nu(\nu)}(P,q)\big].$$

(b) Show that any term in $W^{\mu\nu(\nu)}$ proportional to q^μ or q^ν gives zero when contracted with the lepton momentum tensor in the formula above. Thus we can expand $W^{\mu\nu(\nu)}$ with three scalar form factors,

$$W^{\mu\nu(\nu)} = -g^{\mu\nu} W_1^{(\nu)} + P^\mu P^\nu W_2^{(\nu)} + i\epsilon^{\mu\nu\lambda\sigma} P_\lambda q_\sigma W_3^{(\nu)} + \cdots,$$

where the additional terms do not contribute to the deep inelastic cross section. Find the formula for the deep inelastic cross section in terms of the imaginary parts of $W_1^{(\nu)}$, $W_2^{(\nu)}$, and $W_3^{(\nu)}$.

(c) Evaluate the form factors $W_i^{(\nu)}$ in the parton model, and show that

$$\begin{aligned}
\text{Im}\, W_1^{(\nu)} &= \pi(f_d(x) + f_{\bar{u}}(x)), \\
\text{Im}\, W_2^{(\nu)} &= \frac{4\pi}{ys} x(f_d(x) + f_{\bar{u}}(x)), \\
\text{Im}\, W_3^{(\nu)} &= \frac{2\pi}{ys}(f_d(x) - f_{\bar{u}}(x)).
\end{aligned}$$

Insert these expressions into the formula derived in part (b) and show that the result reproduces the first line of Eq. (17.35).

(d) This analysis motivates the following definition: For a single quark flavor f, let

$$J^\mu_{fL} = \bar{f}\gamma^\mu\left(\frac{1-\gamma^5}{2}\right) f.$$

Define

$$W^{\mu\nu}_{fL} = 2i\int d^4x\, e^{iq\cdot x}\, \langle P|\, T\{J^\mu_{fL}(x) J^\nu_{fL}(0)\}\, |P\rangle\,.$$

Decompose this tensor according to

$$W^{\mu\nu}_{fL} = -g^{\mu\nu} W_{1fL} + P^\mu P^\nu W_{2fL} + i\epsilon^{\mu\nu\lambda\sigma} P_\lambda q_\sigma W_{3fL} + \cdots,$$

where the remaining terms are proportional to q^μ or q^ν. Evaluate the W_{iL} in the parton model. Show that the quantities W_{1fL} and W_{2fL} reproduce the expressions for W_{1f} and W_{2f} given by Eqs. (18.120) and (18.144), and that W_{3fL} is given by

$$\text{Im}\, W_{3fL} = \frac{2\pi}{ys}(f_f(x) - f_{\bar{f}}(x)).$$

(e) Compute the operator product of the currents in the expression for $W^{\mu\nu}_{fL}$, and write the terms in this product that involve twist-2 operators. Show that the expressions for W_{1fL} and W_{2fL} that follow from this analysis reproduce the expressions for W_{1f} and W_{2f} given by Eqs. (18.144) and (18.145). Find the corresponding expression for W_{3fL}.

(f) Define the parton distribution f_f^- by the relation

$$f_f^-(x, Q^2) = \frac{ys}{2\pi} \operatorname{Im} W_{3fL}(x, Q^2).$$

Show that, by virtue of this definition, the distribution function f_f^- satisfies the sum rule (18.155) for odd n.

18.3 Anomalous dimensions of gluon twist-2 operators.

(a) Compute the divergent parts of the diagrams in Fig. 18.14, and use these to derive the second line of Eq. (18.181). Notice that this result holds only for n even. Show that the two diagrams cancel for n odd.

(b) Compute the divergent parts of the diagrams in Fig. 18.5, and use these to derive the third and fourth lines of Eq. (18.181).

18.4 Deep inelastic scattering from a photon. Consider the problem of deep-inelastic scattering of an electron from a photon. This process can actually be measured by analyzing the reaction $e^+e^- \to e^+e^- + X$ in the regime where the positron goes forward, with emission of a collinear photon, which then has a hard reaction with the electron. Let us analyze this process to leading order in QED and to leading-log order in QCD. To predict the photon structure functions, it is reasonable to integrate the renormalization group equations with the initial condition that the parton distribution for photons in the photon is $\delta(x-1)$ at $Q^2 = (\frac{1}{2}\ \text{GeV})^2$. Take $\Lambda = 150$ MeV. Assume for simplicity that there are four flavors of quarks, u, d, c, and s, with charges 2/3, $-1/3$, 2/3, $-1/3$, respectively, and that it is always possible to ignore the masses of these quarks.

(a) Use the Altarelli-Parisi equations to compute the parton distributions for quarks and antiquarks in the photon, to leading order in QED and to zeroth order in QCD. Compute also the probability that the photon remains a photon as a function of Q^2.

(b) Formulate the problem of computing the moments of W_2 for the photon as a problem in operator mixing. Compute the relevant anomalous dimension matrix γ. You should be able to assemble this matrix from familiar ingredients without doing further Feynman diagram computations.

(c) Compute the $n = 2$ moments of the photon structure functions as a function of Q^2.

(d) Describe qualitatively the evolution of the photon structure function as a function of x and Q^2.

Chapter 19

Perturbation Theory Anomalies

In many examples, we have seen that loop corrections can have an important effect on the predictions of quantum field theory. We have studied examples in which the relative importance of operators is shifted by radiative corrections, and in which the form of the interactions they mediate is altered. However, in specific circumstances, radiative corrections can have an even more significant effect: They can destroy symmetries of the classical equations of motion.

The most important effect of this type involves the chiral symmetries of theories with massless fermions. In Section 3.4, we saw that the massless Dirac Lagrangian has an enhanced symmetry associated with the separate number conservation of left- and right-handed fermions. This symmetry is generated by the axial vector current $j^{\mu 5} = \bar{\psi}\gamma^\mu\gamma^5\psi$. Classically,

$$\partial_\mu j^{\mu 5} = 0 \tag{19.1}$$

for zero-mass fermions. This equation of motion is true not only in free fermion theory but also, as a classical field equation, in massless QED and QCD.

However, in this chapter, we will see that the true picture is not so simple. We will show that, in gauge theories, the conservation of the axial vector current is actually incompatible with gauge invariance, and that radiative corrections in gauge theories supply a nonzero operator that appears on the right-hand side of Eq. (19.1). This new conservation equation for the axial current has a number of remarkable consequences, which we will discuss in Sections 19.3 and 19.4.

19.1 The Axial Current in Two Dimensions

Eventually, we will want to analyze the current conservation equation for the axial current in massless QCD. However, this discussion will involve some technical complication, so we will first study the physics that violates axial current conservation in a context in which the calculations are relatively simple. A particularly simple model problem is that of two-dimensional massless QED.

The Lagrangian of two-dimensional QED is

$$\mathcal{L} = \bar{\psi}(i\not{D})\psi - \tfrac{1}{4}(F_{\mu\nu})^2, \tag{19.2}$$

with $\mu, \nu = 0, 1$ and $D_\mu = \partial_\mu + ieA_\mu$. The Dirac matrices must be chosen to satisfy the Dirac algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{19.3}$$

In two dimensions, this set of relations can be represented by 2×2 matrices; we choose

$$\gamma^0 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{19.4}$$

The Dirac spinors will be two-component fields.

The product of the Dirac matrices, which anticommutes with each of the γ^μ, is

$$\gamma^5 = \gamma^0\gamma^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{19.5}$$

Then, just as in four dimensions, there are two possible currents,

$$j^\mu = \bar{\psi}\gamma^\mu\psi, \qquad j^{\mu 5} = \bar{\psi}\gamma^\mu\gamma^5\psi, \tag{19.6}$$

and both are conserved if there is no mass term in the Lagrangian.

To make the conservation laws quite explicit, we label the components of the fermion field ψ in this spinor basis as

$$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}. \tag{19.7}$$

The subscript indicates the γ^5 eigenvalue. Then, using the explicit representations (19.4) and (19.7), we can rewrite the fermionic part of (19.2) as

$$\mathcal{L} = \psi_+^\dagger i(D_0 + D_1)\psi_+ + \psi_-^\dagger i(D_0 - D_1)\psi_-. \tag{19.8}$$

In the free theory, the field equation of ψ_+ would be

$$i(\partial_0 + \partial_1)\psi_+ = 0; \tag{19.9}$$

the solutions to this equation are waves that move to the right in the one-dimensional space at the speed of light. We will thus refer to the particles associated with ψ_+ as *right-moving* fermions. The quanta associated with ψ_- are, similarly, *left-moving*. This distinction is analogous to the distinction between left- and right-handed particles which gives the physical interpretation of γ^5 in four dimensions. Since the Lagrangian (19.8) contains no terms that mix left- and right-moving fields, it seems obvious that the number currents for these fields are separately conserved. Thus,

$$\partial_\mu\Big(\bar{\psi}\gamma^\mu\big(\tfrac{1-\gamma^5}{2}\big)\psi\Big) = 0, \qquad \partial_\mu\Big(\bar{\psi}\gamma^\mu\big(\tfrac{1+\gamma^5}{2}\big)\psi\Big) = 0. \tag{19.10}$$

It is a curious property of two-dimensional spacetime that the vector and axial vector fermionic currents are not independent of each other. Let $\epsilon^{\mu\nu}$ be the totally antisymmetric symbol in two dimensions, with $\epsilon^{01} = +1$. Then the two-dimensional Dirac matrices obey the identity

$$\gamma^\mu \gamma^5 = -\epsilon^{\mu\nu} \gamma_\nu. \tag{19.11}$$

The currents $j^{\mu 5}$ and j^μ have the same relation. Thus we can study the properties of the axial vector current by using results that we have already derived for the vector current.

Vacuum Polarization Diagrams

In Section 7.5, we computed the lowest-order vacuum polarization of QED in dimensional regularization. In the limit of zero mass, we found, in Eq. (7.90),

$$i\Pi^{\mu\nu}(q) = -i(q^2 g^{\mu\nu} - q^\mu q^\nu)\frac{2e^2}{(4\pi)^{d/2}} \operatorname{tr}[1] \int_0^1 dx\, x(1-x) \frac{\Gamma(2-\frac{d}{2})}{(-x(1-x)q^2)^{2-d/2}}, \tag{19.12}$$

where $\operatorname{tr}[1] = 4$ gives the convention for tracing over Dirac matrices given in Eq. (7.88). If we set $\operatorname{tr}[1] = 2$ to be consistent with (19.4) and then set $d = 2$ in (19.12), we find the finite and well-defined result

$$\begin{aligned} i\Pi^{\mu\nu}(q) &= i(q^2 g^{\mu\nu} - q^\mu q^\nu)\frac{2e^2}{4\pi} \cdot 2 \cdot \frac{1}{q^2} \\ &= i\Big(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\Big)\frac{e^2}{\pi}. \end{aligned} \tag{19.13}$$

Notice that this expression has the structure of a photon mass term; the photon receives the mass

$$m_\gamma^2 = \frac{e^2}{\pi}. \tag{19.14}$$

Schwinger showed that this result is exact, and that the photon of two-dimensional QED is a free massive boson.* In the discussion below Eq. (7.72), we pointed out that it is not possible for a vacuum polarization amplitude consistent with the Ward identity to generate a mass for the photon unless it also contains a pole at $q^2 = 0$. In two dimensions, such a pole can arise from the infrared behavior of the fermion-antifermion intermediate state, and we see this behavior explicitly in (19.13).

Once we have an explicit expression for the vacuum polarization, we can find the expectation value of the current induced by a background electromagnetic field. This quantity is generated by the diagram of Fig. 19.1, which gives

$$\int d^2x\, e^{iq\cdot x} \langle j^\mu(x)\rangle = \frac{i}{e}\big(i\Pi^{\mu\nu}(q)\big)A_\nu(q) = -\Big(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\Big) \cdot \frac{e}{\pi} A_\nu(q), \tag{19.15}$$

*J. Schwinger, *Phys. Rev.* **128**, 2425 (1962).

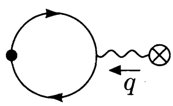


Figure 19.1. Computation of $\langle j^\mu \rangle$ in a background electromagnetic field.

where $A_\nu(q)$ is the Fourier transform of the background field. This quantity manifestly satisfies the current conservation relation $q_\mu \langle j^\mu(q) \rangle = 0$.

The identity (19.11) between the vector and axial vector currents allows us to derive from (19.15) the corresponding expectation value of $j^{\mu 5}$. We find

$$\begin{aligned} \langle j^{\mu 5}(q) \rangle &= -\epsilon^{\mu\nu} \langle j_\nu(q) \rangle \\ &= \epsilon^{\mu\nu} \frac{e}{\pi} \Big(A_\nu(q) - \frac{q_\nu q^\lambda}{q^2} A_\lambda(q) \Big). \end{aligned} \tag{19.16}$$

If the axial vector current were conserved, this object would satisfy the Ward identity. Instead, we find

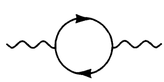
$$q_\mu \langle j^{\mu 5}(q) \rangle = \frac{e}{\pi} \epsilon^{\mu\nu} q_\mu A_\nu(q). \tag{19.17}$$

This is the Fourier transform of the field equation

$$\partial_\mu j^{\mu 5} = \frac{e}{2\pi} \epsilon^{\mu\nu} F_{\mu\nu}. \tag{19.18}$$

Apparently, the axial vector current is not conserved in the presence of electromagnetic fields, as the result of an anomalous behavior of its vacuum polarization diagram.

How could this happen? The Feynman diagrams formally satisfy the Ward identity both for the vector and for the axial vector current. The problem must come in the regularization of the vacuum polarization diagram. By dimensional analysis, we know that this diagram has the form



$$= ie^2 \Big(A g^{\mu\nu} - B \frac{q^\mu q^\nu}{q^2} \Big). \tag{19.19}$$

The coefficient B is a finite integral, and is, in any event, unambiguously determined by the low-energy structure of the theory since it is the residue of the pole in q^2. However, the integral A is logarithmically divergent, so its value depends on the regularization. Dimensional regularization automatically subtracts this integral to set $A = B$; then the vector current Ward identity is satisfied. But then we are led directly to (19.17). We could, alternatively, regularize the integral A so that $A = 0$. Working through the steps of the previous paragraph with this modification, we now find $q_\mu \langle j^{\mu 5}(q) \rangle = 0$, but

$$q_\mu \langle j^\mu(q) \rangle = \frac{e}{\pi} q^\nu A_\nu(q). \tag{19.20}$$

Though the result (19.17) is unpleasant, the result (19.20) would be a complete disaster, since it depends on the unphysical gauge degrees of freedom of the vector potential. We conclude that it is not possible to regularize two-dimensional QED so that, simultaneously, the theory is gauge invariant and the axial vector current is conserved. The price of requiring gauge invariance is the anomalous nonconservation of the axial current shown in (19.18).

The Axial Vector Current Operator Equation

To understand what happened to the axial current from another viewpoint, we now study the operator equation for the divergence of $j^{\mu 5}$. Varying the Lagrangian (19.2), we find the following equations of motion for the fermion fields:

$$\partial\!\!\!/ \psi = -ie A\!\!\!/ \psi, \qquad \partial_\mu \bar{\psi}\gamma^\mu = ie\bar{\psi}A\!\!\!/. \tag{19.21}$$

By using these equations of motion in the most straightforward way, it is easy to conclude that $\partial_\mu j^{\mu 5} = 0$. However, a closer look at these manipulations reveals some subtleties, which alter the final conclusion.

The axial vector current is a composite operator built out of fermion fields. In the previous chapter we saw that products of local operators are often singular, so we will define the current by placing the two fermion fields at distinct points separated by a distance ϵ and then carefully taking the limit as the two fields approach each other. Explicitly, we define

$$j^{\mu 5} = \underset{\epsilon\to 0}{\text{symm lim}} \left\{ \bar{\psi}(x+\tfrac{\epsilon}{2})\gamma^\mu\gamma^5 \exp\Big[-ie\int_{x-\epsilon/2}^{x+\epsilon/2} dz\cdot A(z)\Big]\psi(x-\tfrac{\epsilon}{2})\right\}. \tag{19.22}$$

Notice that, because we have placed ψ and $\bar{\psi}$ at different points, we must introduce a Wilson line (15.53) in order that the operator be locally gauge invariant. To give $j^{\mu 5}$ the correct transformation properties under Lorentz transformations, the limit $\epsilon \to 0$ should be taken symmetrically,

$$\underset{\epsilon\to 0}{\text{symm lim}} \left\{\frac{\epsilon^\mu}{\epsilon^2}\right\} = 0, \qquad \underset{\epsilon\to 0}{\text{symm lim}} \left\{\frac{\epsilon^\mu\epsilon^\nu}{\epsilon^2}\right\} = \frac{1}{d}g^{\mu\nu}, \tag{19.23}$$

with $d = 2$ in this case.

We now compute the divergence of the axial current defined as in (19.22):

$$\begin{aligned}\partial_\mu j^{\mu 5} = \underset{\epsilon\to 0}{\text{symm lim}} \Big\{ &(\partial_\mu\bar{\psi}(x+\tfrac{\epsilon}{2}))\gamma^\mu\gamma^5 \exp\Big[-ie\int_{x-\epsilon/2}^{x+\epsilon/2} dz\cdot A(z)\Big]\psi(x-\tfrac{\epsilon}{2}) \\ &+ \bar{\psi}(x+\tfrac{\epsilon}{2})\gamma^\mu\gamma^5 \exp\Big[-ie\int_{x-\epsilon/2}^{x+\epsilon/2} dz\cdot A(z)\Big](\partial_\mu\psi(x-\tfrac{\epsilon}{2})) \\ &+ \bar{\psi}(x+\tfrac{\epsilon}{2})\gamma^\mu\gamma^5\big[-ie\epsilon^\nu\partial_\mu A_\nu(x)\big]\psi(x-\tfrac{\epsilon}{2})\Big\}.\end{aligned} \tag{19.24}$$

Using the equations of motion (19.21), and keeping terms up to order ϵ, we can reduce this to

$$\begin{aligned}\partial_\mu j^{\mu 5} &= \operatorname{symm}\lim_{\epsilon\to 0}\Big\{\bar\psi(x+\tfrac{\epsilon}{2})\big[ie\not{A}(x+\tfrac{\epsilon}{2}) - ie\not{A}(x-\tfrac{\epsilon}{2}) \\ &\qquad\qquad - ie\epsilon^\nu\gamma^\mu\partial_\mu A_\nu(x)\big]\gamma^5\psi(x-\tfrac{\epsilon}{2})\Big\} \\ &= \operatorname{symm}\lim_{\epsilon\to 0}\Big\{\bar\psi(x+\tfrac{\epsilon}{2})\big[-ie\gamma^\mu\epsilon^\nu(\partial_\mu A_\nu - \partial_\nu A_\mu)\big]\gamma^5\psi(x-\tfrac{\epsilon}{2})\Big\}.\end{aligned} \tag{19.25}$$

Expression (19.25) seems to vanish in the limit $\epsilon\to 0$. However, we must take account of the fact that the product of the fermion operators is singular. In two dimensions, the contraction of fermion fields is

$$\begin{aligned}\overbrace{\psi(y)\bar\psi(z)} &= \int\frac{d^2k}{(2\pi)^2}e^{-ik\cdot(y-z)}\frac{i\not{k}}{k^2} \\ &= -\not{\partial}\Big(\frac{i}{4\pi}\log(y-z)^2\Big) \\ &= \frac{-i}{2\pi}\frac{\gamma^\alpha(y-z)_\alpha}{(y-z)^2}.\end{aligned} \tag{19.26}$$

Thus,

$$\overbrace{\bar\psi(x+\tfrac{\epsilon}{2})\Gamma\psi(x-\tfrac{\epsilon}{2})} = \frac{-i}{2\pi}\operatorname{tr}\Big[\frac{\gamma^\alpha\epsilon_\alpha}{\epsilon^2}\Gamma\Big]. \tag{19.27}$$

Notice that the result (19.27) contains an extra minus sign from the interchange of fermion operators.

Because the contraction of fermion fields is singular as $\epsilon\to 0$, the terms of order ϵ in the last line of (19.25) can give a finite contribution. Taking the contraction according to (19.27), we find

$$\partial_\mu j^{\mu 5} = \operatorname{symm}\lim_{\epsilon\to 0}\Big\{\frac{-i}{2\pi}\operatorname{tr}\Big[\frac{\gamma^\alpha\epsilon_\alpha}{\epsilon^2}\gamma^\mu\gamma^5\Big]\cdot\big(-ie\epsilon^\nu F_{\mu\nu}\big)\Big\}. \tag{19.28}$$

In two dimensions, $\operatorname{tr}[\gamma^\alpha\gamma^\mu\gamma^5] = 2\epsilon^{\alpha\mu}$. Thus,

$$\partial_\mu j^{\mu 5} = \frac{e}{2\pi}\operatorname{symm}\lim_{\epsilon\to 0}\Big\{2\frac{\epsilon_\mu\epsilon^\nu}{\epsilon^2}\Big\}\epsilon^{\mu\alpha}F_{\nu\alpha}. \tag{19.29}$$

Now take the symmetric limit according to the prescription (19.23). We find precisely the anomalous nonconservation equation (19.18). In this derivation, (19.18) appears as an operator relation, rather than in a simple matrix element. Notice that, as in our first derivation of this equation, the assumption of local gauge invariance played a crucial role. If we had defined the axial vector current by reversing the sign of the Wilson line in (19.22), a prescription that would have done violence to local gauge invariance, we would have found the various contributions canceling on the right-hand side of (19.29).

An Example with Fermion Number Nonconservation

To complete our discussion of the two-dimensional axial vector current, we will show that the nonconservation equation (19.18) also has a global aspect. In free fermion theory, the integral of the axial current conservation law gives

$$\int d^2x\, \partial_\mu j^{\mu 5} = N_R - N_L = 0. \tag{19.30}$$

This relation implies that the difference in the number of right-moving and left-moving fermions cannot be changed in any possible process. Combining this with the conservation law for the vector current, we conclude that the number of each type of fermion is separately conserved. From (19.8), we might conclude that these separate conservation laws hold also in two-dimensional QED. However, we have already found that we must be careful in making statements about the axial current.

In two-dimensional QED, the conservation equation for the axial current is replaced by the anomalous nonconservation equation (19.18). If the right-hand side of this equation were the total derivative of a quantity falling off sufficiently rapidly at infinity, its integral would vanish and we would still retain the global conservation law. In fact, $\epsilon^{\mu\nu}F_{\mu\nu}$ is a total derivative:

$$\epsilon^{\mu\nu}F_{\mu\nu} = 2\partial_\mu\big(\epsilon^{\mu\nu}A_\nu\big). \tag{19.31}$$

However, it is easy to imagine examples where the integral of this quantity does not vanish, for example, a world with a constant background electric field. In such a world, the conservation law (19.30) must be violated. But how can this happen?

Let us analyze this problem by thinking about fermions in one space dimension in a background A^1 field that is constant in space and has a very slow time dependence. We will assume that the system has a finite length L, with periodic boundary conditions. Notice that the constant A^1 field cannot be removed by a gauge transformation that satisfies the periodic boundary conditions. One way to see this is to note that the system gives a nonzero value to the Wilson line

$$\exp\Big[-ie\int_0^L dx\, A_1(x)\Big], \tag{19.32}$$

which forms a gauge-invariant closed loop due to the periodic boundary conditions.

Following the derivation of the three-dimensional Hamiltonian, Eq. (3.84), we find that the Hamiltonian of this one-dimensional system is

$$H = \int dx\, \psi^\dagger\big(-i\alpha^1 D_1\big)\psi, \tag{19.33}$$

where $\alpha = \gamma^0\gamma^1 = \gamma^5$. In the components (19.7),

$$H = \int dx \Big\{ -i\psi_+^\dagger(\partial_1 - ieA^1)\psi_+ + i\psi_-^\dagger(\partial_1 - ieA^1)\psi_- \Big\}. \tag{19.34}$$

For a constant A^1 field, it is easy to diagonalize this Hamiltonian. The eigenstates of the covariant derivatives are wavefunctions

$$e^{ik_n x}, \qquad \text{with} \quad k_n = \frac{2\pi n}{L}, \; n = -\infty, \ldots, \infty, \tag{19.35}$$

to satisfy the periodic boundary conditions. Then the single-particle eigenstates of H have energies

$$\begin{aligned} \psi_+ : \qquad & E_n = +(k_n - eA^1), \\ \psi_- : \qquad & E_n = -(k_n - eA^1). \end{aligned} \tag{19.36}$$

Each type of fermion has an infinite tower of equally spaced levels. To find the ground state of H, we fill the negative energy levels and interpret holes created among these filled states as antiparticles.

Now adiabiatically change the value of A^1. The fermion energy levels slowly shift in accord with the relations (19.36). If A^1 changes by the finite amount

$$\Delta A^1 = \frac{2\pi}{eL}, \tag{19.37}$$

which brings the Wilson loop (19.32) back to its original value, the spectrum of H returns to its original form. In this process, each level of ψ_+ moves down to the next position, and each level of ψ_- moves up to the next position, as shown in Fig. 19.2. The occupation numbers of levels should be maintained in this adiabatic process. Thus, remarkably, one right-moving fermion disappears from the vacuum and one extra left-moving fermion appears. At the same time,

$$\begin{aligned} \int d^2x \left(\frac{e}{\pi}\epsilon^{\mu\nu}F_{\mu\nu}\right) &= \int dt\, dx\, \frac{e}{\pi}\partial_0 A_1 \\ &= \frac{e}{\pi}L(-\Delta A^1) \\ &= -2, \end{aligned} \tag{19.38}$$

where we have inserted (19.37) in the last line. Thus the integrated form of the anomalous nonconservation equation (19.18) is indeed satisfied:

$$N_R - N_L = \int d^2x \left(\frac{e}{2\pi}\epsilon^{\mu\nu}F_{\mu\nu}\right). \tag{19.39}$$

Even in this simple example, we see that it is not possible to escape the question of ultraviolet regularization in analyzing the chiral conservation law. Right-moving fermions are lost and left-moving fermions appear from the depths of the fermionic spectrum, $E \to -\infty$. In computing the changes in the separate fermion numbers, we have assumed that the vacuum cannot change the charge it contains at large negative energies. This prescription is gauge invariant, but it leads to the nonconservation of the axial vector current.

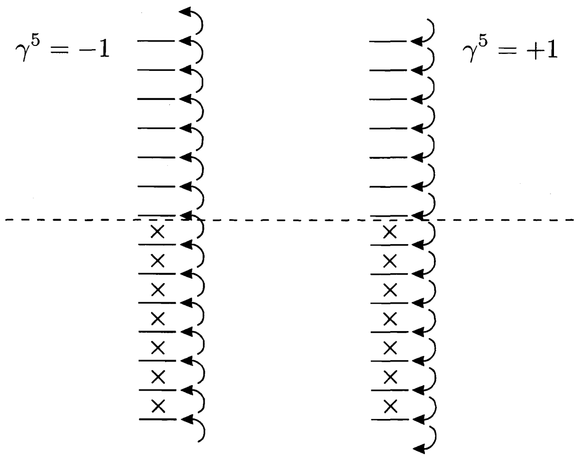


Figure 19.2. Effect on the vacuum state of the Hamiltonian H of one-dimensional QED due to an adiabatic change in the background A^1 field.

19.2 The Axial Current in Four Dimensions

All of the derivations we have just given for the two-dimensional axial current have analogues in four dimensions. In Eq. (3.40), we showed that, in the case of massless fermions, the four-dimensional Dirac equation splits neatly into separate equations for left- and right-handed fermions. If we couple the Dirac equation to a gauge field, we replace derivatives by covariant derivatives. This does not seem to affect the manifest separation between the two helicity components. Thus it seems clear that both the vector and axial vector currents should remain conserved. However, after the analysis we have just completed for the two-dimensional case, we know that we should not take these conservation laws for granted. We will now make a more careful analysis of the axial vector conservation law in four dimensions.

The Axial Vector Current Operator Equation

We begin with the case of massless four-dimensional QED. Of the three arguments that we gave in the previous section for the two-dimensional axial current conservation law, the operator derivation generalizes most easily. The fermion field equations (19.21) are identical in the four-dimensional case. We can again adopt the gauge-invariant definition of the axial vector current (19.22). When we take the divergence of this current, all of the manipulations leading to Eq. (19.25) are still correct.

From this point, we must compute the singular terms in the operator product of the two fermion fields in the limit $\epsilon \to 0$. As in two dimensions, the leading term is given by contracting the two operators using a free-field

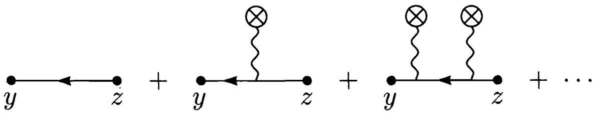


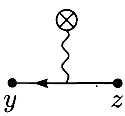
Figure 19.3. Expansion of $\psi(y)\bar{\psi}(z)$ in the presence of a background gauge field.

propagator. This contribution gives

$$\begin{aligned}\overbracket{\psi(y)\bar{\psi}}(z) &= \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot(y-z)} \frac{i\not{k}}{k^2} \\ &= -\not{\partial}\Big(\frac{i}{4\pi^2}\frac{1}{(y-z)^2}\Big) \\ &= \frac{-i}{2\pi^2}\frac{\gamma^\alpha (y-z)_\alpha}{(y-z)^4}. \end{aligned} \tag{19.40}$$

This is highly singular as $(y-z) \to 0$, but it gives zero when traced with $\gamma^\mu\gamma^5$. To find a nonzero result, we must consider terms of higher order in the expansion of the product of operators.

In a nonzero background gauge field, the contraction of fermion fields is given by the series of diagrams shown in Fig. 19.3. We have computed the leading term in this series in (19.40). The higher terms give less singular terms as $(y-z) \to 0$. The second term in the series is given by



$$= \int \frac{d^4k}{(2\pi)^4}\frac{d^4p}{(2\pi)^4} e^{-i(k+p)\cdot y} e^{ik\cdot z} \frac{i(\not{k}+\not{p})}{(k+p)^2}\big(-ie\not{A}(p)\big)\frac{i\not{k}}{k^2}. \tag{19.41}$$

This contribution leads to

$$\begin{aligned}&\big\langle \bar{\psi}(x+\tfrac{\epsilon}{2})\gamma^\mu\gamma^5\psi(x-\tfrac{\epsilon}{2})\big\rangle \\ &\quad = \int \frac{d^4k}{(2\pi)^4}\frac{d^4p}{(2\pi)^4} e^{ik\cdot\epsilon} e^{-ip\cdot x}\, \mathrm{tr}\Big[(-\gamma^\mu\gamma^5)\frac{i(\not{k}+\not{p})}{(k+p)^2}\big(-ie\not{A}(p)\big)\frac{i\not{k}}{k^2}\Big] \\ &\quad = \int \frac{d^4k}{(2\pi)^4}\frac{d^4p}{(2\pi)^4} e^{ik\cdot\epsilon} e^{-ip\cdot x}\frac{4e\epsilon^{\mu\alpha\beta\gamma}(k+p)_\alpha A_\beta(p) k_\gamma}{k^2(k+p)^2}. \end{aligned} \tag{19.42}$$

To evaluate the limit $\epsilon \to 0$, we can expand the integrand for large k. Then

$$\begin{aligned}\big\langle \bar{\psi}(x+\tfrac{\epsilon}{2})\gamma^\mu\gamma^5\psi(x-\tfrac{\epsilon}{2})\big\rangle &\sim 4e\epsilon^{\mu\alpha\beta\gamma}\int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot x} p_\alpha A_\beta(p) \int \frac{d^4k}{(2\pi)^4} e^{ik\cdot\epsilon}\frac{k_\gamma}{k^4} \\ &= 4e\epsilon^{\alpha\beta\mu\gamma}\big(\partial_\alpha A_\beta(x)\big)\frac{\partial}{\partial\epsilon^\gamma}\Big(\frac{i}{16\pi^2}\log\frac{1}{\epsilon^2}\Big) \\ &= 2e\epsilon^{\alpha\beta\mu\gamma}F_{\alpha\beta}(x)\Big(\frac{-i}{8\pi^2}\frac{\epsilon_\gamma}{\epsilon^2}\Big). \end{aligned} \tag{19.43}$$

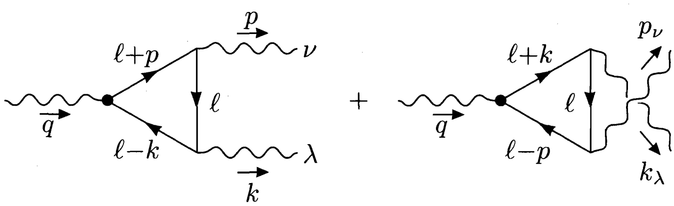


Figure 19.4. Diagrams contributing to the two-photon matrix element of the divergence of the axial vector current.

Substituting this expression into (19.25), we find

$$\partial_\mu j^{\mu 5} = \operatorname{symm}\lim_{\epsilon\to 0}\left\{\frac{e}{4\pi^2}\epsilon^{\alpha\beta\mu\gamma}F_{\alpha\beta}\Big(\frac{-i\epsilon_\gamma}{\epsilon^2}\Big)\big(-ie\epsilon^\nu F_{\mu\nu}\big)\right\}. \tag{19.44}$$

Now take the symmetric limit $\epsilon\to 0$ in four dimensions. We find

$$\partial_\mu j^{\mu 5} = -\frac{e^2}{16\pi^2}\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}. \tag{19.45}$$

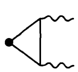
This equation, which expresses the anomalous nonconservation of the four-dimensional axial current, is known as the Adler-Bell-Jackiw anomaly. Adler and Bardeen proved that this operator relation is actually correct to all orders in QED perturbation theory and receives no further radiative corrections.†

Triangle Diagrams

We can confirm the Adler-Bell-Jackiw relation by checking, in standard perturbation theory, that the divergence of the axial vector current has a nonzero matrix element to create two photons. To do this, we must analyze the matrix element

$$\int d^4x e^{-iq\cdot x}\langle p,k|\, j^{\mu 5}(x)\,|0\rangle = (2\pi)^4\delta^{(4)}(p+k-q)\epsilon^*_\nu(p)\epsilon^*_\lambda(k)\mathcal{M}^{\mu\nu\lambda}(p,k). \tag{19.46}$$

The leading-order diagrams contributing to $\mathcal{M}^{\mu\nu\lambda}$ are shown in Fig. 19.4. The first diagram gives the contribution



$$= (-1)(-ie)^2\int\frac{d^4\ell}{(2\pi)^4}\operatorname{tr}\left[\gamma^\mu\gamma^5\frac{i(\not\ell-\not k)}{(\ell-k)^2}\gamma^\lambda\frac{i\not\ell}{\ell^2}\gamma^\nu\frac{i(\not\ell+\not p)}{(\ell+p)^2}\right], \tag{19.47}$$

and the second diagram gives an identical contribution with (p,ν) and (k,λ) interchanged.

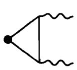
It is easy to give a formal argument that the matrix element of the divergence of the axial current vanishes at this order. Taking the divergence of the axial current in (19.46) is equivalent to dotting this quantity with iq_μ.

†S. Adler and W. A. Bardeen, *Phys. Rev.* **182**, 1517 (1969); S. Adler, in Deser, et. al. (1970).

Now we operate on the right-hand side of (19.47) as we do to prove a Ward identity. Replace

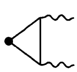
$$q_\mu \gamma^\mu \gamma^5 = (\not{\ell} + \not{p} - \not{\ell} + \not{k})\gamma^5 = (\not{\ell} + \not{p})\gamma^5 + \gamma^5(\not{\ell} - \not{k}). \tag{19.48}$$

Each momentum factor combines with the numerator adjacent to it to cancel the corresponding denominator. This brings (19.47) into the form

$iq_\mu \cdot$ 

$$= e^2 \int \frac{d^4\ell}{(2\pi)^4} \operatorname{tr}\left[\gamma^5 \frac{(\not{\ell} - \not{k})}{(\ell - k)^2} \gamma^\lambda \frac{\not{\ell}}{\ell^2} \gamma^\nu + \gamma^5 \gamma^\lambda \frac{\not{\ell}}{\ell^2} \gamma^\nu \frac{(\not{\ell} + \not{p})}{(\ell + p)^2}\right]. \tag{19.49}$$

Now pass γ^ν through γ^5 in the second term and shift the integral over the first term according to $\ell \to (\ell + k)$:

$iq_\mu \cdot$ 

$$= e^2 \int \frac{d^4\ell}{(2\pi)^4} \operatorname{tr}\left[\gamma^5 \frac{\not{\ell}}{\ell^2} \gamma^\lambda \frac{(\not{\ell} + \not{k})}{(\ell + k)^2} \gamma^\nu - \gamma^5 \frac{\not{\ell}}{\ell^2} \gamma^\nu \frac{(\not{\ell} + \not{p})}{(\ell + p)^2} \gamma^\lambda\right]. \tag{19.50}$$

This expression is manifestly antisymmetric under the interchange of (p, ν) and (k, λ), so the contribution of the second diagram in Fig. 19.4 precisely cancels (19.47).

However, because this derivation involves a shift of the integration variable, we should look closely at whether this shift is allowed by the regularization. From (19.47), we see that the integral that must be shifted is divergent. If the diagram is regulated with a simple momentum cutoff, or even with Pauli-Villars regularization, it turns out that the shift leaves over a finite, nonzero term. In Chapter 7, we encountered a similar problem in our discussion of the QED vacuum polarization diagram. We evaded the problem there by using dimensional regularization. Dimensional regularization of the diagrams of Fig. 19.4 will automatically insure the validity of the QED Ward identities for the photon emission vertices,

$$p_\nu \mathcal{M}^{\mu\nu\lambda} = k_\lambda \mathcal{M}^{\mu\nu\lambda} = 0. \tag{19.51}$$

But in the analysis of the axial vector current, even dimensional regularization has an extra subtlety, because γ^5 is an intrinsically four-dimensional object. In their original paper on dimensional regularization,‡ 't Hooft and Veltman suggested using the definition

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{19.52}$$

in d dimensions. This definition has the consequence that γ^5 anticommutes with γ^μ for $\mu = 0, 1, 2, 3$ but commutes with γ^μ for other values of μ.

In the evaluation of (19.47), the external indices and the momenta p, k, q all live in the physical four dimensions, but the loop momentum ℓ has components in all dimensions. Write

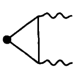
$$\ell = \ell_\parallel + \ell_\perp, \tag{19.53}$$

‡G. 't Hooft and M. J. G. Veltman, *Nucl. Phys.* **B44**, 189 (1972).

where the first term has nonzero components in dimensions $0, 1, 2, 3$ and the second term has nonzero components in the other $d-4$ dimensions. Because γ^5 commutes with the γ^μ in these extra dimensions, identity (19.48) is modified to

$$q_\mu \gamma^\mu \gamma^5 = (\not{\ell} + \not{k})\gamma^5 + \gamma^5(\not{\ell} - \not{p}) - 2\gamma^5 \not{\ell}_\perp. \tag{19.54}$$

The first two terms cancel according to the argument given above; the shift in (19.50) is justified by the dimensional regularization. However, the third term of (19.54) gives an additional contribution:

$$iq_\mu \cdot$$  $$= e^2 \int \frac{d^4\ell}{(2\pi)^4} \operatorname{tr}\left[-2\gamma^5 \not{\ell}_\perp \frac{(\not{\ell} - \not{k})}{(\ell - k)^2} \gamma^\lambda \frac{\not{\ell}}{\ell^2} \gamma^\nu \frac{(\not{\ell} + \not{p})}{(\ell + p)^2}\right]. \tag{19.55}$$

To evaluate this contribution, combine denominators in the standard way, and shift the integration variable $\ell \to \ell + P$, where $P = xk - yp$. In expanding the numerator, we must retain one factor each of γ^ν, γ^λ, $\not{p}$, and $\not{k}$ to give a nonzero trace with γ^5. This leaves over one factor of $\not{\ell}_\perp$ and one factor of $\not{\ell}$ which must also be evaluated with components in extra dimensions in order to give a nonzero integral. The factors $\not{\ell}_\perp$ anticommute with the other Dirac matrices in the problem and thus can be moved to adjacent positions. Then we must evaluate the integral

$$\int \frac{d^4\ell}{(2\pi)^4} \frac{\not{\ell}_\perp \not{\ell}_\perp}{(\ell^2 - \Delta)^3}, \tag{19.56}$$

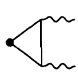
where Δ is a function of k, p, and the Feynman parameters. Using

$$(\not{\ell}_\perp)^2 = \ell_\perp^2 \to \frac{(d-4)}{d}\ell^2 \tag{19.57}$$

under the symmetrical integration, we can evaluate (19.56) as

$$\frac{i}{(4\pi)^{d/2}} \frac{(d-4)}{2} \frac{\Gamma(2-\frac{d}{2})}{\Gamma(3)\Delta^{2-d/2}} \underset{d\to 4}{\longrightarrow} \frac{-i}{2(4\pi)^2}. \tag{19.58}$$

Notice the behavior in which a logarithmically divergent integral contributes a factor $(d-4)$ in the denominator and allows an anomalous term, formally proportional to $(d-4)$, to give a finite contribution. The remainder of the algebra in the evaluation of (19.55) is straightforward. The terms involving the momentum shift P cancel, and we find

$$iq_\mu \cdot$$  $$\begin{aligned} &= e^2 \left(\frac{-i}{2(4\pi)^2}\right) \operatorname{tr}\left[2\gamma^5(-\not{k})\gamma^\lambda \not{p} \gamma^\nu\right] \\ &= \frac{e^2}{4\pi^2} \epsilon^{\alpha\lambda\beta\nu} k_\alpha p_\beta. \end{aligned} \tag{19.59}$$

This term is symmetric under the interchange of (p,ν) with (k,λ), so the second diagram of Fig. 19.4 gives an equal contribution. Thus,

$$\begin{aligned}\langle p,k|\,\partial_\mu j^{\mu 5}(0)\,|0\rangle &= -\frac{e^2}{2\pi^2}\epsilon^{\alpha\nu\beta\lambda}(-ip_\alpha)\epsilon^*_\nu(p)(-ik_\beta)\epsilon^*_\lambda(k)\\ &= -\frac{e^2}{16\pi^2}\,\langle p,k|\,\epsilon^{\alpha\nu\beta\lambda}F_{\alpha\nu}F_{\beta\lambda}(0)\,|0\rangle\,,\end{aligned} \tag{19.60}$$

as we would expect from the Adler-Bell-Jackiw anomaly equation.

Chiral Transformation of the Functional Integral

A third way of understanding the Adler-Bell-Jackiw anomaly comes from analyzing the conservation law for the axial vector current from the functional integral for the fermion field. In Section 9.6, we used the functional integral to derive the current conservation equations and the Ward identities associated with any symmetry of the Lagrangian. It is instructive to see how this argument breaks down when we apply it to the chiral symmetry of massless fermions.

We first review the standard derivation of the axial vector Ward identities following the method of Section 9.6. Starting from the fermionic functional integral

$$Z = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\exp\Big[i\int d^4x\,\bar{\psi}(i\not{D})\psi\Big], \tag{19.61}$$

make the change of variables

$$\begin{aligned}\psi(x) &\to \psi'(x) = (1+i\alpha(x)\gamma^5)\psi(x),\\ \bar{\psi}(x) &\to \bar{\psi}'(x) = \bar{\psi}(1+i\alpha(x)\gamma^5).\end{aligned} \tag{19.62}$$

Since the global chiral rotation, with constant α, is a symmetry of the Lagrangian, the only new terms in the Lagrangian that result from (19.62) contain derivatives of α. Thus,

$$\begin{aligned}\int d^4x\,\bar{\psi}'(i\not{D})\psi' &= \int d^4x\big[\bar{\psi}(i\not{D})\psi - \partial_\mu\alpha(x)\bar{\psi}\gamma^\mu\gamma^5\psi\big]\\ &= \int d^4x\big[\bar{\psi}(i\not{D})\psi + \alpha(x)\partial_\mu(\bar{\psi}\gamma^\mu\gamma^5\psi)\big].\end{aligned} \tag{19.63}$$

Then, by varying the Lagrangian with respect to $\alpha(x)$, we derive the classical conservation equation for the axial current. By carrying out a similar manipulation on the functional expression for a correlation function, as in Eq. (9.102), we would derive the associated Ward identities.

In the argument just given, we assumed that the functional measure does not change when we change variables from $\psi'(x)$ to ψ. This seems reasonable, because the relation of ψ' and ψ in (19.62) looks like a unitary transformation. However, we should examine this point more closely.* First, we must carefully

*K. Fujikawa, *Phys. Rev. Lett.* **42**, 1195 (1979); *Phys. Rev.* **D21**, 2848 (1980).

define the functional measure. To do this, expand the fermion field in a basis of eigenstates of $\not{D}$. Define right and left eigenvectors of $\not{D}$ by

$$(i\not{D})\phi_m = \lambda_m\phi_m, \qquad \hat{\phi}_m(i\not{D}) = -iD_\mu\hat{\phi}_m\gamma^\mu = \lambda_m\hat{\phi}_m. \tag{19.64}$$

For zero background A_μ field, these eigenstates are Dirac wavefunctions of definite momentum; the eigenvalues satisfy

$$\lambda_m^2 = k^2 = (k^0)^2 - (\mathbf{k})^2. \tag{19.65}$$

For a fixed A_μ field, this is also the asymptotic form of the eigenvalues for large k. These eigenfunctions give us a basis that we can use to expand ψ and $\bar{\psi}$:

$$\psi(x) = \sum_m a_m\phi_m(x), \qquad \bar{\psi}(x) = \sum_m \hat{a}_m\hat{\phi}_m(x), \tag{19.66}$$

where a_m, $\hat{a}_m$ are anticommuting coefficients multiplying the c-number eigenfunctions (19.64). The functional measure over ψ, $\bar{\psi}$ can then be defined as

$$\mathcal{D}\psi\mathcal{D}\bar{\psi} = \prod_m da_m d\hat{a}_m, \tag{19.67}$$

and the functional measure over ψ', $\bar{\psi}'$ can be defined in the same way.

If $\psi'(x) = (1 + i\alpha(x)\gamma^5)\psi(x)$, the expansion coefficients of ψ and ψ' are related by a infinitesimal linear transformation $(1 + C)$, computed as follows:

$$a'_m = \sum_n \int d^4x\phi_m^\dagger(x)(1 + i\alpha(x)\gamma^5)\phi_n(x)a_n = \sum_n(\delta_{mn} + C_{mn})a_n. \tag{19.68}$$

Then

$$\mathcal{D}\psi'\mathcal{D}\bar{\psi}' = \mathcal{J}^{-2}\cdot\mathcal{D}\psi\mathcal{D}\bar{\psi}, \tag{19.69}$$

where $\mathcal{J}$ is the Jacobian determinant of the transformation $(1 + C)$. The inverse of $\mathcal{J}$ appears in (19.69) as a result of the rule (9.63) or (9.69) for fermionic integration. To evaluate $\mathcal{J}$, we write

$$\mathcal{J} = \det\bigl(1 + C\bigr) = \exp\bigl[\mathrm{tr}\log\bigl(1 + C\bigr)\bigr] = \exp\Bigl[\sum_n C_{nn} + \cdots\Bigr], \tag{19.70}$$

and we can ignore higher order terms in the last line because C is infinitesimal. Thus,

$$\log\mathcal{J} = i\int d^4x\,\alpha(x)\sum_n \phi_n^\dagger(x)\gamma^5\phi_n(x). \tag{19.71}$$

The coefficient of $\alpha(x)$ looks like $\mathrm{tr}[\gamma^5] = 0$. However, we must regularize the sum over eigenstates n in a gauge-invariant way. The natural choice is

$$\sum_n \phi_n^\dagger(x)\gamma^5\phi_n(x) = \lim_{M\to\infty}\sum_n \phi_n^\dagger(x)\gamma^5\phi_n(x)e^{\lambda_n^2/M^2}. \tag{19.72}$$

As (19.65) indicates, the sign of λ_n^2 will be negative at large momentum after a Wick rotation; thus, the sign in the exponent of the convergence factor is given correctly. We can write (19.72) in an operator form

$$\begin{aligned}\sum_n \phi_n^\dagger(x)\gamma^5\phi_n(x) &= \lim_{M\to\infty}\sum_n \phi_n^\dagger(x)\gamma^5 e^{(i\not{D})^2/M^2}\phi_n(x) \\ &= \lim_{M\to\infty}\langle x|\,\mathrm{tr}\big[\gamma^5 e^{(i\not{D})^2/M^2}\big]\,|x\rangle\,,\end{aligned} \tag{19.73}$$

where, in the second line, we trace over Dirac indices.

To evaluate (19.73), we rewrite $(i\not{D})^2$ according to (16.107). In our present conventions, this equation reads

$$(i\not{D})^2 = -D^2 + \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}, \tag{19.74}$$

with $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu,\gamma^\nu]$. Since we are taking the limit $M \to \infty$, we can concentrate our attention on the asymptotic part of the spectrum, where the momentum k is large and we can expand in powers of the gauge field. To obtain a nonzero trace with γ^5, we must bring down four Dirac matrices from the exponent. The leading term is given by expanding the exponent to order $(\sigma\cdot F)^2$, and then ignoring the background A_μ field in all other terms. This gives

$$\begin{aligned}&\lim_{M\to\infty}\langle x|\,\mathrm{tr}\big[\gamma^5 e^{(-D^2+(e/2)\sigma\cdot F)/M^2}\big]\,|x\rangle \\ &\quad= \lim_{M\to\infty}\mathrm{tr}\Big[\gamma^5\frac{1}{2!}\Big(\frac{e}{2M^2}\sigma^{\mu\nu}F_{\mu\nu}(x)\Big)^2\Big]\,\langle x|\,e^{-\partial^2/M^2}\,|x\rangle\,.\end{aligned} \tag{19.75}$$

The matrix element in (19.75) can be evaluated by a Wick rotation:

$$\begin{aligned}\langle x|\,e^{-\partial^2/M^2}\,|x\rangle &= \lim_{x\to y}\int\frac{d^4k}{(2\pi)^4}e^{-ik\cdot(x-y)}e^{k^2/M^2} \\ &= i\int\frac{d^4k_E}{(2\pi)^4}e^{-k_E^2/M^2} \\ &= i\frac{M^4}{16\pi^2}.\end{aligned} \tag{19.76}$$

Then (19.75) reduces to

$$\begin{aligned}&\lim_{M\to\infty}\frac{-ie^2}{8\cdot 16\pi^2}M^4\,\mathrm{tr}\Big[\gamma^5\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma\frac{1}{(M^2)^2}F_{\mu\nu}F_{\lambda\sigma}(x)\Big] \\ &\quad= -\frac{e^2}{32\pi^2}\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}(x).\end{aligned} \tag{19.77}$$

Thus,

$$\mathcal{J} = \exp\Big[-i\int d^4x\,\alpha(x)\Big(\frac{e^2}{32\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}(x)\Big)\Big]. \tag{19.78}$$

In all, we find that, after the change of variables (19.62), the functional integral (19.61) takes the form

$$Z = \int \mathcal{D}\psi\mathcal{D}\bar{\psi} \exp\left[i\int d^4x\left(\bar{\psi}(i\not{D})\psi + \alpha(x)\left\{\partial_\mu j^{\mu 5} + \frac{e^2}{16\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}\right\}\right)\right]. \tag{19.79}$$

Varying the exponent with respect to $\alpha(x)$, we find precisely the Adler-Bell-Jackiw anomaly equation.

This derivation of the axial vector anomaly is especially interesting because it generalizes readily to any even dimensionality. The functional derivation always picks out for the right-hand side of the anomaly equation the pseudoscalar operator built from gauge fields that has the same dimension, d, as the divergence of the current. In two dimensions, this derivation leads immediately to (19.18). As long as d is even, we can always construct a matrix γ^5 that anticommutes with all of the Dirac matrices by taking their product. Then, the functional derivation leads straightforwardly to the result

$$\partial_\mu j^{\mu 5} = (-1)^{n+1}\frac{2e^n}{n!(4\pi)^n}\epsilon^{\mu_1\mu_2\cdots\mu_{2n}}F_{\mu_1\mu_2}\cdots F_{\mu_{2n-1}\mu_{2n}}, \tag{19.80}$$

where $n = d/2$.

At the end of the previous section, we argued that the axial vector anomaly leads to global nonconservation of fermionic charges in a two-dimensional system with a macroscopic electric field. In the same way, the four-dimensional anomaly equation leads to global nonconservation of the number of left- and right-handed fermions in background fields in which the right-hand side of (19.45) is nonzero. These are field configurations with parallel electric and magnetic fields. In Problem 19.1, we work out an example of four-dimensional massless fermions in a simple situation of this type and show that the fermion numbers are indeed violated, in a manner similar to what we saw at the end of Section 19.1, in accord with the Adler-Bell-Jackiw anomaly.

19.3 Goldstone Bosons and Chiral Symmetries in QCD

The Adler-Bell-Jackiw anomaly has a number of important implications for QCD. To describe these, we must first discuss the chiral symmetries of QCD systematically. In this discussion, we will ignore all but the lightest quarks u and d. In many analyses of the low-energy structure of the strong interactions, one also treats the s quark as light; this gives results that naturally generalize the ones we will find below.

The fermionic part of the QCD Lagrangian is

$$\mathcal{L} = \bar{u}i\not{D}u + \bar{d}i\not{D}d - m_u\bar{u}u - m_d\bar{d}d. \tag{19.81}$$

If the u and d quarks are very light, the last two terms are small and can be neglected. Let us study the implications of making this approximation. If we ignore the u and d masses, the Lagrangian (19.81) of course has isospin symmetry, the symmetry of an $SU(2)$ unitary transformation mixing the u and d fields. However, because the classical Lagrangian for massless fermions contains no coupling between left- and right-handed quarks, this Lagrangian actually is symmetric under the separate unitary transformations

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \to U_L \begin{pmatrix} u \\ d \end{pmatrix}_L, \qquad \begin{pmatrix} u \\ d \end{pmatrix}_R \to U_R \begin{pmatrix} u \\ d \end{pmatrix}_R. \tag{19.82}$$

It is useful to separate the $U(1)$ and $SU(2)$ parts of these transformations; then the symmetry group of the classical, massless QCD Lagrangian is $SU(2) \times SU(2) \times U(1) \times U(1)$. Let Q denote the quark doublet, with chiral components

$$Q_L = \left(\frac{1-\gamma^5}{2}\right)\begin{pmatrix} u \\ d \end{pmatrix}, \qquad Q_R = \left(\frac{1+\gamma^5}{2}\right)\begin{pmatrix} u \\ d \end{pmatrix}. \tag{19.83}$$

Then we can write the currents associated with these symmetries as

$$\begin{aligned} j_L^\mu &= \bar{Q}_L\gamma^\mu Q_L, & j_R^\mu &= \bar{Q}_R\gamma^\mu Q_R, \\ j_L^{\mu a} &= \bar{Q}_L\gamma^\mu \tau^a Q_L, & j_R^{\mu a} &= \bar{Q}_R\gamma^\mu \tau^a Q_R, \end{aligned} \tag{19.84}$$

where $\tau^a = \sigma^a/2$ represent the generators of $SU(2)$. The sums of left- and right-handed currents give the baryon number and isospin currents

$$j^\mu = \bar{Q}\gamma^\mu Q, \qquad j^{\mu a} = \bar{Q}\gamma^\mu \tau^a Q. \tag{19.85}$$

The corresponding symmetries are the transformations (19.82) with $U_L = U_R$. The differences of the currents (19.84) give the corresponding axial vector currents $j^{\mu 5}$, $j^{\mu 5a}$:

$$j^{\mu 5} = \bar{Q}\gamma^\mu\gamma^5 Q, \qquad j^{\mu 5a} = \bar{Q}\gamma^\mu\gamma^5\tau^a Q. \tag{19.86}$$

In the discussion to follow, we will derive conclusions about the strong interactions by assuming that the classical conservation laws for these currents are not spoiled by anomalies. We will show below that this assumption is correct for the isotriplet currents $j^{\mu 5a}$ but not for $j^{\mu 5}$.

The vector $SU(2) \times U(1)$ transformations are manifest symmetries of the strong interactions, and the associated currents lead to familiar conservation laws. What about the orthogonal, axial vector, transformations? These do not correspond to any obvious symmetry of the strong interactions. In 1960, Nambu and Jona-Lasinio hypothesized that these are accurate symmetries of the strong interactions that are spontaneously broken.† This idea has led to a correct and surprisingly detailed description of the properties of the strong interactions at low energy.

†Y. Nambu and G. Jona-Lasinio, *Phys. Rev.* **122**, 345 (1961).

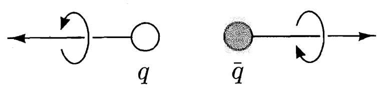


Figure 19.5. A quark-antiquark pair with zero total momentum and angular momentum.

Spontaneous Breaking of Chiral Symmetry

Before we describe the consequences of spontaneously broken chiral symmetry, let us ask why we might expect the chiral symmetries to be spontaneously broken in the first place. In the theory of superconductivity, a small electron-electron attraction leads to the appearance of a condensate of electron pairs in the ground state of a metal. In QCD, quarks and antiquarks have strong attractive interactions, and, if these quarks are massless, the energy cost of creating an extra quark-antiquark pair is small. Thus we expect that the vacuum of QCD will contain a condensate of quark-antiquark pairs. These fermion pairs must have zero total momentum and angular momentum. Thus, as Fig. 19.5 shows, they must contain net chiral charge, pairing left-handed quarks with the antiparticles of right-handed quarks. The vacuum state with a quark pair condensate is characterized by a nonzero vacuum expectation value for the scalar operator

$$\langle 0|\, \bar{Q}Q \,|0\rangle = \langle 0|\, \bar{Q}_L Q_R + \bar{Q}_R Q_L \,|0\rangle \neq 0, \tag{19.87}$$

which transforms under (19.82) with $U_L \neq U_R$. The expectation value signals that the vacuum mixes the two quark helicities. This allows the u and d quarks to acquire effective masses as they move through the vacuum. Inside quark-antiquark bound states, the u and d quarks would appear to move as if they had a sizable effective mass, even if they had zero mass in the original QCD Lagrangian.

The vacuum expectation value (19.87) signals the spontaneous breaking of the full symmetry group (19.82) down to the subgroup of vector symmetries with $U_L = U_R$. Thus there are four spontaneously broken continuous symmetries, associated with the four axial vector currents. At the end of Section 11.1, we proved Goldstone's theorem, which states that every spontaneously broken continuous symmetry of a quantum field theory leads to a massless particle with the quantum numbers of a local symmetry rotation. This means that, in QCD with massless u and d quarks, we should find four spin-zero particles with the correct quantum numbers to be created by the four axial vector currents.

The real strong interactions do not contain any massless particles, but they do contain an isospin triplet of relatively light mesons, the pions. These particles are known to have odd parity (as we expect if they are quark-antiquark bound states). Thus, they can be created by the axial isospin currents. We can parametrize the matrix element of $j^{\mu 5a}$ between the vacuum and an on-shell pion by writing

$$\langle 0|\, j^{\mu 5a}(x)\, |\pi^b(p)\rangle = -ip^\mu f_\pi \delta^{ab} e^{-ip\cdot x}, \tag{19.88}$$

where a, b are isospin indices and f_π is a constant with the dimensions of $(\text{mass})^1$. We show in Problem 19.2 that the value of f_π can be determined from the rate of π^+ decay through the weak interaction; one finds $f_\pi = 93$ MeV. For this reason, f_π is often called the *pion decay constant*. If we contract (19.88) with p_μ and use the conservation of the axial currents, we find that an on-shell pion must satisfy $p^2 = 0$, that is, it must be massless, as required by Goldstone's theorem.

If we now restore the quark mass terms in (19.81), the axial currents are no longer exactly conserved. The equation of motion of the quark field is now

$$i\not{D}Q = \mathbf{m}Q, \qquad -iD_\mu \bar{Q}\gamma^\mu = \bar{Q}\mathbf{m}, \tag{19.89}$$

where

$$\mathbf{m} = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix} \tag{19.90}$$

is the quark mass matrix. Then one can readily compute

$$\partial_\mu j^{\mu 5a} = i\bar{Q}\{\mathbf{m}, \tau^a\}\gamma^5 Q. \tag{19.91}$$

Using this equation together with (19.88), we find

$$\langle 0|\, \partial_\mu j^{\mu 5a}(0)\, |\pi^b(p)\rangle = -p^2 f_\pi \delta^{ab} = \langle 0|\, i\bar{Q}\{\mathbf{m}, \tau^a\}\gamma^5 Q\, |\pi^b(p)\rangle\,. \tag{19.92}$$

The last expression is an invariant quantity times

$$\text{tr}\big[\{\mathbf{m}, \tau^a\}\tau^b\big] = \tfrac{1}{2}\delta^{ab}(m_u + m_d). \tag{19.93}$$

Thus, the quark mass terms give the pions masses of the form

$$m_\pi^2 = (m_u + m_d)\frac{M^2}{f_\pi}. \tag{19.94}$$

The mass parameter M has been estimated to be of order 400 MeV. Thus, to give the observed pion mass of 140 MeV, one needs only $(m_u + m_d) \sim 10$ MeV. This is a small perturbation on the strong interactions.

This argument has an interesting implication for the nature of the isospin symmetry of the strong interactions. In the limit in which the u and d quarks have zero mass in the Lagrangian, these quarks acquire large, equal effective masses from the vacuum with spontaneously broken chiral symmetry. As long as the masses m_u and m_d in the Lagrangian are small compared to the effective mass, the u and d quarks will behave inside hadrons as though they are approximately degenerate. Thus the isospin symmetry of the strong interactions need have nothing to do with a fundamental symmetry linking u and d; it follows for any arbitrary relation between m_u and m_d, provided that both of these parameters are much less than 300 MeV. Similarly, the approximate $SU(3)$ symmetry of the strong interactions follows if the fundamental mass of the s quark is also small compared to the strong interaction scale. The best

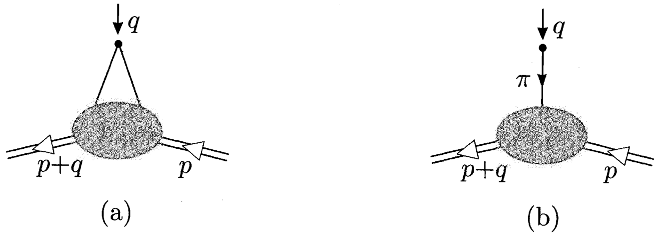


Figure 19.6. Matrix element of the axial isospin current in the nucleon: (a) kinematics of the amplitude; (b) contribution that leads to a pole in q^2.

current estimates of the mass ratios $m_u : m_d : m_s$ are in fact $1 : 2 : 40$, so that the fundamental Lagrangian of the strong interactions shows no sign of flavor symmetry among the quark masses.‡

The identification of the pions as Goldstone bosons of broken chiral symmetry has a number of implications for hadronic matrix elements. Here we will give only one example. In the following argument, we will work in the limit of exact chiral symmetry, ignoring the small corrections from the u and d masses.

The matrix element of the axial isospin current in the nucleon, a quantity that enters the theory of neutron and nuclear β decay, can be written in terms of form factors as follows:

$$\langle N|\, j^{\mu 5a}(q)\,|N\rangle = \bar{u}\Big[\gamma^\mu\gamma^5 F_1^5(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}\gamma^5 F_2^5(q^2) + q^\mu\gamma^5 F_3^5(q^2)\Big]\tau^a u. \tag{19.95}$$

The kinematics of the vertex is shown in Fig. 19.6. Notice that there is one more possible form factor than in the vector case, Eq. (6.33). The value of F_1^5 at $q^2 = 0$ is not restricted by the value of any manifestly conserved charge. Conventionally, one writes simply

$$F_1^5(0) = g_A. \tag{19.96}$$

However, we will now show that the value of this quantity can be computed.

If we ignore quark masses, the axial vector current in (19.95) is conserved, so the form factors satisfy

$$\begin{aligned} 0 &= \bar{u}(p')\Big[\not{q}\gamma^5 F_1^5(q^2) + q^2\gamma^5 F_3^5(q^2)\Big]u(p) \\ &= \bar{u}(p')\Big[(\not{p}' - \not{p})\gamma^5 F_1^5(q^2) + q^2\gamma^5 F_3^5(q^2)\Big]u(p) \\ &= \bar{u}(p')\Big[2m_N\gamma^5 F_1^5(q^2) + q^2\gamma^5 F_3^5(q^2)\Big]u(p). \end{aligned} \tag{19.97}$$

‡The determination of the fundamental quark masses is reviewed by J. Gasser and H. Leutwyler, *Phys. Repts.* **87**, 77 (1982).

Thus, we find

$$g_A = \lim_{q^2\to 0} \frac{q^2}{2m_N} F_3^5(q^2). \tag{19.98}$$

This equation implies that $g_A = 0$ unless F_3^5 contains a pole in q^2. Such a pole would imply the presence of a physical massless particle, but fortunately, there is one available—the massless pion. The process in which the current creates a pion that is then absorbed by the nucleon indeed leads to a pole in $F_3^5(q^2)$, as shown in Fig. 19.6(b).

Let us now compute this pole term and use it to determine g_A. The low-energy pion-nucleon interaction is conventionally parametrized by the Lagrangian

$$\Delta\mathcal{L} = ig_{\pi NN}\pi^a\overline{N}\gamma^5\sigma^a N. \tag{19.99}$$

The amplitude for the current $j^{\mu 5a}$ to create the pion is given by (19.88). Then the contribution of Fig. 19.6(b) to the current vertex is

$$-g_{\pi NN}\bar{u}(2\tau^a\gamma^5)u\cdot\frac{i}{q^2}\cdot(iq^\mu f_\pi). \tag{19.100}$$

Thus,

$$F_3^5(q^2) = \frac{1}{q^2}\cdot 2f_\pi g_{\pi NN}. \tag{19.101}$$

We find that g_A is given by a combination of f_π, the nucleon mass, and the pion-nucleon coupling constant:

$$g_A = \frac{f_\pi}{m_N}g_{\pi NN}. \tag{19.102}$$

This strange identity, called the *Goldberger-Treiman relation*, is satisfied experimentally to 5% accuracy.

The identification of the pion as the Goldstone boson of spontaneously broken chiral symmetry leads to numerous other predictions for current matrix elements and pion scattering amplitudes. In particular, the leading terms of the pion-pion and pion-nucleon scattering amplitudes at low energy can be computed directly in terms of f_π by arguments similar to one just given.*

Anomalies of Chiral Currents

Up to this point, we have discussed the chiral symmetries of QCD according to the classical current conservation equations. We must now ask whether these equations are affected by the Adler-Bell-Jackiw anomaly, and what the consequences of that modification are.

To begin, we study the modification of the chiral conservation laws due to the coupling of the quark currents to the gluon fields of QCD. The arguments given in the previous section go through equally well in the case of

*The detailed consequences of spontaneously broken chiral symmetry are worked out in a very clear manner in Georgi (1984).

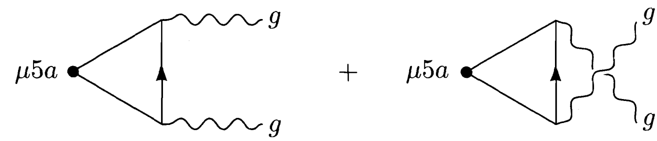


Figure 19.7. Diagrams that lead to an axial vector anomaly for a chiral current in QCD.

massless fermions coupling to a non-Abelian gauge field, so we expect that an axial vector current will receive an anomalous contribution from the diagrams shown in Fig. 19.7. The anomaly equation should be the Abelian result, supplemented by an appropriate group theory factor. In addition, since the axial current is gauge invariant, the anomaly must also be gauge invariant. That is, it must contain the full non-Abelian field strength, including its nonlinear terms. These terms are actually included in the functional derivation of the anomaly given at the end of Section 19.2.

For the axial currents of QCD, written in (19.86), we can read the group theory factors for the Adler-Bell-Jackiw anomaly from the diagrams of Fig. 19.7. For the axial isospin currents,

$$\partial_\mu j^{\mu 5a} = -\frac{g^2}{16\pi^2}\epsilon^{\alpha\beta\mu\nu}F^c_{\alpha\beta}F^d_{\mu\nu}\cdot \mathrm{tr}\big[\tau^a t^c t^d\big], \tag{19.103}$$

where $F^c_{\mu\nu}$ is a gluon field strength, τ^a is an isospin matrix, t^c is a color matrix, and the trace is taken over colors and flavors. In this case, we find

$$\mathrm{tr}\big[\tau^a t^c t^d\big] = \mathrm{tr}[\tau^a]\,\mathrm{tr}[t^c t^d] = 0, \tag{19.104}$$

since the trace of a single τ^a vanishes. Thus the conservation of the axial isospin currents is unaffected by the Adler-Bell-Jackiw anomaly of QCD. However, in the case of the isospin singlet axial current, the matrix τ^a is replaced by the matrix 1 on flavors, and we find

$$\partial_\mu j^{\mu 5} = -\frac{g^2 n_f}{32\pi^2}\epsilon^{\alpha\beta\mu\nu}F^c_{\alpha\beta}F^c_{\mu\nu}, \tag{19.105}$$

where n_f is the number of flavors; $n_f = 2$ in our current model.

Thus, the isospin singlet axial current is not in fact conserved in QCD. The divergence of this current is equal to a gluon operator with nontrivial matrix elements between hadron states. Some subtle questions remain concerning the effects of this operator. In particular, it can be shown, as we saw for the two-dimensional axial anomaly in Eq. (19.31), that the right-hand side of (19.105) is a total divergence. Nevertheless, again in accord with our experience in two dimensions, there are physically reasonable field configurations in which the four-dimensional integral of this term takes a nonzero value. This topic is discussed further at the end of Section 22.3. In any event, Eq. (19.105) indeed implies that QCD has no isosinglet axial symmetry and no associated Goldstone boson. This equation explains why the strong interactions contain no light isosinglet pseudoscalar meson with mass comparable to that of the pions.

Though the axial isospin currents have no axial anomaly from QCD interactions, they do have an anomaly associated with the coupling of quarks to electromagnetism. Again referring to the diagrams of Fig. 19.7, we see that the electromagnetic anomaly of the axial isospin currents is given by

$$\partial_\mu j^{\mu 5a} = -\frac{e^2}{16\pi^2}\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}\cdot \mathrm{tr}\big[\tau^a Q^2\big], \tag{19.106}$$

where $F_{\mu\nu}$ is the electromagnetic field strength, Q is the matrix of quark electric charges,

$$Q = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & -\frac{1}{3}\end{pmatrix}, \tag{19.107}$$

and the trace again runs over flavors and colors. Since the matrices in the trace do not depend on color, the color sum simply gives a factor of 3. The flavor trace is nonzero only for $a = 3$; in that case, the electromagnetic anomaly is

$$\partial_\mu j^{\mu 53} = -\frac{e^2}{32\pi^2}\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu}. \tag{19.108}$$

Because the current $j^{\mu 53}$ annihilates a π^0 meson, Eq. (19.108) indicates that the axial vector anomaly contributes to the matrix element for the decay $\pi^0 \to 2\gamma$. We will now show that, in fact, it gives the leading contribution to this amplitude. Again, we work in the limit of massless u and d quarks, so that the chiral symmetries are exact up to the effects of the anomaly.

Consider the matrix element of the axial current between the vacuum and a two-photon state:

$$\langle p, k|\, j^{\mu 53}(q)\,|0\rangle = \epsilon^*_\nu \epsilon^*_\lambda \mathcal{M}^{\mu\nu\lambda}(p,k). \tag{19.109}$$

This is the same matrix element (19.46) that we studied in QED perturbation theory in Section 19.2. Now, however, we will study the general properties of this matrix element by expanding it in form factors. In general, the amplitude can be decomposed by writing all possible tensor structures and applying the restrictions that follow from symmetry under the interchange of (p,ν) and (k,λ) and the QED Ward identities (19.51). This leaves three possible structures:

$$\begin{aligned}\mathcal{M}^{\mu\nu\lambda} &= q^\mu \epsilon^{\nu\lambda\alpha\beta}p_\alpha k_\beta \mathcal{M}_1 + \big(\epsilon^{\mu\nu\alpha\beta}k^\lambda - \epsilon^{\mu\lambda\alpha\beta}p^\nu\big)k_\alpha p_\beta \mathcal{M}_2 \\ &\quad + \Big[\big(\epsilon^{\mu\nu\alpha\beta}p^\lambda - \epsilon^{\mu\lambda\alpha\beta}k^\nu\big)k_\alpha p_\beta - \epsilon^{\mu\nu\lambda\sigma}(p-k)_\sigma p\cdot k\Big]\mathcal{M}_3.\end{aligned} \tag{19.110}$$

The second term satisfies (19.51) by virtue of the on-shell conditions $p^2 = k^2 = 0$.

Now contract (19.110) with (iq_μ) to take the divergence of the axial vector current. We find

$$iq_\mu \mathcal{M}^{\mu\nu\lambda} = iq^2\epsilon^{\nu\lambda\alpha\beta}p_\alpha k_\beta \mathcal{M}_1 - i\epsilon^{\mu\nu\lambda\sigma}q_\mu(p-k)_\sigma p\cdot k\mathcal{M}_3; \tag{19.111}$$

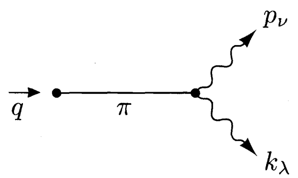


Figure 19.8. Contribution that leads to a pole in the axial vector current form factor $\mathcal{M}_1$.

the other terms automatically give zero. Using $q = p + k$, $q^2 = 2p \cdot k$, we can simplify this to

$$iq_\mu \mathcal{M}^{\mu\nu\lambda} = iq^2 \epsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta (\mathcal{M}_1 + \mathcal{M}_3). \tag{19.112}$$

The whole quantity is proportional to q^2 and apparently vanishes in the limit $q^2 \to 0$. This contrasts with the prediction of the axial vector anomaly. Taking the matrix element of the right-hand side of (19.108), we find

$$iq_\mu \mathcal{M}^{\mu\nu\lambda} = -\frac{e^2}{4\pi^2} \epsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta. \tag{19.113}$$

The conflict can be resolved if one of the form factors appearing in (19.112) contains a pole in q^2. Such a pole can arise through the process shown in Fig. 19.8, in which the current creates a π^0 meson which subsequently decays to two photons. The amplitude for the current to create the meson is given by (19.88). Let us parametrize the pion decay amplitude as

$$i\mathcal{M}(\pi^0 \to 2\gamma) = iA\epsilon_\nu^* \epsilon_\lambda^* \epsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta, \tag{19.114}$$

where A is a constant to be determined. Then the contribution of the process of Fig. 19.8 to the amplitude $\mathcal{M}^{\mu\nu\lambda}$ defined in (19.109) is

$$\big(iq^\mu f_\pi\big) \frac{i}{q^2} \big(iA\epsilon^{\nu\lambda\alpha\beta} p_\alpha k_\beta\big). \tag{19.115}$$

This is a contribution to the form factor $\mathcal{M}_1$,

$$\mathcal{M}_1 = \frac{-i}{q^2} f_\pi \cdot A, \tag{19.116}$$

plus terms regular at $q^2 = 0$. Now, by equating (19.112) to (19.113), we determine A in terms of the coefficient of the anomaly:

$$A = \frac{e^2}{4\pi^2} \frac{1}{f_\pi}. \tag{19.117}$$

From the decay matrix element (19.114), it is straightforward to work out the decay rate of the π^0. Note that, though we have worked out the decay matrix element in the limit of a massless π^0, we must supply the physically correct kinematics which depends on the π^0 mass. Including a factor 1/2 for the phase space of identical particles, we find

$$\begin{aligned}\Gamma(\pi^0 \to 2\gamma) &= \frac{1}{2m_\pi}\frac{1}{8\pi}\frac{1}{2}\sum_{\text{pols.}}\left|\mathcal{M}(\pi^0 \to 2\gamma)\right|^2 \\ &= \frac{1}{32\pi m_\pi}\cdot A^2 \cdot 2(p\cdot k)^2 \\ &= A^2 \cdot \frac{m_\pi^3}{64\pi}.\end{aligned} \tag{19.118}$$

Thus, finally,

$$\Gamma(\pi^0 \to 2\gamma) = \frac{\alpha^2}{64\pi^3}\frac{m_\pi^3}{f_\pi^2}. \tag{19.119}$$

This relation, which provides a direct measure of the coefficient of the Adler-Bell-Jackiw anomaly, is satisfied experimentally to an accuracy of a few percent.

19.4 Chiral Anomalies and Chiral Gauge Theories

Up to this point, we have coupled gauge fields to fermions in a parity-symmetric manner, replacing the derivative in the Dirac equation by a covariant derivative. This procedure couples the gauge field to the vector current of fermions. However, this procedure gives only a subset of the possible couplings of fermions to gauge bosons. In this section we will construct more general, parity-asymmetric, couplings and discuss their interplay with the axial vector anomaly.

We will focus primarily on theories of massless fermions. If the Lagrangian contains no fermion mass terms, it has no terms that mix the two helicity states of a Dirac fermion. Thus, in a theory that contains massless Dirac fermions ψ_i, we can write the kinetic energy term in the helicity basis (3.36) as

$$\mathcal{L} = \psi_{Li}^\dagger i\bar{\sigma}\cdot\partial\psi_{Li} + \psi_{Ri}^\dagger i\sigma\cdot\partial\psi_{Ri}. \tag{19.120}$$

There is no difficulty in coupling this system to a gauge field by assigning the left-handed fields ψ_{Li} to one representation of the gauge group G and assigning the right-handed fields to a different representation. For example, we might assign the left-handed fields to a representation r of G and take the right-handed fields to be invariant under G. This gives

$$\mathcal{L} = \psi_{Li}^\dagger i\bar{\sigma}\cdot D\psi_{Li} + \psi_{Ri}^\dagger i\sigma\cdot\partial\psi_{Ri}, \tag{19.121}$$

with $D_\mu = \partial_\mu - igA_\mu^a t_r^a$. In more conventional notation, (19.121) becomes

$$\mathcal{L} = \bar{\psi}i\gamma^\mu\Big(\partial_\mu - igA_\mu^a t_r^a\Big(\frac{1-\gamma^5}{2}\Big)\Big)\psi. \tag{19.122}$$

It is straightforward to verify that the classical Lagrangian (19.122) is invariant to the local gauge transformation

$$\begin{aligned} \psi &\to \Big(1 + i\alpha^a t^a \Big(\frac{1-\gamma^5}{2}\Big)\Big)\psi, \\ A^a_\mu &\to A^a_\mu + \frac{1}{g}\partial_\mu \alpha^a + f^{abc} A^b_\mu \alpha^c, \end{aligned} \qquad (19.123)$$

which generalizes (15.46). Since the right-handed fields are free fields, we can even eliminate these fields and write a gauge-invariant Lagrangian for purely left-handed fermions.

The idea of gauge fields that couple only to left-handed fermions plays a central role in the construction of a theory of weak interactions. The coupling of the W boson to quarks and leptons described in (17.31) can be derived by assigning the left-handed components of quarks and leptons to doublets of an $SU(2)$ gauge symmetry

$$Q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \qquad L_L = \begin{pmatrix} \nu \\ \ell \end{pmatrix}_L, \qquad (19.124)$$

and then identifying the W bosons as gauge fields that couple to this $SU(2)$ group. In this picture, it is the restriction of the symmetry to left-handed fields that leads to the helicity structure of the weak interaction effective Lagrangian. We will discuss a complete, explicit model of weak interactions, incorporating this idea, in the next chapter.

To work out the general properties of chirally coupled fermions, it is useful to rewrite their Lagrangian with one further transformation. Below Eq. (3.38), we noted that the quantity $\sigma^2\psi_R^*$ transforms under Lorentz transformations as a left-handed field. Thus it is useful to rewrite the right-handed components in (19.120) as new left-handed fermions, by defining

$$\psi'_{Li} = \sigma^2 \psi^*_{Ri}, \qquad \psi'^{\dagger}_{Li} = \psi^T_{Ri}\sigma^2. \qquad (19.125)$$

This transformation relabels the right-handed fermions as antifermions and calls their left-handed antiparticles a new species of left-handed fermions. By using (3.38), we can rewrite the Lagrangian for the right-handed fermions as

$$\int d^4x\, \psi^\dagger_{Ri} i\sigma\cdot\partial\psi_{Ri} = \int d^4x\, \psi'^\dagger_{Li} i\bar{\sigma}\cdot\partial\psi'_{Li}. \qquad (19.126)$$

The minus sign from fermion interchange cancels the minus sign from integration by parts. Notice that, if the fermions are coupled to gauge fields in the representation r, this manipulation changes the covariant derivative as follows:

$$\begin{aligned} \psi^\dagger_R i\sigma\cdot(\partial - igA^a t^a_r)\psi_R &= \psi'^\dagger_L i\bar{\sigma}\cdot\big(\partial + igA^a (t^a_r)^T\big)\psi'_L \\ &= \psi'^\dagger_L i\bar{\sigma}\cdot\big(\partial - igA^a t^a_{\bar{r}}\big)\psi'_L. \end{aligned} \qquad (19.127)$$

Thus the new fields ψ'_L belong to the conjugate representation to r, for which the representation matrices are given by (15.82). In this notation, QCD with

n_f flavors of massless fermions is rewritten as an $SU(3)$ gauge theory coupled to n_f massless fermions in the 3 and n_f massless fermions in the $\bar{3}$ representation of $SU(3)$. The most general gauge theory of massless fermions would simply assign left-handed fermions to an arbitrary, reducible representation R of the gauge group G. We have just seen that rewriting a system of Dirac fermions leads to $R = r \oplus \bar{r}$, a *real* representation in the sense described below (15.82). Conversely, if R is not a real representation, then the theory cannot be rewritten in terms of Dirac fermions and is intrinsically chiral.

The rewriting (19.125) transforms the mass term of the QCD Lagrangian as follows:

$$m\bar{\psi}_i\psi_i = m\big(\psi_R^\dagger\psi_L + \text{h.c.}\big) = -m\big(\psi_{Li}'^T\sigma^2\psi_{Li} + \text{h.c.}\big). \tag{19.128}$$

This has the form of the *Majorana* mass term that we encountered in Problem 3.4. The most general mass term that can be built purely from left-handed fermion fields is

$$\Delta\mathcal{L}_M = M_{ij}\psi_{Li}^T\sigma^2\psi_{Lj} + \text{h.c.} \tag{19.129}$$

The matrix M_{ij} is symmetric under $i \leftrightarrow j$, since the minus sign from the antisymmetry of σ^2 is compensated by a minus sign from fermion interchange. This mass term is gauge invariant if M_{ij} is invariant under G. For example, the mass term in (19.128) couples 3 and $\bar{3}$ indices together in an $SU(3)$ singlet combination. In general, a gauge-invariant mass term exists if the representation containing the fermions is *strictly real*, in the sense described below (15.82). In an intrinsically chiral theory, there is no possible gauge-invariant mass term. We will see in the next chapter that, in the gauge theory of the weak interactions, mass terms for the quarks and leptons are forbidden by gauge invariance. We will present a solution to this problem in Section 20.2.

At the classical level, there is no restriction on the representation R of the left-handed fermions. However, at the level of one-loop corrections, many possible choices become inconsistent due to the axial vector anomaly. In a gauge theory of left-handed massless fermions, consider computing the diagrams of Fig. 19.9, in which the external fields are non-Abelian gauge bosons and the marked vertex represents the gauge symmetry current

$$j^{\mu a} = \bar{\psi}\gamma^\mu\Big(\frac{1-\gamma^5}{2}\Big)t^a\psi. \tag{19.130}$$

The gauge boson vertices also contain factors of $(1-\gamma^5)/2$. The three projectors can be moved together into a single factor. Then, if we regularize this diagram as in Section 19.2, the term containing a γ^5 has an axial vector anomaly that leads to the relation

$$\langle p,\nu,b;k,\lambda,c|\,\partial_\mu j^{\mu a}\,|0\rangle = \frac{g^2}{8\pi^2}\epsilon^{\alpha\nu\beta\lambda}p_\alpha k_\beta\cdot\mathcal{A}^{abc}, \tag{19.131}$$

where $\mathcal{A}^{abc}$ is a trace over group matrices in the representation R:

$$\mathcal{A}^{abc} = \text{tr}\big[t^a\{t^b,t^c\}\big]. \tag{19.132}$$

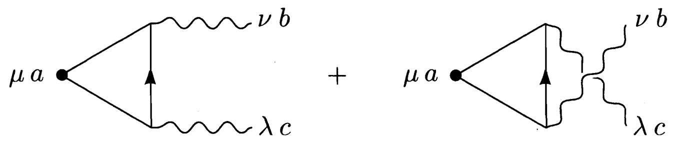


Figure 19.9. Diagrams contributing to the anomaly of a gauge symmetry current in a chiral gauge theory.

This equation implies that, unless $\mathcal{A}^{abc}$ vanishes, the current $j^{\mu a}$ is not conserved. The factor (19.132) is totally symmetric in (a, b, c), so this condition is independent of which current is treated as an external operator. As we described in Sections 19.1 and 19.2, we can change the regularization of the diagram so that the external current is conserved, but only at the price of violating the conservation of one of the other two currents in the diagram.

Since the whole construction of a theory with local gauge invariance is based on the existence of an exact global symmetry, the violation of the conservation of $j^{\mu a}$ does violence to the structure of the theory. For example, triangle diagrams of the form of Fig. 19.9 will now generate divergent gauge boson mass terms and will upset the delicate relations between three- and four-point vertices discussed in Chapter 16. These relations, following from the Ward identity, were necessary to insure the cancellation of unphysical states and the unitarity of the S-matrix. The only way to avoid this problem is to insist that $\mathcal{A}^{abc} = 0$ as a fundamental consistency condition for chiral gauge theories.† Gauge theories satisfying this condition are said to be *anomaly free*.

As an example of the application of this condition, consider the prototype weak interaction gauge theory that we presented in (19.124). If the two gauge bosons in Fig. 19.9 are $SU(2)$ gauge bosons and the current $j^{\mu a}$ is an $SU(2)$ gauge current, we would evaluate (19.132) by substituting $t^a = \tau^a = \sigma^a/2$ and using the relation $\{\sigma^b, \sigma^c\} = 2\delta^{bc}$. This gives

$$\mathcal{A}^{abc} = \frac{1}{8}\,\mathrm{tr}\big[\sigma^a \cdot 2\delta^{bc}\big] = 0, \tag{19.133}$$

so the consistency condition is satisfied. If the fermions in (19.124) also couple to electromagnetism, there is an additional consistency condition that we would find by taking the current in Fig. 19.9 to be the electromagnetic current. The factor $\mathcal{A}^{abc}$ for this case is

$$\mathrm{tr}\big[Q\{\tau^b, \tau^c\}\big], \tag{19.134}$$

where Q is the matrix of electric charges. If we simplify as in (19.133), the trace (19.134) becomes

$$\tfrac{1}{2}\,\mathrm{tr}\big[Q\big]\,\delta^{bc}. \tag{19.135}$$

†D. J. Gross and R. Jackiw, *Phys. Rev.* **D6**, 477 (1972).

This factor is proportional to the sum of the fermion electric charges, which does not vanish either for quarks or for leptons. However, if we sum over one quark doublet and one lepton doublet, with a factor 3 for colors, we find

$$\operatorname{tr}\big[Q\big] = 3\cdot(\tfrac{2}{3}-\tfrac{1}{3}) + (0-1) = 0. \tag{19.136}$$

Remarkably, the weak interaction gauge theory described by (19.124) can be consistently combined with QED only if the theory contains equal numbers of quark and lepton doublets.

We complete this section by working out more generally the condition that a chiral gauge theory be anomaly free. We will first derive some basic properties of the anomaly factor $\mathcal{A}^{abc}$ and then apply these to chiral gauge theories with simple gauge groups.

If the fermion representation R is real, R is equivalent to its conjugate reprsentation $\overline{R}$. Thus, as we described below (15.82), t^a_R is related by a unitary transformation to $t^a_{\overline{R}} = -(t^a_R)^T$. Since (19.132) is invariant to unitary transformations of the t^a, we can replace t^a_R by $t^a_{\overline{R}}$. Then

$$\begin{aligned}\mathcal{A}^{abc} &= \operatorname{tr}\big[(-t^a)^T\{(-t^b)^T,(-t^c)^T\}\big]\\ &= -\operatorname{tr}\big[\{t^c,t^b\}t^a\big]\\ &= -\mathcal{A}^{abc}.\end{aligned} \tag{19.137}$$

Thus, if R is real, the gauge theory is automatically anomaly free. As a special case, any gauge theory of Dirac fermions is anomaly free.

In more general circumstances, we can simplify the calculation of $\mathcal{A}^{abc}$ by noting that it is an invariant of the gauge group G that is totally symmetric with three indices in the adjoint representation. For some possible groups, a suitable invariant may not exist, and in those cases $\mathcal{A}^{abc}$ must vanish. For example, in $SU(2)$ the adjoint representation has spin 1. The symmetric product of two spin-1 multiplets gives spin 0 plus spin 2, with no spin-1 component. Thus, there is no symmetric tensor coupling two spin-1 indices to give a spin 1. The factor (19.132) must then vanish in any $SU(2)$ gauge theory. We saw this happen in an explicit example in Eq. (19.133).

In $SU(n)$ groups, $n \geq 3$, there is a unique symmetric invariant d^{abc} of the required type. It appears in the anticommutator of representation matrices of the fundamental representation:

$$\big\{t^a_n, t^b_n\big\} = \tfrac{1}{n}\delta^{ab} + d^{abc}t^c_n. \tag{19.138}$$

The uniqueness of this invariant implies that, in an $SU(n)$ gauge theory, any trace of the form of (19.132) is proportional to d^{abc}. For each representation r, we can define an *anomaly coefficient* $A(r)$ by

$$\operatorname{tr}\big[t^a_r\{t^b_r,t^c_r\}\big] = \tfrac{1}{2}A(r)d^{abc}. \tag{19.139}$$

For the fundamental representation, we can see from (19.138) that $A(n) = 1$. It follows from the argument of (19.137) that

$$A(\bar{r}) = -A(r). \tag{19.140}$$

For higher representations, the anomaly coefficients can be worked out using methods similar to those we used in Section 15.4 to compute $C_2(r)$. For example, we show in Problem 19.3 that, if a and s are the $SU(n)$ representations corresponding to antisymmetric and symmetric two-index tensors, then

$$A(a) = n - 4, \qquad A(s) = n + 4. \tag{19.141}$$

An $SU(n)$ gauge theory is anomaly free if the anomaly coefficients of the various irreducible components of the fermion multiplet R sum to zero. For example, the $SU(n)$ gauge theory of left-handed fermions with representation content

$$R = a + (n - 4)\bar{n} \tag{19.142}$$

is anomaly free.

Of the various simple Lie groups listed below (15.72), only $SU(n)$. $SO(4n+2)$, and E_6 have complex representations. Of these, only $SU(n)$ and $SO(6)$, which has the same Lie algebra as $SU(4)$, have a symmetric invariant of the type required to build the anomaly. Gauge theories based on $SO(4n+2)$, $n \geq 2$, and on E_6 are automatically anomaly free. The groups $SO(10)$ and E_6 have been suggested as candidates for the *grand unified* gauge symmetry of particle physics, which we will discuss in Section 22.2.

There is one further constraint on the representation content of a chiral gauge theory, which comes from considering its coupling to gravity. It is possible to show that the diagrams of Fig. 19.9 give an anomaly contribution when computed with a gauge current $j^{\mu a}$ and external gravitational fields. The group-theory factor that multiplies this diagram is

$$\mathrm{tr}\big[t_R^a\big]. \tag{19.143}$$

This factor automatically vanishes if the gauge group is non-Abelian. However, if the gauge group of the theory contains $U(1)$ factors, the theory cannot be consistently coupled to gravity unless each of the $U(1)$ generators is traceless.‡

Once we have constructed a consistent chiral gauge theory, we have an additional problem of finding a prescription for calculating in this theory consistently. In a vector-like gauge theory, we can define ultraviolet-divergent diagrams with dimensional regularization. This guarantees that the divergent diagrams will be regulated in a way that respects the Ward identities of local gauge invariance. To generalize dimensional regularization to chiral gauge theories, we need to introduce a dimensional continuation of γ^5. The 't Hooft-Veltman definition used to define the chiral current in Section 19.2 is not satisfactory, because this definition does not manifestly respect the conservation of the gauge currents. A useful alternative procedure is to define γ^5 formally as an object that anticommutes with all of the γ^μ. This prescription gives unambiguous, gauge-invariant results for amplitudes that are not proportional to $\epsilon^{\mu\nu\lambda\sigma}$, at least through two-loop order. In Section 21.3, we will

‡L. Alvarez-Gaumé and E. Witten, *Nucl. Phys.* **B234**, 269 (1984).

use this prescription to compute loop diagrams in weak interaction theory. As a last resort, one can always compute with a non-gauge-invariant regulator and add non-gauge-invariant counterterms to the theory so that the gauge theory Ward identities remain valid.

19.5 Anomalous Breaking of Scale Invariance

There is one more important example of a symmetry that is an invariance at the classical level and is broken by quantum corrections. This is the classical scale invariance of a massless field theory with a dimensionless coupling constant. In Chapter 12, we saw that a quantum field theory with no classical dimensionful parameters still depends on a mass scale through the regularization of ultraviolet divergences, or, equivalently, through the running of coupling constants. We have already seen how to analyze this induced dependence on the renormalization scale using the Callan-Symanzik equation. In this section, we will show how the violation of classical scale invariance by quantum corrections can be described as a current conservation anomaly.

In this book we have avoided giving a careful treatment of the energy-momentum tensor of a quantum field theory. In Section 2.2, we used Noether's theorem to demonstrate that the invariance of a quantum field theory under spacetime translations implies the presence of a conserved tensor $T^{\mu\nu}$. In Section 9.6, we gave an alternative derivation of this result using the functional integral formalism. However, to discuss the theory of scale invariance, we will need some more detailed properties of the energy-momentum tensor. We will now simply state these properties and refer elsewhere for their derivations.*

The tensor $T^{\mu\nu}$ defined by expressions (2.17) and (9.99) is called the *canonical energy-momentum tensor*. The expressions that defined this tensor do not imply that $T^{\mu\nu}$ is symmetric. In fact, this tensor need not be symmetric, and, in a gauge theory, it need not be gauge-invariant. However, it is always possible to convert $T^{\mu\nu}$ into a symmetric and gauge-invariant tensor $\Theta^{\mu\nu}$ by the addition

$$\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\lambda \Sigma^{\mu\nu\lambda}, \tag{19.144}$$

where $\Sigma^{\mu\nu\lambda}$ is antisymmetric under interchange of μ and λ. The form of the added term implies that $\Theta^{\mu\nu}$ is conserved if $T^{\mu\nu}$ is, and that the global energy-momentum four-vector is unchanged,

$$P^\nu = \int d^3x\, T^{0\nu} = \int d^3x\, \Theta^{0\nu}. \tag{19.145}$$

A scale transformation of a scalar field theory can be defined as a transformation of variables

$$\phi(x) \to e^{-D\sigma}\phi(xe^{-\sigma}), \tag{19.146}$$

---

*The conclusions presented in the next three paragraphs are derived with care in C. G. Callan, S. Coleman, and R. Jackiw, *Ann. Phys.* **59**, 42 (1970).

with $D = 1$, the canonical mass dimension of the field. The scale transformation is defined similarly in theories of fermion and gauge fields. If this transformation is an invariance of the classical Lagrangian, as it will be if there are no dimensionful couplings, this theory will possess a conserved current D^μ, called the *dilatation current.* The dilatation current has a simple relation to the symmetric energy-momentum tensor $\Theta^{\mu\nu}$: $D^\mu = \Theta^{\mu\nu}x_\nu$, so that

$$\partial_\mu D^\mu = \Theta^\mu{}_\mu. \tag{19.147}$$

The derivation of these results from Noether's theorem is not straightforward. There is a simpler derivation, which, however, uses formalism beyond the scope of this book. If the quantum field theory under consideration is coupled to gravity, then the energy-momentum tensor can be identified as the source of the gravitational field. This energy-momentum tensor can be found by varying the Lagrangian $\mathcal{L}_m$ of matter fields with respect to the spacetime metric $g_{\mu\nu}(x)$. This construction gives a manifestly symmetric and gauge-invariant tensor, which turns out to be $\Theta^{\mu\nu}$:

$$\Theta^{\mu\nu} = 2\frac{\delta}{\delta g_{\mu\nu}(x)}\int d^4x\,\mathcal{L}_m. \tag{19.148}$$

A scale transformation can be represented as a change in the spacetime metric

$$g_{\mu\nu}(x) \to e^{2\sigma}g_{\mu\nu}(x). \tag{19.149}$$

Combining (19.148) and (19.149), we see that the change in the Lagrangian induced by this transformation is just the trace of $\Theta^{\mu\nu}$. This will be equal by Noether's theorem to the divergence of the corresponding current, giving us back Eq. (19.147).

In QED, it is not hard to guess the form of the symmetric energy-momentum tensor:

$$\Theta^{\mu\nu} = -F^{\mu\lambda}F^\nu{}_\lambda + \tfrac{1}{4}g^{\mu\nu}(F_{\lambda\sigma})^2 + \tfrac{1}{2}\bar{\psi}i(\gamma^\mu D^\nu + \gamma^\nu D^\mu)\psi - g^{\mu\nu}\bar{\psi}(i\not{D} - m)\psi. \tag{19.150}$$

This is a gauge-invariant symmetric tensor that leads to the familiar expression for the total energy,

$$H = \int d^3x\Big\{\tfrac{1}{2}(E^2 + B^2) + \psi^\dagger(-i\gamma^0\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} + m)\psi\Big\}. \tag{19.151}$$

For future reference, we note that these results are true at the classical level in any spacetime dimension d. In four dimensions, the trace of the gauge field terms cancels automatically. After using the Dirac equation, which is valid as an operator equation of motion, we find that the trace of $\Theta^{\mu\nu}$ is given by

$$\Theta^\mu{}_\mu = m\bar{\psi}\psi \tag{19.152}$$

and indeed vanishes, classically, if $m = 0$.

When quantum corrections are included, we know that a scale transformation is not a symmetry of the theory, since the same theory referred to a larger scale contains a different value of the renormalized coupling constant. The shift of the renormalized coupling is

$$g \to g + \sigma\beta(g), \tag{19.153}$$

and the corresponding change in the Lagrangian is

$$\sigma\beta(g)\frac{\partial}{\partial g}\mathcal{L}. \tag{19.154}$$

Thus, when quantum corrections are included, the equation for the dilatation current in a classically scale-invariant theory should read

$$\partial_\mu D^\mu = \Theta^\mu{}_\mu = \beta(g)\frac{\partial}{\partial g}\mathcal{L}. \tag{19.155}$$

In massless QED, we can write this formula most usefully by rescaling the gauge fields so that the coupling constant e is removed from the covariant derivative: $eA^\mu \to A^\mu$. Then e appears only in the term

$$\mathcal{L} = -\frac{1}{4e^2}(F_{\lambda\sigma})^2 + \cdots, \tag{19.156}$$

and Eq. (19.155) reads

$$\Theta^\mu{}_\mu = \frac{\beta(e)}{2e^3}(F_{\lambda\sigma})^2. \tag{19.157}$$

This relation, which says that the trace of the symmetric energy-momentum tensor takes a nonzero value as a result of quantum corrections, is known as the *trace anomaly.*

We should be able to check the trace anomaly equation (19.157) directly in perturbation theory. We now evaluate the trace of $\Theta^{\mu\nu}$ explicitly to one-loop order. The formalism we have set up is very similar to that of the background field calculation of the β function done in Section 16.6. As in that section, we will integrate over the fluctuating parts of quantum fields in the presence of a background field with a nonzero $F_{\mu\nu}$. Equation (19.157) predicts that this integration will lead to the expression

$$\langle\Theta^\mu{}_\mu\rangle = C\int\frac{d^4k}{(2\pi)^4}A_\mu(-k)(k^2g^{\mu\nu} - k^\mu k^\nu)A_\nu(k), \tag{19.158}$$

where A_μ is the background field and the constant C is equal to $\beta(e)/e^3$.

Since we will be using dimensional regularization, we should begin by writing the trace of $\Theta^{\mu\nu}$ in d dimensions:

$$\Theta^\mu{}_\mu = -\frac{4-d}{4}(F_{\lambda\sigma})^2 + (1-d)\bar{\psi}i\not{D}\psi. \tag{19.159}$$

The one-loop matrix element of this quantity proportional to two powers of the background field arises from the three diagrams shown in Fig. 19.10. Since the second term on the right-hand side of (19.159) vanishes by the equation of motion of $\psi(x)$, one might expect that this term gives zero contribution to the trace. Indeed, it is easy to check that the first two diagrams in Fig.

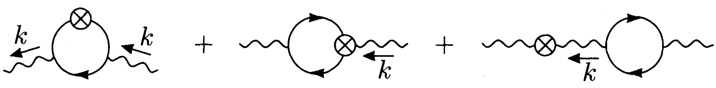
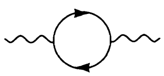


Figure 19.10. One-loop diagrams contributing to the anomalous trace of $\Theta^{\mu\nu}$.

19.10 cancel: These diagrams have the same structure, since the first has an extra propagator and an extra factor $\not{p}$ from the operator matrix element, and opposite overall signs.

The first term on the left-hand side of (19.159) is unexpected, since it apparently vanishes in four dimensions. However, the fermion loop diagram is divergent, and in dimensional regularization, this introduces a factor $1/(2-d/2)$. As a result, this diagram gives a nonzero contribution to the operator matrix element. In massless QED, the fermion loop diagram has the value



$$= -i(k^2 g^{\mu\nu} - k^\mu k^\nu)\frac{4}{3(4\pi)^2}\big(\Gamma(2-\tfrac{d}{2}) + (\text{finite})\big). \qquad (19.160)$$

Then the complete expression for the third diagram in Fig. 19.10 is

$$\int \frac{d^4k}{(2\pi)^4} A_\mu(-k)\Big(-2\frac{4-d}{4}\Big)(k^2 g^{\mu\nu} - k^\mu k^\nu)\frac{-i}{k^2}\Big(-ik^2\frac{4}{3(4\pi)^2}\frac{1}{2-d/2}\Big)A_\nu(k). \qquad (19.161)$$

This is of the form of (19.158), with

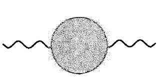
$$C = \frac{1}{12\pi^2}, \qquad (19.162)$$

which is indeed the first β function coefficient in massless QED.

This discussion generalizes to QCD and other gauge theories. In a non-Abelian gauge theory, $\Theta^{\mu\nu}$ is given by the obvious generalization of (19.150) with the Abelian field-strength tensor $F_{\mu\nu}$ replaced by the non-Abelian expression $F^a_{\mu\nu}$. The trace of $\Theta^{\mu\nu}$ is again given by

$$\Theta^\mu{}_\mu = -\frac{4-d}{4}(F^a_{\lambda\sigma})^2, \qquad (19.163)$$

plus terms that vanish by the equations of motion. In the background field gauge, the one-loop diagrams with the operator $\Theta^\mu{}_\mu$ inserted into the loop cancel as above. We saw in Section 16.6 that the two-point functions in this gauge sum to



$$= -i(k^2 g^{\mu\nu} - k^\mu k^\nu)\Big[\frac{-b_0}{(4\pi)^2}\Big]\big(\Gamma(2-\tfrac{d}{2}) + (\text{finite})\big), \qquad (19.164)$$

where $\beta(g) = -b_0 g^3/(4\pi)^2$. Following through the logic of the previous paragraph, we again find the result (19.158) with the identification of C as the first β function coefficient.

As with the axial vector anomaly, the trace anomaly can be found in many different ways. For each possible method of regulating a quantum field theory, there is a derivation of the trace anomaly that exploits the possible pathology of that particular regulator. For example, if one uses a Pauli-Villars regulator with heavy fermions to cancel the divergence of the QED fermion loop diagram, the heavy fermions Ψ contribute a term $M\overline{\Psi}\Psi$ to the trace of $\Theta^{\mu\nu}$. The loop diagram with this term inserted turns out to have a finite limit as $M \to \infty$, which precisely reproduces the trace anomaly. This computation is worked out in Problem 19.4.

As with the axial vector anomaly, each derivation of the anomaly with a different regulator, taken individually, seems artificial, as if there were a problem with the field theory that we are not quite clever enough to fix. Eventually, though, we are forced to conclude that the quantum field theory is trying to tell us something. The anomalous symmetries of the classical theory cannot be promoted to symmetries of the quantum theory. Instead, the anomalous conservation laws require profound and qualitative changes in the theory from the classical to the quantum level.

Problems

19.1 Fermion number nonconservation in parallel E and B fields.

(a) Show that the Adler-Bell-Jackiw anomaly equation leads to the following law for global fermion number conservation: If N_R and N_L are, respectively, the numbers of right- and left-handed massless fermions, then

$$\Delta N_R - \Delta N_L = -\frac{e^2}{2\pi^2}\int d^4x\, \mathbf{E}\cdot\mathbf{B}.$$

To set up a solvable problem, take the background field to be $A^\mu = (0, 0, Bx^1, A)$, with B constant and A constant in space and varying only adiabatically in time.

(b) Show that the Hamiltonian for massless fermions represented in the components (3.36) is

$$H = \int d^3x \Big[\psi_R^\dagger(-i\boldsymbol{\sigma}\cdot\mathbf{D})\psi_R - \psi_L^\dagger(-i\boldsymbol{\sigma}\cdot\mathbf{D})\psi_L\Big],$$

with $D^i = \nabla^i - ieA^i$. Concentrate on the term in the Hamiltonian that involves right-handed fermions. To diagonalize this term, one must solve the eigenvalue equation $-i\boldsymbol{\sigma}\cdot\mathbf{D}\psi_R = E\psi_R$.

(c) The ψ_R eigenvectors can be written in the form

$$\psi_R = \begin{pmatrix}\phi_1(x^1)\\ \phi_2(x^1)\end{pmatrix} e^{i(k_2x^2 + k_3x^3)}.$$

The functions ϕ_1 and ϕ_2, which depend only on x^1, obey coupled first-order differential equations. Show that, when one of these functions is eliminated, the other obeys the equation of a simple harmonic oscillator. Use this observation to find the single-particle spectrum of the Hamiltonian. Notice that the eigenvalues do not depend on k_2.

(d) If the system of fermions is set up in a box with sides of length L and periodic boundary conditions, the momenta k_2 and k_3 will be quantized:

$$k_i = \frac{2\pi n_i}{L}.$$

By looking back to the harmonic oscillator equation in part (c), show that the condition that the center of the oscillation is inside the box leads to the condition

$$k_2 < eBL.$$

Combining these two conditions, we see that each level found in part (c) has a degeneracy of

$$\frac{eL^2B}{2\pi}.$$

(e) Now consider the effect of changing the background A adiabatically by an amount (19.37). Show that the vacuum loses right-handed fermions. Repeating this analysis for the left-handed spectrum, one sees that the vacuum gains the same number of left-handed fermions. Show that these numbers are in accord with the global nonconservation law given in part (a).

19.2 Weak decay of the pion.

(a) In the effective Lagrangian for semileptonic weak interactions (18.28), the hadronic part of the operator is a left-handed current involving the u and d quarks. Show that this current is related to the quark currents of Section 19.3 as follows:

$$\bar{u}_L\gamma^\mu d_L = \frac{1}{2}(j^{\mu 1} + ij^{\mu 2} - j^{\mu 51} - ij^{\mu 52}),$$

where 1, 2 are isospin indices. Using this identification and (19.88), show that the amplitude for the decay $\pi^+ \to \ell^+\nu$ is given by

$$i\mathcal{M} = G_F f_\pi \bar{u}(q)\not{p}(1-\gamma^5)v(k),$$

where p, k, q are the momenta of the π^+, ℓ^+, ν.

(b) Compute the decay rate of the pion. Show that this rate vanishes in the limit of zero lepton mass, and that the relative rate of pion decay to muons and electrons is given by

$$\frac{\Gamma(\pi^+ \to e^+\nu)}{\Gamma(\pi^+ \to \mu^+\nu)} = \left(\frac{m_e}{m_\mu}\right)^2 \frac{(1-m_e^2/m_\pi^2)^2}{(1-m_\mu^2/m_\pi^2)^2} = 10^{-4}.$$

From the measured pion lifetime, $\tau_\pi = 2.6 \times 10^{-8}$ sec, and the pion and muon masses, $m_\pi = 140$ MeV, $m_\mu = 106$ MeV, determine the value of f_π.

19.3 Computation of anomaly coefficients.

(a) Consider a product $r_1 \times r_2$ of $SU(n)$ representations, which is decomposed into irreducible representations as in (15.98). Using the explicit form of the generators given in (15.99), show that the anomaly coefficients satisfy

$$d_1 A(r_2) + d_2 A(r_1) = \sum_i A(r_i).$$

(b) As we saw in Problem 15.5, the two-index symmetric and antisymmetric tensors form irreducible representations of $SU(n)$, which we will call s and a, respectively. In $SU(3)$, the representation a is three-dimensional. Show that it is equivalent to the $\bar{3}$. Compute the anomaly coefficients for a and s, making use of the identity in part (a).

(c) Since $SU(n)$ has a unique three-index symmetric tensor d^{abc} which is already nonvanishing in an $SU(3)$ subgroup, we can compute the anomaly coefficient in $SU(n)$ by restricting our attention to three generators in this subgroup. By decomposing $SU(n)$ representations into $SU(3)$ representations, compute the anomaly coefficients for a and s in $SU(n)$ and derive Eq. (19.141). Find the anomaly coefficient of the j-index totally antisymmetric tensor representation of $SU(n)$. Why does the result always vanish when $2j = n$?

19.4 Large fermion mass limits. In the text, we derived the Adler-Bell-Jackiw and trace anomalies by the use of dimensional regularization. As an alternative, one could imagine deriving these results using Pauli-Villars regularization. In that technique, one regularizes the value of a fermion loop integral by subtracting the value of the same loop diagram computed with fermions Ψ of large mass M. The parameter M plays the role of the cutoff and should be taken to infinity at the end of the calculation. The anomalies arise because some pieces of the diagrams computed for very heavy fermions do not disappear in the limit $M \to \infty$. These nontrivial $M \to \infty$ limits are interesting in their own right and can have physical applications (for example, in part (c) of the Final Project for Part III).

(a) Show that the Adler-Bell-Jackiw anomaly equation is equivalent to the following large-mass limit of a fermion matrix element between the vacuum and a two-photon state:

$$\lim_{M\to\infty} \left\{ \langle p,k|\, 2iM\bar{\Psi}\gamma^5\Psi\, |0\rangle \right\} = -\frac{e^2}{2\pi^2}\epsilon^{\alpha\nu\beta\lambda}p_\alpha\epsilon^*_\nu(p)k_\beta\epsilon^*_\lambda(k).$$

(b) Show that the trace anomaly, at one-loop order, is equivalent to the following large-mass limit:

$$\lim_{M\to\infty} \left\{ \langle p,k|\, M\bar{\Psi}\Psi\, |0\rangle \right\} = +\frac{e^2}{6\pi^2}\big[p\cdot k\,\epsilon^*(p)\cdot\epsilon^*(k) - p\cdot\epsilon^*(k)\,k\cdot\epsilon^*(p)\big].$$

(c) Show that the matrix element in part (a) is ultraviolet-finite before the $M \to \infty$ limit is taken. Evaluate the matrix element explicitly at one-loop order and verify the limit claimed in part (a).

(d) The matrix element in part (b) has a potential ultraviolet divergence. However, show that the coefficient of $(p\cdot\epsilon^*(k)k\cdot\epsilon^*(p))$ is ultraviolet-finite, and that the rest of the expression is determined by gauge invariance. Compute the full matrix element using dimensional regularization as a gauge-invariant regulator and verify the result claimed in part (b).

Chapter 20

Gauge Theories with Spontaneous Symmetry Breaking

In the course of this book, we have discussed three distinct fashions in which symmetries can be realized in a quantum field theory. The simplest case is a global symmetry that is manifest, leading to particle multiplets with restricted interactions. A second possibility is a global symmetry that is spontaneously broken. Then, as discussed in Chapter 11,* the symmetry currents are still conserved and interactions are similarly restricted, but the vacuum state does not respect the symmetry and the particles do not form obvious symmetry multiplets. Instead, such a theory contains massless particles, Goldstone bosons, one for each generator of the spontaneously broken symmetry. The third case is that of a local, or gauge, symmetry. As we saw in Chapter 15, such a symmetry requires the existence of a massless vector field for each symmetry generator, and the interactions among these fields are highly restricted.

It is now only natural to consider a fourth possibility: What happens if we include both local gauge invariance and spontaneous symmetry breaking in the same theory? In this chapter and the next, we will find that this combination of ingredients leads to new possibilities for the construction of quantum field theory models. We will see that spontaneous symmetry breaking requires gauge vector bosons to acquire mass. However, the interactions of these massive bosons are still constrained by the underlying gauge symmetry, and these constraints can have observable consequences.

In elementary particle physics, the principal application of spontaneously broken local symmetry is in the currently accepted model of weak interactions. This model, due to Glashow, Weinberg, and Salam, is introduced in Section 20.2. There we will see that it makes a number of precise and successful predictions for weak interaction phenomena. Remarkably, this model also unifies the weak interactions with electromagnetism in a single larger gauge theory.

*Section 11.1 is necessary background for the present chapter, but the rest of Chapter 11 is not.

20.1 The Higgs Mechanism

In this section we analyze some simple examples of gauge theories with spontaneous symmetry breaking. We begin with an Abelian gauge theory, and then study several examples of non-Abelian models.

An Abelian Example

As our first example, consider a complex scalar field coupled both to itself and to an electromagnetic field:

$$\mathcal{L} = -\tfrac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 - V(\phi), \tag{20.1}$$

with $D_\mu = \partial_\mu + ieA_\mu$. This Lagrangian is invariant under the local $U(1)$ transformation

$$\phi(x) \to e^{i\alpha(x)}\phi(x), \qquad A_\mu(x) \to A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x). \tag{20.2}$$

If we choose the potential in $\mathcal{L}$ to be of the form

$$V(\phi) = -\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2, \tag{20.3}$$

with $\mu^2 > 0$, the field ϕ will acquire a vacuum expectation value and the $U(1)$ global symmetry will be spontaneously broken. The minimum of this potential occurs at

$$\langle\phi\rangle = \phi_0 = \left(\frac{\mu^2}{\lambda}\right)^{1/2}, \tag{20.4}$$

or at any other value related by the $U(1)$ symmetry (20.2).

Let us expand the Lagrangian (20.1) about the vacuum state (20.4). Decompose the complex field $\phi(x)$ as

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}}\big(\phi_1(x) + i\phi_2(x)\big). \tag{20.5}$$

The potential (20.3) is rewritten

$$V(\phi) = -\frac{1}{2\lambda}\mu^4 + \frac{1}{2}\cdot 2\mu^2\phi_1^2 + \mathcal{O}(\phi_i^3), \tag{20.6}$$

so that the field ϕ_1 acquires the mass $m = \sqrt{2}\mu$ and ϕ_2 is the massless Goldstone boson. So far, this whole analysis follows that in Section 11.1.

But now consider how the kinetic energy term of ϕ is transformed. Inserting the expansion (20.5), we rewrite

$$|D_\mu\phi|^2 = \tfrac{1}{2}(\partial_\mu\phi_1)^2 + \tfrac{1}{2}(\partial_\mu\phi_2)^2 + \sqrt{2}e\phi_0\cdot A_\mu\partial^\mu\phi_2 + e^2\phi_0^2A_\mu A^\mu + \cdots, \tag{20.7}$$

where we have omitted terms cubic and quartic in the fields A_μ, ϕ_1, and ϕ_2. The last term written explicitly in (20.7) is a photon mass term

$$\Delta\mathcal{L} = \tfrac{1}{2}m_A^2 A_\mu A^\mu, \tag{20.8}$$

where the mass

$$m_A^2 = 2e^2\phi_0^2 \tag{20.9}$$

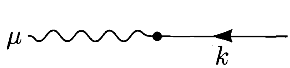
arises from the nonvanishing vacuum expectation value of ϕ. Notice that the sign of this mass term is correct; the physical spacelike components of A^μ appear in (20.8) as

$$\Delta\mathcal{L} = -\tfrac{1}{2}m_A^2(A^i)^2,$$

with the correct sign for a potential energy term.

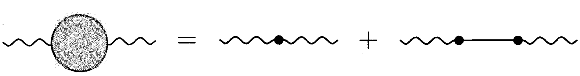
In Chapter 7, and again in Chapter 16 for the non-Abelian case, we argued that a gauge boson cannot obtain a mass, unless this mass term is associated with a pole in the vacuum polarization amplitude. There is a counterexample to this result in two-dimensional spacetime; there, as we saw in Section 19.1, a pole of the required form can arise from the infrared singularity generated by a massless fermion pair. However, in four dimensions, a pole in the vacuum polarization amplitude can be created only by a massless scalar particle. Typically, in situations with unbroken symmetry, no such particle is available.

However, a model with a spontaneously broken continuous symmetry must have massless Goldstone bosons. These scalar particles have the quantum numbers of the symmetry currents, and therefore have just the right quantum numbers to appear as intermediate states in the vacuum polarization. In the model we are now discussing, we can see this pole arise explicitly in the following way: The third term in Eq. (20.7) couples the gauge boson directly to the Goldstone boson ϕ_2; this gives a vertex of the form



$$= i\sqrt{2}e\phi_0(-ik^\mu) = m_A k^\mu. \tag{20.10}$$

If we also treat the mass term (20.8) as a vertex in perturbation theory, then the leading-order contributions to the vacuum polarization amplitude give the expression



$$\begin{aligned} &= im_A^2 g^{\mu\nu} + (m_A k^\mu)\frac{i}{k^2}(-m_A k^\nu) \\ &= im_A^2\Big(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\Big). \end{aligned} \tag{20.11}$$

The Goldstone boson supplies exactly the right pole to make the vacuum polarization amplitude properly transverse.

Although the Goldstone boson plays an important formal role in this theory, it does not appear as an independent physical particle. The easiest way to see this is to make a particular choice of gauge, called the *unitarity gauge.* Using the local $U(1)$ gauge symmetry (20.2), we can choose $\alpha(x)$ in such a way that $\phi(x)$ becomes real-valued at every point x. With this choice, the field ϕ_2 is removed from the theory. The Lagrangian (20.1) becomes

$$\mathcal{L} = -\tfrac{1}{4}(F_{\mu\nu})^2 + (\partial_\mu\phi)^2 + e^2\phi^2 A_\mu A^\mu - V(\phi). \tag{20.12}$$

If the potential $V(\phi)$ favors a nonzero vacuum expectation value of ϕ, the gauge field acquires a mass; it also retains a coupling to the remaining, physical field ϕ_1.

This mechanism, by which spontaneous symmetry breaking generates a mass for a gauge boson, was explored and generalized to the non-Abelian case by Higgs, Kibble, Guralnik, Hagen, Brout, and Englert, and is now known as the *Higgs mechanism.* However, this mechanism had an earlier application to the theory of superconductivity. In Chapter 8, we constructed the Landau description of a second-order phase transition. To describe a superconductor, Landau and Ginzburg coupled this theory to an external electromagnetic field; they obtained precisely the Lagrangian (20.1). Since the gauge field acquires a nonzero mass, external electromagnetic fields penetrate a superconductor only to the depth m_A^{-1}. This explains the *Meissner effect*, the observed exclusion of macroscopic magnetic fields from a superconductor.

The role of the Goldstone boson in the Higgs mechanism is intricate, and seems mysterious at this level of the discussion. We first saw that the involvement of the Goldstone boson is necessary, as a matter of principle, in order for the gauge boson to acquire a mass. We then saw that the Goldstone boson can be formally eliminated from the theory. However, we might argue that the Goldstone boson has not completely disappeared. A massless vector boson has only two physical polarization states; we saw in Chapter 16 that the longitudinal polarization state cannot be produced, and appears in the formalism only to cancel other unphysical contributions. However, a massive vector boson must have three physical polarization states: In its rest frame, it is a spin-1 object, which can make no distinction between transverse and longitudinal polarizations. It is tempting to say that the gauge boson acquired its extra degree of freedom by *eating* the Goldstone boson. In Sections 21.1 and 21.2 we will clarify this picture, by studying the quantization and gauge-fixing of spontaneously broken gauge theories.

Systematics of the Higgs Mechanism

The Higgs mechanism extends straightforwardly to systems with non-Abelian gauge symmetry. It is not difficult to derive the general relation by which a set of scalar field vacuum expectation values leads to the appearance of gauge boson masses. Let us work out this relation and then apply it in a number of examples.

Consider a system of scalar fields ϕ_i that appear in a Lagrangian invariant under a symmetry group G, represented by the transformation

$$\phi_i \to (1 + i\alpha^a t^a)_{ij}\phi_j. \tag{20.13}$$

It is convenient to write the ϕ_i as real-valued fields, for example, writing n complex fields as $2n$ real fields. Then the group representation matrices t^a must be pure imaginary and, since they are Hermitian, antisymmetric. Let us write

$$t^a_{ij} = iT^a_{ij}, \tag{20.14}$$

so that the T^a are real and antisymmetric.

If we promote the symmetry group G to a local gauge symmetry, the covariant derivative on the ϕ_i is

$$D_\mu \phi = (\partial_\mu - igA^a_\mu t^a)\phi = (\partial_\mu + gA^a_\mu T^a)\phi.$$

Then the kinetic energy term for the ϕ_i is

$$\tfrac{1}{2}(D_\mu \phi_i)^2 = \tfrac{1}{2}(\partial_\mu \phi_i)^2 + gA^a_\mu(\partial_\mu \phi_i T^a_{ij}\phi_j) + \tfrac{1}{2}g^2 A^a_\mu A^{b\mu}(T^a\phi)_i(T^b\phi)_i. \tag{20.15}$$

Now let the ϕ_i acquire vacuum expectation values

$$\langle \phi_i \rangle = (\phi_0)_i, \tag{20.16}$$

and expand the ϕ_i about these values. The last term in Eq. (20.15) contains a term with the structure of a gauge boson mass,

$$\Delta\mathcal{L} = \tfrac{1}{2}m^2_{ab}A^a_\mu A^{b\mu}, \tag{20.17}$$

with the mass matrix

$$m^2_{ab} = g^2(T^a\phi_0)_i(T^b\phi_0)_i. \tag{20.18}$$

This matrix is positive semidefinite, since any diagonal element, in any basis, has the form

$$m^2_{aa} = g^2(T^a\phi_0)^2 \geq 0 \qquad \text{(no sum)}.$$

Thus, generically, all of the gauge bosons will receive positive masses. However, it may be that some particular generator T^a of G leaves the vacuum invariant:

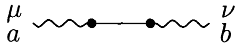
$$T^a\phi_0 = 0.$$

In that case, the generator T^a will give no contribution to (20.18), and the corresponding gauge boson will remain massless.

As in the Abelian case, the gauge boson propagator receives a contribution from the Goldstone bosons, which is necessary to make the vacuum polarization amplitude transverse. To compute this contribution, we need the vertex that mixes gauge bosons and Goldstone bosons. This comes from the second term of the Lagrangian (20.15). When we insert the vacuum expectation value of the scalar field (20.16), this term becomes

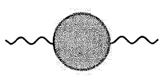
$$\Delta\mathcal{L} = gA^a_\mu \partial_\mu \phi_i (T^a\phi_0)_i. \tag{20.19}$$

This interaction term does not involve all of the components of ϕ—only those that are parallel to a vector $T^a\phi_0$ for some choice of T^a. These vectors represent the infinitesimal rotations of the vacuum; thus the components ϕ_i that appear in (20.19) are precisely the Goldstone bosons. Using the fact that these bosons are massless, we can compute the counterpart, for the non-Abelian case, of the Goldstone boson diagram in Eq. (20.11):



$$= \sum_j \big(gk^\mu (T^a\phi_0)_j\big) \frac{i}{k^2} \big(-gk^\nu (T^b\phi_0)_j\big). \tag{20.20}$$

The sum runs over those components j with a nonzero projection onto the space spanned by the $T^a\phi_0$, or equally well, over all j. This diagram is therefore proportional to the mass matrix (20.18). Combining this expression with the contribution to the vacuum polarization from (20.17), we find a properly transverse result,



$$= i m_{ab}^2 \Big(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\Big), \tag{20.21}$$

where m_{ab}^2 is given by Eq. (20.18).

Non-Abelian Examples

Let us now apply this general formalism to some specific examples of non-Abelian gauge theories. Consider first a model with an $SU(2)$ gauge field coupled to a scalar field ϕ that transforms as a spinor of $SU(2)$. The covariant derivative acting on ϕ is

$$D_\mu \phi = (\partial_\mu - igA_\mu^a \tau^a)\phi, \tag{20.22}$$

where $\tau^a = \sigma^a/2$. The square of this expression is the scalar field kinetic energy term.

If ϕ acquires a vacuum expectation value, we can use the freedom of $SU(2)$ rotations to write this expectation value in the form

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{20.23}$$

Then the gauge boson masses arise from

$$|D_\mu \phi|^2 = \frac{1}{2} g^2 \begin{pmatrix} 0 & v \end{pmatrix} \tau^a \tau^b \begin{pmatrix} 0 \\ v \end{pmatrix} A_\mu^a A^{b\mu} + \cdots. \tag{20.24}$$

We can symmetrize the matrix product under the interchange of a and b; using $\{\tau^a, \tau^b\} = \frac{1}{2}\delta^{ab}$, we find the mass term

$$\Delta\mathcal{L} = \frac{g^2 v^2}{8} A_\mu^a A^{a\mu}. \tag{20.25}$$

All three gauge bosons receive the mass

$$m_A = \frac{gv}{2}, \tag{20.26}$$

signaling that all three generators of $SU(2)$ are broken equally well by (20.23).

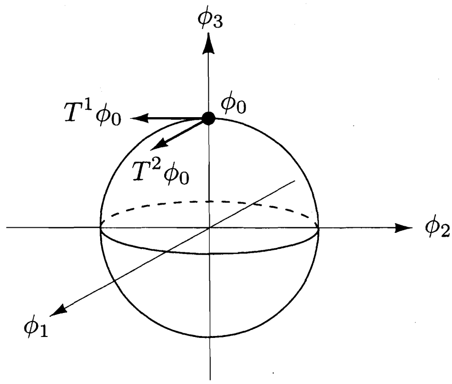


Figure 20.1. The space of configurations for a scalar field in the vector representation of $SU(2)$. When the $SU(2)$ symmetry is spontaneously broken, the allowed vacuum states lie on a spherical surface. If the vacuum expectation value ϕ_0 lies in the 3 direction, then the generator T^3 leaves ϕ_0 unchanged, while T^1 and T^2 rotate ϕ_0 in the directions shown.

What if we had taken ϕ to transform according to the vector representation of $SU(2)$? If we take ϕ to be a real-valued vector under $SU(2)$, we must assign it the covariant derivative

$$(D_\mu \phi)_a = \partial_\mu \phi_a + g\epsilon_{abc} A^b_\mu \phi_c. \tag{20.27}$$

Again, the ϕ kinetic energy term is the square of this object, and so, if ϕ acquires a vacuum expectation value, we find the gauge boson mass term

$$\Delta\mathcal{L} = \frac{1}{2}(D_\mu\phi)^2 = \frac{g^2}{2}\big(\epsilon_{abc}A^b_\mu(\phi_0)_c\big)^2 + \cdots. \tag{20.28}$$

If a vector of $SU(2)$ acquires an expectation value ϕ_0, we can choose our coordinates so that this vector points in any particular direction in the internal space. We will choose it to point in the 3 direction, as indicated in Fig. 20.1:

$$\langle\phi_c\rangle = (\phi_0)_c = V\delta_{c3}. \tag{20.29}$$

Inserting (20.29) into (20.28), we find

$$\Delta\mathcal{L} = \frac{g^2}{2}V^2\big(\epsilon_{ab3}A^b_\mu\big)^2 = \frac{g^2}{2}V^2\big((A^1_\mu)^2 + (A^2_\mu)^2\big). \tag{20.30}$$

The gauge bosons corresponding to the generators 1 and 2 acquire masses

$$m_1 = m_2 = gV, \tag{20.31}$$

while the boson corresponding to the generator 3 remains massless. It is easy to see the reason for this distinction by glancing at Fig. 20.1. The vacuum expectation value of ϕ_c destroys the symmetry of rotation about the axes 1 and 2, but it preserves the symmetry of rotation about the 3 axis. As we saw in our general analysis, gauge bosons corresponding to unbroken symmetry generators remain massless.

It is interesting that this model contains both massive and massless gauge bosons, with the distinction between these bosons created by spontaneous symmetry breaking. If we interpret the massive bosons as W bosons and the massless gauge boson as the photon, it is tempting to interpret this theory as a unified model of weak and electromagnetic interactions. Georgi and Glashow proposed this model as a serious candidate for the theory of weak interactions.† However, Nature chooses a different model, which we will discuss in the next section.

We turn next to a more complicated example. Consider an $SU(3)$ gauge theory with a scalar field in the adjoint representation. The covariant derivative of ϕ takes the form

$$D_\mu \phi_a = \partial_\mu \phi_a + g f_{abc} A^b_\mu \phi_c, \tag{20.32}$$

and so the gauge field masses arise from the term

$$\Delta\mathcal{L} = \frac{g^2}{2}\left(f_{abc} A^b_\mu \phi_c\right)^2. \tag{20.33}$$

We can write this more clearly by defining the quantity

$$\Phi = \phi_c t^c, \tag{20.34}$$

where t^c are the 3×3 traceless Hermitian matrices that represent the generators of $SU(3)$. Using this notation and the definition (15.68) of the structure constants, we can rewrite the mass term (20.33) as

$$\Delta\mathcal{L} = -g^2 \,\mathrm{tr}\big[[t^a, \Phi][t^b, \Phi]\big] A^a_\mu A^{b\mu}. \tag{20.35}$$

Now let Φ acquire a vacuum expectation value

$$\langle \Phi \rangle = \Phi_0. \tag{20.36}$$

Since Φ_0 is a traceless Hermitian matrix, we should analyze its effects by diagonalizing it. In principle, Φ_0 could have three arbitrary eigenvalues that sum to zero. However, when one minimizes explicit potential energy functions, one often finds expectation values that preserve some of the original symmetry. We will consider two examples.

First, Φ_0 might have the orientation

$$\Phi_0 = |\phi| \cdot \begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix}. \tag{20.37}$$

†H. Georgi and S. L. Glashow, *Phys. Rev. Lett.* **28**, 1494 (1972).

This matrix commutes with the four $SU(3)$ generators

$$t^a = \begin{pmatrix} \tau^a & 0 \\ 0 & 0 \end{pmatrix}, \qquad t^8 = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix}. \tag{20.38}$$

Thus, the expectation value (20.37) breaks $SU(3)$ spontaneously to $SU(2) \times U(1)$ and leaves the gauge bosons corresponding to these four generators massless. The remaining four generators of $SU(3)$,

$$\begin{aligned} t^4 &= \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & t^5 &= \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \\ t^6 &= \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & t^7 &= \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \end{aligned} \tag{20.39}$$

acquire the masses

$$m^2 = \left(3g|\phi|\right)^2, \tag{20.40}$$

as one can check by substituting these matrices into Eq. (20.35).

Another possible orientation for Φ_0 is

$$\Phi_0 = |\phi| \cdot \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}. \tag{20.41}$$

In this case, only t^3 and t^8 commute with Φ_0, so the original $SU(3)$ symmetry is broken down to $U(1)\times U(1)$. By substituting into (20.35), one can determine that the gauge bosons corresponding to the remaining generators of $SU(3)$ acquire the masses

$$t^1, t^2: \quad m^2 = \left(2g|\phi|\right)^2, \qquad t^4, t^5, t^6, t^7: \quad m^2 = \left(g|\phi|\right)^2. \tag{20.42}$$

Still larger symmetry groups offer a wider variety of symmetry-breaking patterns, and more complex mass matrices. We consider one further example in Problem 20.1.

Formal Description of the Higgs Mechanism

Up to this point, our study of the Higgs mechanism has been based on the explicit analysis of scalar field Lagrangians coupled to gauge fields. Scalar field theories provide the simplest examples of systems with spontaneous symmetry breaking, and the explicit calculations they allow are useful for visualization. But symmetries can be broken in other ways. In the theory of superconductivity, for example, the Abelian gauge invariance of electromagnetism is broken by pairs of electrons that condense in the ground state of a metal. In Section 19.3, we argued that, in the approximation that quark masses are very small, QCD possesses global symmetries that are spontaneously broken by a condensation of quark-antiquark pairs. In these examples, spontaneous symmetry breaking is the result of strong interactions beyond perturbation theory. We would like to understand whether these more general mechanisms of spontaneous symmetry breaking can also give mass to vector bosons, and, if so, how the masses can be calculated.

To carry out this analysis, we will need to abstract several ideas from the preceding discussion. First, we will discuss in general terms the relations between gauge bosons, Goldstone bosons, and global symmetry currents. Then we will use this information to construct the gauge boson mass matrix without making direct use of the Lagrangian.

Consider, first, an arbitrary quantum field theory $\mathcal{L}_0$ with a global symmetry G. In Section 9.6, we derived the Noether current corresponding to the G symmetry by varying the Lagrangian by a local gauge transformation with infinitesimal parameter $\alpha^a(x)$. Transforming with a constant α^a should leave $\mathcal{L}_0$ unchanged. Then the more general variation of $\mathcal{L}_0$ must take the form

$$\delta\mathcal{L}_0 = (\partial_\mu \alpha^a) J^{\mu a}, \tag{20.43}$$

for some set of vector operators $J^{\mu a}$ built from the fields of $\mathcal{L}_0$. The variational principle then tells us that

$$\partial_\mu J^{\mu a} = 0. \tag{20.44}$$

We can identify the $J^{\mu a}$ as the Noether currents of the global gauge symmetry.

We can now couple this globally symmetric theory to non-Abelian gauge fields, promoting the global symmetry to a local symmetry. To first order in g, the new Lagrangian should take the form

$$\mathcal{L} = \mathcal{L}_0 - gA_\mu^a J^{\mu a} + \mathcal{O}(A^2). \tag{20.45}$$

To check this, note that the transformation (20.43) compensates the variation due to a gauge transformation of A_μ^a, Eq. (15.46), to leading order in g. The terms of order A^2 and higher can in general be arranged to compensate the higher-order terms in the gauge transformation. Thus, matrix elements involving only one insertion of the gauge field can be evaluated using properties of the Noether currents of the original globally symmetric theory. Note in particular that the conservation law for these currents, Eq. (20.44), guarantees that the Ward identities for these matrix elements are satisfied.

If the global symmetry of the theory $\mathcal{L}_0$ is spontaneously broken, this theory will contain Goldstone bosons, which will stand in a special relation to the Noether currents. At long wavelength, the Goldstone bosons become infinitesimal symmetry rotations of the vacuum, $Q^a\left|0\right\rangle$, where Q^a is the global charge associated with $J^{\mu a}$. Thus, the operators $J^{\mu a}$ have the correct quantum numbers to create Goldstone bosons from the vacuum. Let $\left|\pi_k\right\rangle$ denote a Goldstone boson state. In general, there will be a current $J^{\mu a}$ that can create or destroy this boson; we can parametrize the corresponding matrix element as

$$\left\langle 0\right| J^{\mu a}(x) \left|\pi_k(p)\right\rangle = -ip^\mu F^a{}_k e^{-ip\cdot x}, \tag{20.46}$$

where p^μ is the on-shell momentum of the boson and $F^a{}_k$ is a matrix of constants. The elements $F^a{}_k$ vanish when a denotes a generator of an unbroken symmetry. Then the nonvanishing matrix elements of $F^a{}_k$ connect the currents of the spontaneously broken symmetries to their corresponding Goldstone bosons. Since the currents $J^{\mu a}$ are conserved, we find

$$0 = \partial_\mu \langle 0| J^{\mu a}(x) |\pi_k(p)\rangle = -p^2 F^a{}_k e^{-ip\cdot x}, \tag{20.47}$$

which implies that the bosons with nonzero matrix elements (20.46) satisfy $p^2 = 0$ on shell and so are massless. This is another proof of Goldstone's theorem.‡

Since the scalar field theory that we examined earlier in this section should be a special case of this analysis, we should find there an example of the relation given in Eq. (20.46). Comparing Eqs. (20.15) and (20.45), we see that, for the scalar field theory,

$$J^{\mu a} = \partial_\mu \phi_i T^a_{ij} \phi_j, \tag{20.48}$$

which is indeed the Noether current corresponding to the global gauge symmetry. Inserting the vacuum expectation value (20.16), we find

$$J^{\mu a} = \partial_\mu \phi_i (T^a \phi_0)_i, \tag{20.49}$$

which leads to the set of matrix elements

$$\langle 0| J^{\mu a}(x) |\phi_i(p)\rangle = -ip^\mu (T^a\phi_0)_i e^{-ip\cdot x}. \tag{20.50}$$

Using this relation, we can identify

$$F^a{}_i = T^a_{ij}\phi_{0j} \tag{20.51}$$

for the Higgs mechanism in a weakly coupled scalar field theory. To be more precise, the index i runs over all components of the scalar field ϕ. However, we saw in the discussion below Eq. (20.19) that (20.51) is nonzero only for components ϕ_i that are Goldstone bosons, and only for symmetry generators a that are spontaneously broken. Thus the nonzero components of (20.51) form precisely the structure (20.46).

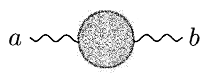
As a concrete illustration of the way that the objects $T^a\phi_0$ link spontaneously broken generators and Goldstone bosons, consider the situation of $SU(2)$ symmetry broken by a scalar field in the vector representation, as in Eq. (20.29) and Fig. 20.1. According to the figure, rotations about the $\hat{1}$ axis tip the vacuum expectation value of ϕ into the $\hat{2}$ direction, rotations about the $\hat{2}$ axis tip this expectation value into the $\hat{1}$ direction, and rotations about $\hat{3}$ leave $\langle\phi\rangle$ invariant. Thus the gauge generators T^1 and T^2 are spontaneously broken, and the scalar field components ϕ^2 and ϕ^1 are the corresponding Goldstone bosons. This accords with the result of computing the elements of $(T^a\phi_0)_i$ explicitly: Using $(T^a)_{bc} = \epsilon^{bac}$, we find

$$(T^a\phi_0)_b = \epsilon_{bac}\langle\phi^c\rangle = V\epsilon^{ba3}. \tag{20.52}$$

‡A special case of this argument appeared in the discussion of Eq. (19.88).

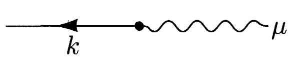
Inserting this result into formula (20.50), we see that the current of each spontaneously broken symmetry creates and destroys its own Goldstone boson.

Now we can use this formalism to study the working of the Higgs mechanism in this general context. Consider the original theory $\mathcal{L}_0$ coupled to gauge bosons of G. To see how the Higgs mechanism operates, we must compute the vacuum polarization amplitude. This amplitude is required by the Ward identity to be transverse, so it is necessarily of the form



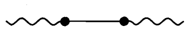
$$= i\left(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\right) \cdot (m_{ab}^2 + \mathcal{O}(k^2)). \qquad (20.53)$$

It is not easy to compute the nonsingular terms in (20.53) in this general situation, but it is straightforward to compute the singular term, which comes from contributions with an intermediate Goldstone boson. Combining Eqs. (20.45) and (20.46), we see that the amplitude for a gauge boson to convert to a Goldstone boson is



$$= -gk^\mu F^a{}_j. \qquad (20.54)$$

Then the pole contribution to the vacuum polarization is



$$= (gk^\mu F^a{}_j)\frac{i}{k^2}(-gk^\nu F^b{}_j). \qquad (20.55)$$

Comparing (20.55) with (20.21), we identify

$$m_{ab}^2 = g^2 F^a{}_j F^b{}_j. \qquad (20.56)$$

Notice that, in the case in which the symmetry is broken by a scalar field, this result reverts to (20.18). However, Eq. (20.56) applies to any theory of spontaneously broken symmetry, whether the symmetry breaking is apparent from the Lagrangian or whether it requires strong interactions or other nonperturbative effects. It is a general result, then, that any gauge boson coupled to the current of a spontaneously broken symmetry acquires a mass.

20.2 The Glashow-Weinberg-Salam Theory of Weak Interactions

We are now ready to write down the spontaneously broken gauge theory that gives the experimentally correct description of the weak interactions, a model introduced by Glashow, Weinberg, and Salam (GWS). Like the second $SU(2)$ model considered in the previous section, this model gives a unified description of weak and electromagnetic interactions, in which the massless photon corresponds to a particular combination of symmetry generators that remains unbroken.

Again we begin with a theory with $SU(2)$ gauge symmetry. To break the symmetry spontaneously, we introduce a scalar field in the spinor representation of $SU(2)$, as in Eq. (20.22). However, we know that this theory leads to a system with no massless gauge bosons. We therefore introduce an additional $U(1)$ gauge symmetry. We assign the scalar field a charge $+1/2$ under this $U(1)$ symmetry, so that its complete gauge transformation is

$$\phi \to e^{i\alpha^a \tau^a} e^{i\beta/2} \phi. \tag{20.57}$$

(Here $\tau^a = \sigma^a/2$.) If the field ϕ acquires a vacuum expectation value of the form

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{20.58}$$

then a gauge transformation with

$$\alpha^1 = \alpha^2 = 0, \qquad \alpha^3 = \beta \tag{20.59}$$

leaves $\langle \phi \rangle$ invariant. Thus, the theory will contain one massless gauge boson, corresponding to this particular combination of generators. The remaining three gauge bosons will acquire masses from the Higgs mechanism.

Gauge Boson Masses

It is straightforward to work out the details of the mass spectrum by using the methods of the previous section. The covariant derivative of ϕ is

$$D_\mu \phi = \left(\partial_\mu - igA^a_\mu \tau^a - i\tfrac{1}{2} g' B_\mu\right)\phi, \tag{20.60}$$

where A^a_μ and B_μ are, respectively, the $SU(2)$ and $U(1)$ gauge bosons. Since the $SU(2)$ and $U(1)$ factors of the gauge group commute with one another, they can have different coupling constants, which we have called g and g'.

The gauge boson mass terms come from the square of Eq. (20.60), evaluated at the scalar field vacuum expectation value (20.58). The relevant terms are

$$\Delta\mathcal{L} = \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} \left(gA^a_\mu \tau^a + \frac{1}{2} g' B_\mu\right)\left(gA^{b\mu} \tau^b + \frac{1}{2} g' B^\mu\right) \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{20.61}$$

If we evaluate the matrix product explicitly, using $\tau^a = \sigma^a/2$, we find

$$\Delta\mathcal{L} = \frac{1}{2} \frac{v^2}{4} \left[g^2 (A^1_\mu)^2 + g^2 (A^2_\mu)^2 + (-gA^3_\mu + g' B_\mu)^2\right]. \tag{20.62}$$

There are three massive vector bosons, which we will notate as follows:

$$\begin{aligned} W^\pm_\mu &= \frac{1}{\sqrt{2}}(A^1_\mu \mp iA^2_\mu) \qquad \text{with mass} \quad m_W = g\frac{v}{2}; \\ Z^0_\mu &= \frac{1}{\sqrt{g^2 + g'^2}}(gA^3_\mu - g' B_\mu) \qquad \text{with mass} \quad m_Z = \sqrt{g^2 + g'^2}\,\frac{v}{2}. \end{aligned} \tag{20.63}$$

The fourth vector field, orthogonal to Z^0_μ, remains massless:

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(g' A^3_\mu + gB_\mu) \qquad \text{with mass} \quad m_A = 0. \tag{20.64}$$

We will identify this field with the electromagnetic vector potential.

From now on it will be more convenient to write all expressions in terms of these mass eigenstate fields. Consider, for instance, the coupling of the vector fields to fermions. For a fermion field belonging to a general $SU(2)$ representation, with $U(1)$ charge Y, the covariant derivative takes the form

$$D_\mu = \partial_\mu - igA_\mu^a T^a - ig'YB_\mu. \tag{20.65}$$

In terms of the mass eigenstate fields, this becomes

$$\begin{aligned} D_\mu = \partial_\mu - i\frac{g}{\sqrt{2}}\left(W_\mu^+ T^+ + W_\mu^- T^-\right) - i\frac{1}{\sqrt{g^2+g'^2}} Z_\mu \left(g^2 T^3 - g'^2 Y\right) \\ - i\frac{gg'}{\sqrt{g^2+g'^2}} A_\mu \left(T^3 + Y\right), \end{aligned} \tag{20.66}$$

where $T^\pm = (T^1 \pm iT^2)$. The normalization is chosen so that, in the spinor representation of $SU(2)$,

$$T^\pm = \tfrac{1}{2}(\sigma^1 \pm i\sigma^2) = \sigma^\pm. \tag{20.67}$$

The last term of Eq. (20.66) makes explicit the fact that the massless gauge boson A_μ couples to the gauge generator $(T^3 + Y)$, which generates precisely the symmetry operation (20.59).

To put expression (20.66) into a more useful form, we should identify the coefficient of the electromagnetic interaction as the electron charge e,

$$e = \frac{gg'}{\sqrt{g^2+g'^2}}, \tag{20.68}$$

and identify the electric charge quantum number as

$$Q = T^3 + Y. \tag{20.69}$$

These substitutions, with $Q = -1$ for the electron, give the conventional form of the coupling of the electromagnetic field.

To simplify expression (20.66) further, we define the *weak mixing angle*, θ_w, to be the angle that appears in the change of basis from (A^3, B) to (Z^0, A):

$$\begin{pmatrix} Z^0 \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta_w & -\sin\theta_w \\ \sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} A^3 \\ B \end{pmatrix},$$

that is,

$$\cos\theta_w = \frac{g}{\sqrt{g^2+g'^2}}, \qquad \sin\theta_w = \frac{g'}{\sqrt{g^2+g'^2}}. \tag{20.70}$$

Then, with the manipulation in the Z^0 coupling

$$g^2 T^3 - g'^2 Y = (g^2 + g'^2)T^3 - g'^2 Q,$$

we can rewrite the covariant derivative (20.66) in the form

$$D_\mu = \partial_\mu - i\frac{g}{\sqrt{2}}(W_\mu^+ T^+ + W_\mu^- T^-) - i\frac{g}{\cos\theta_w} Z_\mu (T^3 - \sin^2\theta_w Q) - ieA_\mu Q, \tag{20.71}$$

where

$$g = \frac{e}{\sin\theta_w}. \tag{20.72}$$

We see here that the couplings of all of the weak bosons are described by two parameters: the well-measured electron charge e, and a new parameter θ_w. The couplings induced by W and Z exchange will also involve the masses of these particles. However, these masses are not independent, since it follows from Eqs. (20.63) that

$$m_W = m_Z \cos\theta_w. \tag{20.73}$$

All effects of W and Z exchange processes, at least at tree level, can be written in terms of the three basic parameters e, θ_w, and m_W.

Coupling to Fermions

The covariant derivative (20.71) uniquely determines the coupling of the W and Z^0 fields to fermions, once the quantum numbers of the fermion fields are specified. To determine these quantum numbers, we must take account of the fact, mentioned in Section 17.3, that the W boson couples only to left-handed helicity states of quarks and leptons.

At the level of the classical Lagrangian, there is no difficulty in constructing theories in which the left- and right-handed components of a fermion field couple differently to gauge bosons.* Already in Section 3.2 we saw that the kinetic energy term for Dirac fermions splits into separate pieces for the left- and right-handed fields:

$$\bar{\psi} i\not{\partial}\psi = \bar{\psi}_L i\not{\partial}\psi_L + \bar{\psi}_R i\not{\partial}\psi_R. \tag{20.74}$$

When we couple ψ to a gauge field, we can assign ψ_L and ψ_R to different representations of the gauge group. Then the two terms on the right-hand side of (20.74) will contain two different covariant derivatives, and these will imply two different sets of couplings.

In the GWS model, we can use this technique to insure that only the left-handed components of the quark and lepton fields couple to the W bosons. We assign the left-handed fermion fields to doublets of $SU(2)$, while making the right-handed fermion fields singlets under this group. Once we have specified the T^3 value for each fermion field, the value of Y that we must assign follows from Eq. (20.69). This means that the Y assignments will also be different for

*In Section 19.4, we argued that there is a possible problem with this strategy at the level of quantum corrections. We will check below whether the specific model we construct avoids this problem.

the left- and right-handed components of quarks and leptons. For the right-handed fields, $T^3 = 0$, and so we reproduce the standard electric charges by assigning Y to equal the electric charge. For example, for the right-handed u quark field, $Y = +2/3$; for e_R^-, $Y = -1$. For the left-handed fields,

$$E_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \qquad Q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \tag{20.75}$$

the assignments $Y = -1/2$ and $Y = +1/6$, respectively, combine with $T^3 = \pm 1/2$ to give the correct electric charge assignments. Since the left- and right-handed fermions live in different representations of the fundamental gauge group, it is often useful to think of these components as distinct particles, which are mixed by the fermion mass terms.

In fact, the construction of fermion mass terms is a serious problem, because all possible such terms are forbidden by global gauge invariances. For example, we cannot write an electron mass term

$$\Delta\mathcal{L} = -m_e(\bar{e}_L e_R + \bar{e}_R e_L), \tag{20.76}$$

because the fields e_L and e_R belong to different $SU(2)$ representations and have different $U(1)$ charges. For the next few pages, we will ignore this problem by treating all fermion fields as massless. This description will suffice to analyze the structure of the weak interactions at high energies, where the quark and lepton masses can be ignored. At the end of this section we will return to the problem of writing quark and lepton mass terms in the GWS theory. The solution to this problem will reinforce the idea that the left- and right-handed fermion fields are fundamentally independent entities, mixed to form massive fermions by some subsidiary process.

If we ignore fermion masses, the Lagrangian for the weak interactions of quarks and leptons follows directly from the charge assignments given above. The fermion kinetic energy terms for e, ν, u, and d are

$$\mathcal{L} = \overline{E}_L(i\not{D})E_L + \bar{e}_R(i\not{D})e_R + \overline{Q}_L(i\not{D})Q_L + \bar{u}_R(i\not{D})u_R + \bar{d}_R(i\not{D})d_R. \tag{20.77}$$

In each term, the covariant derivative is given by Eq. (20.65), with T^a and Y evaluated in the particular representation to which that fermion field belongs. For example,

$$\overline{Q}_L(i\not{D})Q_L = \overline{Q}_L i\gamma^\mu(\partial_\mu - igA_\mu^a\tau^a - i\tfrac{1}{6}g'B_\mu)Q_L. \tag{20.78}$$

A right-handed neutrino would have zero coupling both to $SU(2)$ and to $U(1)$, so we have simply omitted this field from Eq. (20.77).

To work out the physical consequences of the fermion-vector boson couplings, we should write Eq. (20.77) in terms of the vector boson mass eigenstates, using the form of the covariant derivative given in Eq. (20.71). Equation (20.77) then takes the form

$$\begin{aligned}\mathcal{L} = \overline{E}_L(i\not{\partial})E_L + \bar{e}_R(i\not{\partial})e_R + \overline{Q}_L(i\not{\partial})Q_L + \bar{u}_R(i\not{\partial})u_R + \bar{d}_R(i\not{\partial})d_R \\ + g(W_\mu^+ J_W^{\mu+} + W_\mu^- J_W^{\mu-} + Z_\mu^0 J_Z^\mu) + eA_\mu J_{EM}^\mu,\end{aligned} \tag{20.79}$$

where

$$\begin{aligned}
J_W^{\mu+} &= \frac{1}{\sqrt{2}}\left(\bar{\nu}_L\gamma^\mu e_L + \bar{u}_L\gamma^\mu d_L\right);\\
J_W^{\mu-} &= \frac{1}{\sqrt{2}}\left(\bar{e}_L\gamma^\mu \nu_L + \bar{d}_L\gamma^\mu u_L\right);\\
J_Z^\mu &= \frac{1}{\cos\theta_w}\Big[\bar{\nu}_L\gamma^\mu\left(\tfrac{1}{2}\right)\nu_L + \bar{e}_L\gamma^\mu\left(-\tfrac{1}{2}+\sin^2\theta_w\right)e_L + \bar{e}_R\gamma^\mu\left(\sin^2\theta_w\right)e_R\\
&\qquad + \bar{u}_L\gamma^\mu\left(\tfrac{1}{2}-\tfrac{2}{3}\sin^2\theta_w\right)u_L + \bar{u}_R\gamma^\mu\left(-\tfrac{2}{3}\sin^2\theta_w\right)u_R\\
&\qquad + \bar{d}_L\gamma^\mu\left(-\tfrac{1}{2}+\tfrac{1}{3}\sin^2\theta_w\right)d_L + \bar{d}_R\gamma^\mu\left(\tfrac{1}{3}\sin^2\theta_w\right)d_R\Big];\\
J_{EM}^\mu &= \bar{e}\gamma^\mu(-1)e + \bar{u}\gamma^\mu\left(+\tfrac{2}{3}\right)u + \bar{d}\gamma^\mu\left(-\tfrac{1}{3}\right)d.
\end{aligned} \tag{20.80}$$

Here we have used Eq. (20.67) to simplify the W boson currents. Notice that the current J_{EM}^μ associated with the photon field is indeed the standard electromagnetic current.

Anomaly Cancellation

As we have just seen, there is no difficulty in writing a Lagrangian that couples the GWS gauge bosons to fermions in a chiral fashion. However, these chiral couplings do present a potential problem that appears at the level of one-loop corrections. In Section 19.2, we saw that an axial current that is conserved at the level of the classical equations of motion can acquire a nonzero divergence through one-loop diagrams that couple this current to a pair of gauge bosons. The Feynman diagram that contains this anomalous contribution is a triangle diagram with the axial current and the two gauge currents at its vertices. In a gauge theory in which gauge bosons couple to a chiral current, the dangerous triangle diagrams appear in the one-loop corrections to the three-gauge-boson vertex function. The anomalous terms violate the Ward identity for this amplitude. Thus, as we argued in Section 19.4, theories in which gauge bosons couple to chiral currents can be gauge invariant only if the anomalous contribution somehow disappears. Fortunately, as we saw there, the anomalous terms can be arranged to cancel when one sums over all possible fermion species that can circulate in these diagrams.†

Within the GWS theory, the requirement from experiment that the weak interaction currents are left-handed forced us to choose a chiral gauge coupling. Now we must check that the anomalous terms from the triangle diagrams cancel as required. We will find that they do, but only through a subtle and rather magical interplay of the quantum numbers of quarks and leptons.

The anomalous term of a triangle diagram of three gauge bosons A_μ^a, A_ν^b,

†If you have not read Chapter 19, but you are willing to assume that the fermion triangle diagram contains a contribution that violates gauge invariance, you should still be able to follow the argument that follows.

and A^c_λ is proportional to the group theoretic invariant

$$\mathrm{tr}\left[\gamma^5 t^a \{t^b, t^c\}\right], \tag{20.81}$$

where the trace is taken over all fermion species. The anticommutator comes from taking the sum of two triangle diagrams in which the fermions circle in opposite directions. The factor γ^5 registers the fact that the anomaly is associated with chiral currents; this factor equals -1 for left-handed fermions and $+1$ for right-handed fermions. In theories such as QED or QCD in which the gauge bosons couple equally to right- and left-handed species, the anomalies automatically cancel. This bookkeeping method is a special case of the more general method presented in Section 19.4.

To evaluate the anomalies of the GWS theory, it is easiest to work in the basis of $SU(2) \times U(1)$ gauge bosons, before the mixing to the photon and Z^0 mass eigenstates. It suffices to evaluate the triangle diagrams for massless fermions, so that right- and left-handed fermions have distinct quantum numbers. However, we must consider not only the anomalies of diagrams with three $SU(2) \times U(1)$ gauge bosons, but also diagrams with both weak-interaction gauge bosons and color $SU(3)$ gauge bosons of QCD. If we consider effects of gravity on the weak-interaction gauge theory, there is also a possibly anomalous diagram with one weak-interaction gauge boson and two gravitons. We can omit diagrams, such as the anomaly of three $SU(3)$ bosons or of one $SU(3)$ boson and two gravitons, in which all of the couplings are left-right symmetric. Then the full set of diagrams with possible anomalous terms is shown in Fig. 20.2. All of the possible anomalies must cancel if the Ward identities of the $SU(2) \times U(1)$ gauge theory are to be satisfied.

It is a special property of $SU(2)$ gauge theory that the anomaly of three $SU(2)$ gauge bosons always vanishes; this result follows from the property of Pauli sigma matrices $\{\sigma^a, \sigma^b\} = 2\delta^{ab}$, which implies that the trace (20.81) vanishes. The anomalies containing one $SU(3)$ boson or one $SU(2)$ boson are proportional to

$$\mathrm{tr}[t^a] = 0 \qquad \text{or} \qquad \mathrm{tr}[\tau^a] = 0. \tag{20.82}$$

The remaining nontrivial anomalies are those of one $U(1)$ boson with two $SU(2)$ bosons or two $SU(3)$ bosons, the anomaly of three $U(1)$ bosons, and the gravitational anomaly with one $U(1)$ gauge boson.

The anomaly of one $U(1)$ boson with two $SU(3)$ bosons is proportional to the group theory factor

$$\mathrm{tr}[t^a t^b Y] = \tfrac{1}{2}\delta^{ab} \cdot \sum_q Y_q, \tag{20.83}$$

where the sum runs over left-handed quarks and right-handed quarks, with an extra (-1) for the left-handed contributions. Inserting the charge assignments given above for u_L, d_L, u_R, and d_R, we find

$$\sum_q Y_q = -2 \cdot \tfrac{1}{6} + (\tfrac{2}{3}) + (-\tfrac{1}{3}) = 0. \tag{20.84}$$

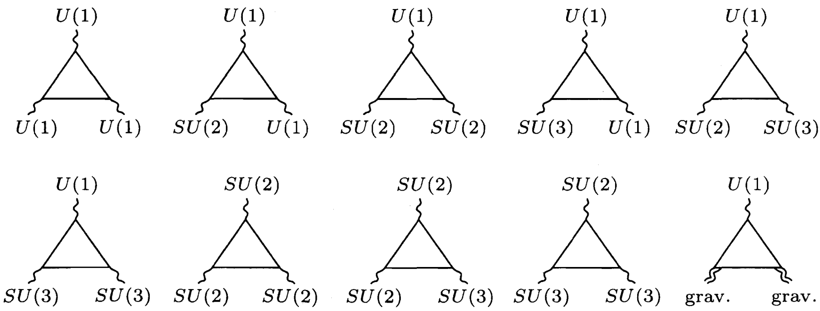


Figure 20.2. Possible gauge anomalies of weak interaction theory. All of these anomalies must vanish for the Glashow-Weinberg-Salam theory to be consistent.

Similarly, the anomaly of a $U(1)$ boson with two $SU(2)$ bosons is proportional to

$$\mathrm{tr}[\tau^a \tau^b Y] = \tfrac{1}{2}\delta^{ab} \sum_{fL} Y_{fL}, \tag{20.85}$$

where the sum runs over the left-handed fermions E_L and Q_L. Thus,

$$\sum_{fL} Y_{fL} = -(-\tfrac{1}{2}) - 3 \cdot \tfrac{1}{6} = 0; \tag{20.86}$$

the factor 3 counts the color states of the quarks. The anomaly of three $U(1)$ gauge bosons is proportional to a sum involving left- and right-handed leptons and quarks:

$$\mathrm{tr}[Y^3] = -2(-\tfrac{1}{2})^3 + (-1)^3 - 3\big[2(\tfrac{1}{6})^3 - (\tfrac{2}{3})^3 - (-\tfrac{1}{3})^3\big] = 0. \tag{20.87}$$

Finally, the gravitational anomaly with one $U(1)$ gauge boson is proportional to

$$\mathrm{tr}[Y] = -2(-\tfrac{1}{2}) + (-1) - 3\big[2(\tfrac{1}{6}) - (\tfrac{2}{3}) - (-\tfrac{1}{3})\big] = 0. \tag{20.88}$$

The Glashow-Weinberg-Salam theory is thus a chiral gauge theory that is completely free of axial vector anomalies among the gauge currents. However, the cancellation of anomalies requires that leptons and quarks appear in complete multiplets with the structure of $(E_L, e_R, Q_L, u_R, d_R)$. This set of fields is often called a *generation* of quarks and leptons. The consistency of the theory requires that quarks and leptons appear in Nature in equal numbers, organizing themselves into successive generations in this way.

Experimental Consequences of the GWS Theory

Now that we have a fundamental theory for the coupling of W and Z bosons to fermions, we can work out the consequences of this theory for observable processes mediated by weak bosons. This analysis should reproduce the effective Lagrangian description of the weak interactions used in Chapters 17

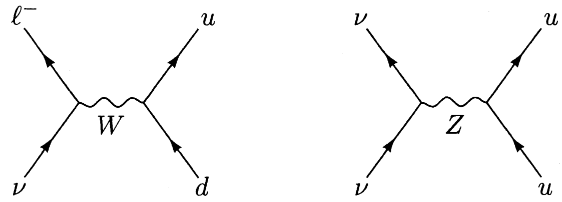


Figure 20.3. Some processes with virtual W and Z boson exchange.

and 18, and also predict additional observable effects of weak boson exchange. In our discussion here, we will derive only the most basic relations in this subject; we do not have space for a systematic survey of the phenomenology of weak interactions. However, we encourage the reader to study the experimental foundations of the weak interactions, which contain many beautiful illustrations of the principles of quantum field theory.[‡]

At energies low compared to the vector boson masses, the couplings of the weak bosons have their major effects through processes that involve virtual weak boson exchange. These processes are shown in Fig. 20.3. We will derive the Feynman rules for massive gauge bosons in Chapter 21. Meanwhile, it is reasonable to guess that the W and Z boson propagators are given by

$$\big\langle W^{\mu+}(p)W^{\nu-}(-p)\big\rangle = \frac{-ig^{\mu\nu}}{p^2 - m_W^2}, \qquad \langle Z^\mu(p)Z^\nu(-p)\rangle = \frac{-ig^{\mu\nu}}{p^2 - m_Z^2}. \tag{20.89}$$

We will see in Section 21.1 that these propagators give correct expressions for diagrams with W and Z exchange up to terms of order (m_f/m_W), where m_f is a fermion mass.

First consider the W exchange diagram in Fig. 20.3, in the limit of energies low compared to the W mass. We can then neglect the p^2 term in the denominator of the W propagator (20.89). Taking the W coupling from Eq. (20.79), we find that the diagram can be described by the effective Lagrangian

$$\begin{aligned}\Delta\mathcal{L}_W &= \frac{g^2}{m_W^2} J_W^{\mu-} J_{\mu W}^{+} \\ &= \frac{g^2}{2m_W^2}\big(\bar{e}_L\gamma^\mu\nu_L + \bar{d}_L\gamma^\mu u_L\big)\big(\bar{\nu}_L\gamma_\mu e_L + \bar{u}_L\gamma_\mu d_L\big).\end{aligned} \tag{20.90}$$

The coefficient is often written in terms of the *Fermi constant*,

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2}. \tag{20.91}$$

The various terms in this effective Lagrangian reproduce the expressions we have already written in Eqs. (17.31), (18.28), and (18.29). Since these interactions among leptons and quarks are mediated by the exchange of an

---

[‡]The experimental successes of the theory of weak interactions are reviewed in the book of Commins and Bucksbaum (1983).

electrically charged vector boson, they are called collectively *charged-current* interactions. The effective Lagrangian (20.90) turns out to provide an impressively successful description of the phenomenology of charged-current weak interactions. We have described its use in high-energy neutrino scattering, but it has comparable successes in nuclear β-decay, muon decay, and a variety of other processes.

In a similar way, we can work out the effective Lagrangian resulting from virtual Z^0 exchange. We find

$$\begin{aligned}\Delta\mathcal{L}_Z &= \frac{g^2}{2m_Z^2} J_Z^\mu J_{\mu Z} \\ &= \frac{4G_F}{\sqrt{2}}\left(\sum_f \bar{f}\gamma^\mu(T^3 - \sin^2\theta_w Q)f\right)^2,\end{aligned} \tag{20.92}$$

where the sum in the second line runs over all left-handed and right-handed flavors, and we have used relation (20.73) to simplify the prefactor. We say that the effective Lagrangian (20.92) mediates *neutral current* weak interaction processes. Notice that, if we define $SU(2)$ gauge currents as

$$J^{\mu a} = \sum_f \bar{f}\gamma^\mu T^a f, \tag{20.93}$$

then the effective Lagrangians of W and Z exchange can be written together in the simple form

$$\Delta\mathcal{L}_W + \Delta\mathcal{L}_Z = \frac{4G_F}{\sqrt{2}}\left[(J^{\mu 1})^2 + (J^{\mu 2})^2 + (J^{\mu 3} - \sin^2\theta_w J^\mu_{EM})^2\right]. \tag{20.94}$$

This expression becomes manifestly invariant under an unbroken global $SU(2)$ symmetry in the limit $g' \to 0$ or $\sin^2\theta_w \to 0$. We will discuss this observation further at the end of this section.

The neutral current effective Lagrangian (20.92) contains terms that couple together all of the various species of quarks and leptons. These terms violate parity, and so distinguish themselves from the effects of strong and electromagnetic interactions. For example, Eq. (20.92) predicts the existence of neutral current deep inelastic neutrino scattering events, in which a high-energy neutrino shatters a nucleon but does not convert to a final-state muon or electron. This process is analyzed in Problem 20.4. Similarly, the neutral-current interaction predicts the presence of parity-violating effects in electron deep inelastic scattering. It also predicts a parity-violating electron-nucleon interaction that should mix atomic energy levels, and a similar parity-violating nucleon-nucleon interaction. Within the GWS theory, the strengths of all of these various effects are predicted in terms of the Fermi constant and one additional parameter, the value of $\sin^2\theta_w$. Thus, the GWS theory can be tested by observing each of these effects and asking whether a single value of this parameter can account for the strengths of all of these disparate processes.

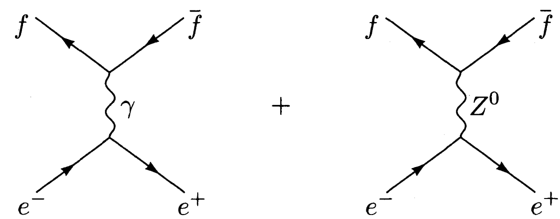


Figure 20.4. Diagrams contributing to the process $e^+e^- \to f\bar{f}$ in the Glashow-Weinberg-Salam theory.

Further tests of the GWS theory are available at higher energies. The process $e^+e^- \to f\bar{f}$ is affected in an essential way, since the theory contains a new diagram with s-channel Z^0 exchange, which interferes with the standard photon exchange diagram, as shown in Fig. 20.4. It is straightforward to work out the effects of this interference using the methods of Section 5.2, so we have left this analysis as Problem 20.3.

As the center-of-mass energy approaches m_Z, the Z^0 appears directly as a resonance in the e^+e^- annihilation cross section. Similarly, both the W and the Z can be observed as resonances in quark-antiquark annihilation, viewed as a parton subprocess in proton-antiproton scattering. The positions of these resonances are predicted from G_F, $\sin^2\theta_w$, and the value of e or α, according to Eqs. (20.72) and (20.91). Using these relations, we find

$$m_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2\theta_w}, \qquad m_Z^2 = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2\theta_w \cos^2\theta_w}. \tag{20.95}$$

The detailed shape of the Z^0 resonance is shown in Fig. 20.5. The experimental measurements shown are compared to a theoretical curve with the resonance position adjusted for the best fit. The height and width of the resonance are then predicted by the GWS theory. The resonance is broadened to higher energies by processes in which the electron and positron radiate collinear photons before annihilation; this correction was discussed in Problem 5.5.

Because the Lagrangian of the GWS theory treats left- and right-handed fermions as distinct species with completely different quantum numbers, the couplings of the Z^0 to left- and right-handed fermions differ signficantly. One manifestation of this is the presence of a *polarization asymmetry*, a net polarization of fermions produced in the decay $Z^0 \to f\bar{f}$, or an asymmetry in the inverse process of Z^0 production. This asymmetry can be read directly from the form of the Z^0 current given in (20.80):

$$\begin{aligned} A_{LR}^f &= \frac{\Gamma(Z^0 \to f_L\bar{f}_R) - \Gamma(Z^0 \to f_R\bar{f}_L)}{\Gamma(Z^0 \to f_L\bar{f}_R) + \Gamma(Z^0 \to f_R\bar{f}_L)} \\ &= \frac{(\frac{1}{2} - |Q_f|\sin^2\theta_w)^2 - (Q_f\sin^2\theta_w)^2}{(\frac{1}{2} - |Q_f|\sin^2\theta_w)^2 + (Q_f\sin^2\theta_w)^2}. \end{aligned} \tag{20.96}$$

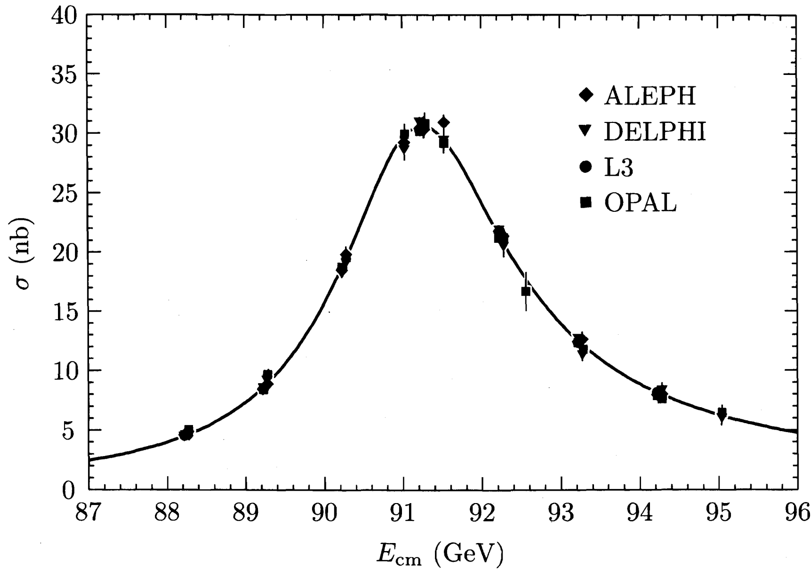


Figure 20.5. The total cross section for e^+e^- annihilation to hadrons for $E_{\rm cm}$ close to the Z^0 boson mass, as measured by the ALEPH, DELPHI, L3, and OPAL experiments and compiled by the Particle Data Group, *Phys. Rev.* **D50**, (1994), Fig. 32.14. References to the original articles are given there. The solid curve is the prediction of the GWS theory.

For a realistic value $\sin^2\theta_w = 0.23$, this expression gives a 15% asymmetry for charged leptons and a 95% asymmetry for d, s, and b quarks. The asymmetry can be checked experimentally for leptons by measuring the polarization of τ leptons at the Z^0 resonance, or by measuring the relative cross sections for producing the resonance using left- versus right-handed electrons. For quarks, the asymmetry can be determined indirectly from the forward-backward production asymmetry on the resonance, as explained in Problem 20.3.

Because the weak neutral current has so many different manifestations, the GWS theory of weak interactions can be subjected to a stringent test by comparing the values of the parameter $\sin^2\theta_w$ needed to account for each of its predicted effects. Table 20.1 presents the values of $\sin^2\theta_w$ extracted from a wide variety of weak interaction neutral current effects and asymmetries. In all cases, one-loop radiative corrections must be included to analyze the experiment at the required level of accuracy. These radiative corrections involve some subtlety. First, one must adopt a specific renormalization convention that defines $\sin^2\theta_w$ and use it consistently in all calculations. The table shows results for two different choices of this convention. In both conventions, the values of weak-interaction observables are taken to be functions of α, G_F, and a third independent parameter. In the first column this parameter is the mass ratio m_W/m_Z, and, following the tree-level expression (20.73), we consider

Table 20.1. Values of $\sin^2\theta_w$ from Weak Interaction Experiments

Observed Quantity or Process	s_W^2	$\sin^2\theta_{w\,\overline{MS}}$
m_Z	0.2247 (21)	0.2320 (6)
m_W	0.2264 (25)	0.2338 (22)
Γ_Z	0.2250(18)	0.2322 (6)
Lepton f-b asymmetries at the Z^0	0.2243 (17)	0.2315 (11)
All pair-production asymmetries at the Z^0	0.2245 (17)	0.2317 (8)
A_{LR}^e at the Z^0	0.2221 (17)	0.2292 (10)
Deep inelastic neutrino scattering	0.2260 (48)	0.233 (5)
Neutrino-proton elastic scattering	0.205 (31)	0.212 (32)
Neutrino-electron elastic scattering	0.224 (9)	0.231 (9)
Atomic parity violation	0.216 (8)	0.223 (8)
Parity violation in inelastic e^- scattering	0.216 (17)	0.223 (18)

The values listed here are obtained by fitting experimental observations by adjusting the value of s_W^2 or $\sin^2\theta_{w\,\overline{MS}}$, taking α and G_F as accurately known parameters. The numbers in parentheses are the standard errors in the last displayed digits. The conversion from the experimentally measured quantities to s_W^2 or $\sin^2\theta_{w\,\overline{MS}}$ depends on the value of the top quark mass and the mass of the Higgs boson. These values assume a top quark mass of 169 GeV and a Higgs mass of 300 GeV; the quoted errors include an uncertainty of 17 GeV in the top quark mass and a range from 60 GeV to 1000 GeV for the Higgs mass. The differences in the relative errors between the two columns reflect the importance of this theoretical uncertainty. Some observables depend weakly on α_s; these values assume $\alpha_s(m_Z) = 0.120 \pm .007$. This table is taken from the article of P. Langacker and J. Erler for the Particle Data Group, *Phys. Rev.* **D50**, 1304 (1994). That article contains a full set of references and a discussion of the sources of uncertainty in these determinations.

this ratio to define a renormalized value of $\sin^2\theta_w$:

$$s_W^2 \equiv 1 - \frac{m_W^2}{m_Z^2}. \tag{20.97}$$

In the second column, the third parameter is $\sin^2\theta_w$ computed from the weak interaction coupling constants defined by minimal subtraction (Eq. (11.77)). The differences between different definitions of $\sin^2\theta_w$ appear at the level of one-loop computations and can reveal interesting physics; this subject is discussed in Section 21.3.

A second subtlety is that the one-loop corrections to weak neutral current processes depend on the value of the t quark mass, which has only recently been determined and is still somewhat poorly known. The dependence on the t quark mass is relatively strong, for interesting reasons that we will discuss in Section 21.3. The one-loop corrections also depend weakly on properties of the particles responsible for the spontaneous symmetry breaking.

We can see from Table 20.1 that a wide variety of effects due to the weak neutral current have been observed, with magnitudes accounted for by a single, consistent value of $\sin^2\theta_w$. This remarkable concordance of theory and experiment gives us confidence that the Glashow-Weinberg-Salam theory is indeed the correct description of weak and electromagnetic interactions.

Fermion Mass Terms

We now return to the problem of writing mass terms for the quarks and leptons. Recall that one cannot put ordinary mass terms into the Lagrangian, because the left- and right-handed components of the various fermion fields have different gauge quantum numbers and so simple mass terms violate gauge invariance. To give masses to the quarks and leptons, we must again invoke the mechanism of spontaneous symmetry breaking.

We began this section by assuming that a scalar field ϕ acquires a vacuum expectation value (20.58), in order to give mass to the W and Z bosons. This scalar field needed to be a spinor under $SU(2)$ and to have $Y = 1/2$ in order to produce the correct pattern of gauge boson masses. With these quantum numbers, we can also write a gauge-invariant coupling linking e_L, e_R, and ϕ, as follows:

$$\Delta\mathcal{L}_e = -\lambda_e \overline{E}_L \cdot \phi\, e_R + \text{h.c.} \tag{20.98}$$

Here the $SU(2)$ indices of the doublets E_L and ϕ are contracted; notice also that the charges Y of the various fields sum to zero. The parameter λ_e is a new dimensionless coupling constant. If we replace ϕ in this expression by its vacuum expectation value (20.58), we obtain

$$\Delta\mathcal{L}_e = -\frac{1}{\sqrt{2}}\lambda_e v\, \bar{e}_L e_R + \text{h.c.} + \cdots . \tag{20.99}$$

This is a mass term for the electron. The size of the mass is set by the vacuum expectation value of ϕ, rescaled by the new dimensionless coupling:

$$m_e = \frac{1}{\sqrt{2}}\lambda_e v. \tag{20.100}$$

Since the electron mass is proportional to v, one might expect that the masses of the electron and the W boson should be of the same order. In fact, taking the observed values, $m_e/m_W \sim 6\times 10^{-6}$. Since λ_e is a renormalizable coupling, it must be treated as an input to the theory. Thus the GWS theory allows the electron to be very light, but it cannot explain why the electron is so light compared to the W boson.

We can write mass terms for the quark fields in the same way. Notice that, in the following expression, both terms are invariant under $SU(2)$ and have zero net Y:

$$\Delta\mathcal{L}_q = -\lambda_d \overline{Q}_L \cdot \phi\, d_R - \lambda_u \epsilon^{ab} \overline{Q}_{La} \phi_b^\dagger u_R + \text{h.c.} \tag{20.101}$$

Substituting the vacuum expectation value of ϕ from Eq. (20.58), these terms become

$$\Delta\mathcal{L}_q = -\frac{1}{\sqrt{2}}\lambda_d v\, \bar{d}_L d_R - \frac{1}{\sqrt{2}}\lambda_u v\, \bar{u}_L u_R + \text{h.c.} + \cdots, \tag{20.102}$$

standard mass terms for the d and u quarks. The GWS theory thus gives the relations

$$m_d = \frac{1}{\sqrt{2}}\lambda_d v, \qquad m_u = \frac{1}{\sqrt{2}}\lambda_u v. \tag{20.103}$$

As with the electron, the theory parametrizes but does not explain the small values of the d and u quark masses observed experimentally.

When additional generations of quarks are introduced into the theory, there can be additional coupling terms that mix generations. Alternatively, we can diagonalize the Higgs couplings by choosing a new basis for the quark fields. We will show that this is always possible in Section 20.3. However, this simplification of the Higgs couplings causes a complication in the gauge couplings. Let

$$u_L^i = (u_L, c_L, t_L)\ , \qquad d_L^i = (d_L, s_L, b_L) \tag{20.104}$$

denote the up- and down-type quarks in their original basis, and let $u_L'^i$ and $d_L'^i$ denote the quarks in the basis that diagonalizes their Higgs couplings. This latter basis is the physical one, since it is the basis that diagonalizes the mass matrix. The two bases are related by unitary transformations:

$$u_L^i = U_u^{ij} u_L'^j, \qquad d_L^i = U_d^{ij} d_L'^j. \tag{20.105}$$

In this new basis, the W boson current takes the form

$$J_W^{\mu +} = \frac{1}{\sqrt{2}}\bar{u}_L^i \gamma^\mu d_L^i = \frac{1}{\sqrt{2}}\bar{u}_L'^i \gamma^\mu (U_u^\dagger U_d)_{ij} d_L'^j. \tag{20.106}$$

This expression is conventionally written

$$J_W^{\mu +} = \frac{1}{\sqrt{2}}\bar{u}_L'^i \gamma^\mu V_{ij} d_L'^j, \tag{20.107}$$

where V_{ij} is a unitary matrix called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The off-diagonal terms in V_{ij} allow weak-interaction transitions between quark generations. For example, restricting to two generations for simplicity and writing

$$V_{1j} d_L'^j = \cos\theta_c d_L' + \sin\theta_c s_L', \tag{20.108}$$

the term proportional to $\sin\theta_c$ allows an s quark to decay weakly to a u quark. We have made use of this structure in our discussion of the effective Lagrangian for K meson decays in Section 18.2. We will discuss CKM flavor mixing and its symmetry properties in more detail in Section 20.3.

It is interesting to note that there is no term within the structure we have described that gives a mass to the neutrino. If we wanted to generalize Eq. (20.98) to allow a neutrino mass term, we would have to introduce a new fermion field ν_R that is completely neutral under $SU(2)\times U(1)$. Then we could write the Higgs coupling

$$\Delta\mathcal{L}_\nu = -\lambda_\nu \epsilon^{ab} \overline{E}_{La} \phi_b^\dagger \nu_R + \text{h.c.}, \tag{20.109}$$

which would give the ν_e a mass, presumably comparable to that of the electron. But we know from experiment that neutrino masses are extremely small; the mass of the ν_e is known to be less than 10 eV. This extreme suppression of the neutrino masses would be naturally explained if the states ν_R do not exist. We will show in Section 20.3 that this assumption also implies that there are no transitions between leptons of different generations; this result is also in accord with very strong experimental bounds.

The Higgs Boson

This discussion of fermion mass generation emphasizes that the scalar field that causes spontaneous breaking of the gauge symmetry is an important ingredient in the structure of the Glashow-Weinberg-Salam theory. We should therefore ask whether it has any more direct manifestations.

To investigate this question, we will work in the *unitarity gauge*, analogous to that used for the Abelian model in Eq. (20.12). Let us parametrize the scalar field ϕ by writing

$$\phi(x) = U(x) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{20.110}$$

The two-component spinor on the right has an arbitrary real-valued lower component, given by the vacuum expectation value of ϕ plus a fluctuating real-valued field $h(x)$ with $\langle h(x) \rangle = 0$. This spinor is acted on by a general $SU(2)$ gauge transformation $U(x)$ to produce the most general complex-valued two-component spinor. We can now make a gauge transformation to eliminate $U(x)$ from the Lagrangian. This reduces ϕ to a field with one physical degree of freedom.

An explicit renormalizable Lagrangian that leads to a vacuum expectation value for ϕ is

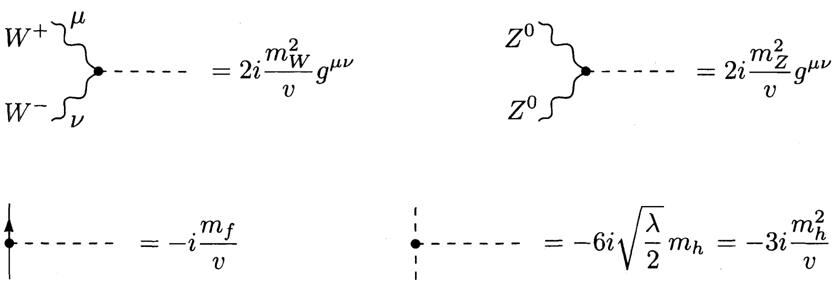
$$\mathcal{L} = \left| D_\mu \phi \right|^2 + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2. \tag{20.111}$$

The minimum of the potential energy occurs at

$$v = \left(\frac{\mu^2}{\lambda} \right)^{1/2}. \tag{20.112}$$

In the unitarity gauge, the potential energy terms in (20.111) take the form

$$\begin{aligned} \mathcal{L}_V &= -\mu^2 h^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4 \\ &= -\frac{1}{2} m_h^2 h^2 - \sqrt{\frac{\lambda}{2}}\, m_h h^3 - \frac{1}{4} \lambda h^4. \end{aligned} \tag{20.113}$$



$$= 2i\frac{m_W^2}{v}g^{\mu\nu} \qquad = 2i\frac{m_Z^2}{v}g^{\mu\nu}$$

$$= -i\frac{m_f}{v} \qquad = -6i\sqrt{\frac{\lambda}{2}}\,m_h = -3i\frac{m_h^2}{v}$$

Figure 20.6. Feynman rules for the couplings of the Higgs boson to vector bosons, to fermions, and to itself.

The quantum of the field $h(x)$ is a scalar particle with mass

$$m_h = \sqrt{2}\mu = \sqrt{2\lambda}\,v. \tag{20.114}$$

This particle is known as the *Higgs boson.* As for the fermions in the GWS theory, the Higgs boson has a mass whose general magnitude is controlled by the vacuum expectation value v, but whose precise value is determined by a new, unspecified, renormalizable coupling constant.

The expansion of the kinetic energy term of (20.111) in unitarity gauge yields the gauge boson mass term (20.62), plus additional terms involving the Higgs boson field. Explicitly,

$$\mathcal{L}_K = \frac{1}{2}(\partial_\mu h)^2 + \Big[m_W^2 W^{\mu+}W_\mu^- + \frac{1}{2}m_Z^2 Z^\mu Z_\mu\Big]\cdot\Big(1+\frac{h}{v}\Big)^2, \tag{20.115}$$

where m_W and m_Z are given by Eqs. (20.63).

Finally, the fermion mass terms in Eqs. (20.98) and (20.101) lead to direct couplings of the Higgs boson to fermions. Evaluating these terms in unitarity gauge, we find that, for any quark or lepton flavor, the Higgs boson couples according to

$$\mathcal{L}_f = -m_f\bar{f}f\Big(1+\frac{h}{v}\Big). \tag{20.116}$$

The Higgs boson couplings in Eqs. (20.113), (20.115), and (20.116) lead to the Feynman rules shown in Fig. 20.6.

In general, the couplings of the Higgs boson to other particles of the weak interaction theory are proportional to the masses of those particles. Thus, the particles that are most easily made in the laboratory have very weak couplings to the Higgs boson, which makes it difficult to observe this particle. In any event, the Higgs boson has not yet been found. As of this writing, the Higgs boson that we have just described has been searched for and excluded for values of m_h below 60 GeV. If the self-coupling λ is large, however, the Higgs boson could have a mass as large as 1000 GeV; thus, a large dynamic range remains unexplored.

The phenomenological properties of the Higgs boson are worked out in more detail in the Final Project of Part III.

A Higgs Sector?

Since there is no experimental evidence for the existence of the simple Higgs boson contained in the GWS model, it is worth asking whether the W and Z bosons might acquire mass by a more complicated mechanism. There are two aspects to this question.

First, is it certain that the W and Z bosons are gauge bosons of a spontaneously broken $SU(2) \times U(1)$ symmetry? The evidence for this idea comes from the universality of the couplings of the various quarks and leptons to the weak interactions. This universality is tested in the fact that the same value of the Fermi constant describes all charged-current weak-interaction processes, and that this same strength of coupling combined with a single value of $\sin^2\theta_w$ describes the whole range of weak neutral current phenomena. We have seen, especially in the discussion of Chapter 16, that the principle of local gauge invariance leads naturally to the prediction of universal, flavor-independent, coupling constants. No other principle is known that would explain this striking regularity. Thus there is compelling evidence that the underlying theory of the weak interactions is a spontaneously broken gauge theory.

However, it is quite possible that the mechanism of the spontaneous breaking of $SU(2) \times U(1)$ is more complicated than the simple model of a single scalar field that we have written in Eq. (20.111). In principle, the breaking of $SU(2) \times U(1)$ might be the result of the dynamics of a complicated new set of particles and interactions, which we will refer to as the *Higgs sector*. Experiment gives us only three properties of this new sector: First, it must generate the masses of the quarks and leptons. Second, it must generate the masses of the W and Z bosons. The third piece of information, which is the only nontrivial one, comes from the relation (20.73) between weak boson masses in the GWS theory,

$$m_W = m_Z \cos\theta_w. \tag{20.117}$$

This relation is satisfied experimentally to better than 1% accuracy, that is, to the level of one-loop radiative corrections. Whatever complicated mechanism we invoke to generate the spontaneous breaking of $SU(2) \times U(1)$, it should reproduce this relation in a natural way.

To understand the implications of relation (20.117), we must analyze the gauge boson mass matrix without assuming that $SU(2) \times U(1)$ is broken by the expectation value of a scalar field. Actually, it is possible to compute the gauge boson mass matrix under much less restrictive assumptions, using the argument given at the very end of Section 20.1. There we constructed the gauge boson mass matrix from the matrix elements for gauge currents to create or destroy Goldstone bosons. We will now show that relation (20.117) follows for a large class of models of $SU(2) \times U(1)$ breaking for which these matrix elements satisfy certain simple properties.

Any model of weak-interaction symmetry breaking must contain some set of fields that is responsible for the spontaneous breaking of $SU(2) \times U(1)$. Think of this sector of the theory as a field theory with a global $SU(2) \times U(1)$ symmetry, which is promoted to a local symmetry through its coupling to gauge bosons. In the theory with global symmetry, this symmetry is spontaneously broken to $U(1)$. Since three continuous symmetries are spontaneously broken, this sector must supply three Goldstone bosons, which will eventually be eaten by W^+, W^-, and Z^0. Call these three bosons π_a, where $a = 1, 2, 3$. Let $J^{\mu a}$ be the $SU(2)$ symmetry currents of the new sector, and let $J^{\mu Y}$ be the $U(1)$ current. The gauge boson mass matrix will then be constructed from the matrix elements (20.46), which here take the form

$$\langle 0| J^{\mu A} |\pi_b(p)\rangle = -ip^\mu F^A{}_b, \tag{20.118}$$

with $A = 1, 2, 3, Y$ and $b = 1, 2, 3$. Using the method of Eq. (20.55), we find that the gauge boson vacuum polarization contains the pole term

$$-\frac{i}{k^2}(g_A F^A{}_c)(g_B F^B{}_c), \tag{20.119}$$

summed over c, where $g_A = g$ for $A = 1, 2, 3$ and $g_A = g'$ for $A = Y$. Then we can identify the gauge boson mass matrix as

$$m^2_{AB} = g_A g_B F^A{}_c F^B{}_c. \tag{20.120}$$

To reproduce the known form of the weak gauge boson mass matrix, we must now place constraints on the $F^A{}_b$. First we must insure that the photon remains massless. This follows if the linear combination of charges (20.69) annihilates the vacuum. In the language of Eq. (20.118), we must insist that the corresponding linear combination of currents cannot excite a Goldstone boson:

$$\langle 0| \left(J^{\mu 3} + J^{\mu Y}\right) |\pi^a(p)\rangle = 0. \tag{20.121}$$

We can also achieve relation (20.117), using the following additional assumption: The symmetry-breaking sector has an $SU(2)$ global symmetry, under which the three Goldstone bosons and the three $SU(2)$ gauge currents transform as triplets, which remains exact when the $SU(2)$ gauge symmetry is spontaneously broken. This global $SU(2)$ symmetry implies that, if $A = a = 1, 2, 3$ in Eq. (20.118),

$$\langle 0| J^{\mu a} \left|\pi^b(p)\right\rangle = -iFp^\mu \delta^{ab}, \tag{20.122}$$

where F is a parameter with the dimensions of mass. Combining (20.122) and (20.121), we have

$$\langle 0| J^{\mu Y} \left|\pi^3(p)\right\rangle = +iFp^\mu. \tag{20.123}$$

Inserting this form for $F^A{}_b$ into (20.120), we find the gauge boson mass matrix

$$m^2 = F^2 \begin{pmatrix} g^2 & & & \\ & g^2 & & \\ & & g^2 & -gg' \\ & & -gg' & g'^2 \end{pmatrix}, \tag{20.124}$$

where the matrix acts on the gauge boson $(A^1_\mu, A^2_\mu, A^3_\mu, B_\mu)$. The eigenvectors of this matrix are precisely (20.63) and (20.64). To reproduce the eigenvalues, we need only define

$$v = 2F. \tag{20.125}$$

We have now shown that the GWS relation between the W and Z boson masses is not special to the situation in which the gauge symmetry is broken by a single scalar field. This relation follows from the much more general assumption of an unbroken global $SU(2)$ symmetry of the Higgs sector. This symmetry is often called *custodial $SU(2)$ symmetry.** We have seen this symmetry already as the global $SU(2)$ symmetry of the weak-interaction effective Lagrangian (20.94).

For the case of a single scalar field, the custodial symmetry arises in the following way: If we write the field ϕ in terms of its four real components, the Lagrangian (20.111) (ignoring the gauge couplings) has $O(4)$ global symmetry. The vacuum expectation value of ϕ breaks this symmetry down to $O(3)$, that is, $SU(2)$.

However, there are many other quantum field theories that break $SU(2)$ spontaneously while leaving another global $SU(2)$ symmetry unbroken. One rather complex example is given by QCD with two massless flavors, if we identify the gauged $SU(2)$ with the symmetry generated by U_L in (19.82) and identify the custodial $SU(2)$ with vectorial isospin symmetry. A copy of the familiar strong interactions with a mass scale large enough to give $F = 125$ GeV would be a perfectly acceptable model for the Higgs sector. (Unfortunately, it is not easy in this model to generate masses for the quarks and leptons.)

The question of the nature of the Higgs sector and the explicit mechanism of $SU(2) \times U(1)$ breaking is probably the most pressing open problem in the theory of elementary particles. We will discuss this question further in the Epilogue.

20.3 Symmetries of the Theory of Quarks and Leptons

Putting together the theory of strong interactions described in Chapter 17 and the theory of weak and electromagnetic interactions described in the previous section, we have now constructed a complete description of elementary particle interactions. It is interesting to investigate the symmetries of this theory, to ask what might be the fundamental symmetries of the underlying description of Nature.

We have already seen, in the arguments leading up to Eq. (15.17), that the Lagrangian of a gauge theory is highly restricted by the conditions of renormalizability and gauge invariance. In this section, we will construct the

*P. Sikivie, L. Susskind, M. Voloshin, and V. Zakharov, *Nucl. Phys.* **B173**, 189 (1980).

most general renormalizable Lagrangian consistent with the $SU(3) \times SU(2) \times U(1)$ gauge symmetries of the strong, weak and electromagnetic interactions. We can then ask what further global symmetries we must impose on this theory in order to give it the global symmetries that we see in Nature.

As a first step, we will ignore the Higgs scalar field and the mass terms of quarks, leptons, and gauge bosons. Then the Lagrangian of the theory of quarks and leptons is entirely specified by gauge invariance and renormalizability. We have

$$\mathcal{L}_K = -\frac{1}{4}\sum_i (F^a_{i\mu\nu})^2 + \sum_J \bar{\psi}_J (i\not{D})\psi_J, \tag{20.126}$$

where the index i runs over the three factors of the gauge group and the index J runs over the various multiplets of chiral fermions.

In principle, we could add to (20.126) the following pseudoscalar pure gauge operators:

$$\Delta\mathcal{L}_\theta = \sum_i \frac{\theta_i g_i^2}{64\pi^2}\epsilon^{\mu\nu\lambda\sigma} F^a_{i\mu\nu} F^a_{i\lambda\sigma}. \tag{20.127}$$

These terms are apparently odd under both P and T. However, we saw at the end of Section 19.2 that terms of this form can be generated or canceled by making a change of variables in the effective action. For example, the change of variables on the right-handed electron field

$$e_R \to e^{i\alpha} e_R \tag{20.128}$$

produces, according to (19.78) or (19.79), a correction to the Lagrangian involving the P- and T-odd combination of field strengths for the $U(1)$ gauge field

$$\Delta\mathcal{L} = \alpha \cdot \frac{g'^2}{32\pi^2}\epsilon^{\mu\nu\lambda\sigma} F_{\mu\nu} F_{\lambda\sigma}. \tag{20.129}$$

The coefficient of (20.129) differs from the corresponding coefficient in (19.79) because we transform only the right-handed chiral component of the electron field. If we were to transform another fermion field, of hypercharge Y, we would find a similar shift, with the coefficient proportional to Y^2. If this new field coupled to the $SU(2)$ or $SU(3)$ gauge fields, we would also find terms proportional to those field strengths. Thus, we can eliminate the term in (20.127) involving the $U(1)$ field strengths by making the change of variables (20.128) with $\alpha = -\frac{1}{2}\theta_1$. We can eliminate all three terms in (20.127) by making appropriate chiral rotations on three fermion multiplets, say, e_R, E_L, and Q_L. The change of variables (20.128), which rotates the right-handed electron field, is not symmetric under parity and, in fact, changes the definition of the parity operation. By making this change of variables, we are choosing new coordinates in which the P and T transformation properties of the whole theory are as simple as possible.

Let us now investigate the properties of the Lagrangian (20.126) under P, C, and T. The couplings of the QCD gauge bosons are invariant to each of these symmetries separately. However, the couplings of the $SU(2)$ gauge bosons violate P and C as much as possible. Recall from Section 3.6 that P converts a left-handed electron to a right-handed electron, and that C converts a left-handed electron to a left-handed positron. Each of these operations converts a particle that couples to $SU(2)$ gauge bosons to one that does not. However, the combination of these two operations interchanges left-handed particles with right-handed antiparticles. Thus the combined operation CP is a symmetry of (20.126). This Lagrangian is also invariant under time reversal.

Thus, we see that the discrete symmetries of C and P, on the one hand, and CP and T, on the other, stand on a very different footing in gauge field theories. Any chiral gauge theory will naturally violate C and P. At this level in our analysis, it is a mystery why C and P should be observed to be approximate symmetries of Nature. On the other hand, every theory of gauge bosons and massless fermions respects CP and T. It is known experimentally that Nature contains some interaction that violates CP, since the CP selection rules are weakly violated in the decays of the K^0 meson. But to find a source for this violation, we must add terms to our basic gauge theory (20.126).

We must, first of all, add dynamics to (20.126) that will cause the spontaneous breaking of $SU(2)\times U(1)$. We will begin by working with the simplest model with one Higgs scalar field ϕ. The most general renormalizable Lagrangian for ϕ is

$$\mathcal{L}_\phi = \left|D_\mu\phi\right|^2 + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2. \tag{20.130}$$

The Hermiticity of $\mathcal{L}_\phi$ implies that the parameters μ^2 and λ are real. Thus this Lagrangian respects P, C, and T. As discussed at the end of the previous section, this Lagrangian also automatically has the custodial $SU(2)$ symmetry required to produce the mass relation (20.117).

Finally, we must add the terms that couple the Higgs field to the quarks and leptons. Here, renormalizability and gauge invariance provide the weakest constraints, and there are many allowed interactions. We will first analyze the coupling of ϕ to the quark fields, and then generalize this discussion to the leptons.

In writing the Higgs field couplings to the quarks, we should recall that there are known to be three generations of quarks and leptons. Thus there are three doublets of left-handed quarks:

$$Q^i_L = \begin{pmatrix} u^i \\ d^i \end{pmatrix}_L = \left(\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L\right). \tag{20.131}$$

There are six right-handed quarks, three with $Y=\frac{2}{3}$ and three with $Y=-\frac{1}{3}$:

$$u^i_R = \left(u_R, c_R, t_R\right), \qquad d^i_R = \left(d_R, s_R, b_R\right). \tag{20.132}$$

When we couple gauge fields to these quarks, we replace the ordinary derivatives with covariant derivatives. This automatically gives all of the quarks the same coupling to QCD and all quarks of the same type the same coupling to the weak interactions. It does not allow mixing between the various quark flavors. However, the coupling of the Higgs field to the quarks does not follow from a gauge principle and so need not have any of these restrictions. Unless we require quark flavor conservation by postulating a new discrete symmetry of the theory, the Higgs couplings will, in general, mix the various flavors.

If we do not impose any additional symmetries on the theory, we must write the most general renormalizable gauge-invariant coupling with the structure of Eq. (20.101):

$$\mathcal{L}_m = -\lambda_d^{ij}\overline{Q}_L^i \cdot \phi\, d_R^j - \lambda_u^{ij}\epsilon^{ab}\overline{Q}_{La}^i\phi_b^\dagger u_R^j + \text{h.c.}, \tag{20.133}$$

where λ_d^{ij} and λ_u^{ij} are general, not necessarily symmetric or Hermitian, complex-valued matrices. The operation of CP interchanges the operators written in (20.133) with their Hermitian conjugates without changing the coefficients; thus, CP is equivalent to the substitutions

$$\lambda_d^{ij} \to (\lambda_d^{ij})^*, \qquad \lambda_u^{ij} \to (\lambda_u^{ij})^*. \tag{20.134}$$

CP would be a symmetry of (20.133) if the matrices λ^{ij} were real-valued; however, there is no principle that requires this. Without the imposition of further symmetry requirements, it seems that (20.133) does maximum violence to all discrete and flavor conservation symmetries.

However, just as we were able to eliminate the T-violating terms (20.127) by making a chiral rotation, we can simplify the form of (20.133) by appropriate chiral transformations. To find the required transformations, diagonalize the Hermitian matrices obtained by squaring λ_d and λ_u. Define unitary matrices U_u and W_u by

$$\lambda_u\lambda_u^\dagger = U_u D_u^2 U_u^\dagger, \qquad \lambda_u^\dagger\lambda_u = W_u D_u^2 W_u^\dagger, \tag{20.135}$$

where D_u^2 is a diagonal matrix with positive eigenvalues. Then

$$\lambda_u = U_u D_u W_u^\dagger, \tag{20.136}$$

where D_u is the diagonal matrix whose diagonal elements are the positive square roots of the eigenvalues of (20.135). We can define unitary matrices U_d and W_d in a similar way and decompose λ_d as

$$\lambda_d = U_d D_d W_d^\dagger. \tag{20.137}$$

Now make the change of variables

$$u_R^i \to W_u^{ij} u_R^j, \qquad d_R^i \to W_d^{ij} d_R^j. \tag{20.138}$$

This eliminates the unitary matrices W_u and W_d from the Higgs coupling (20.133). Since each of the three u_R^i and each of the three d_R^i have the same coupling to the gauge fields, W_u and W_d commute with the corresponding covariant derivatives. Thus, under (20.138),

$$\sum_i \left(\bar{u}_R^i(i\not{D})u_R^i + \bar{d}_R^i(i\not{D})d_R^i\right) \to \sum_i \left(\bar{u}_R^i(i\not{D})u_R^i + \bar{d}_R^i(i\not{D})d_R^i\right), \tag{20.139}$$

and so W_u and W_d disappear from the theory.

The analogous transformation on the left-handed fields also makes a dramatic simplification. Make the change of variables

$$u_L^i \to U_u^{ij} u_L^j, \qquad d_L^i \to U_d^{ij} d_L^j. \tag{20.140}$$

This transformation eliminates U_u, U_d from the terms in (20.133) that involve the lower component of the Higgs field. In unitarity gauge, only these terms survive. By combining the diagonal elements of D_u and D_d with the vacuum expectation value of the Higgs field, we can relate these elements to quark masses:

$$m_u^i = \frac{1}{\sqrt{2}} D_u^{ii} v, \qquad m_d^i = \frac{1}{\sqrt{2}} D_d^{ii} v. \tag{20.141}$$

With this replacement, (20.133) takes the form

$$\mathcal{L}_m = -m_d^i \bar{d}_L^i d_R^i \left(1 + \frac{h}{v}\right) - m_u^i \bar{u}_L^i u_R^i \left(1 + \frac{h}{v}\right) + \text{h.c.} \tag{20.142}$$

This has the standard form of quark mass terms and Higgs boson couplings. The transformations (20.138) and (20.140) thus convert the quark fields to the basis of mass eigenstates. In this basis, the mass terms and Higgs couplings are diagonal in flavor and conserve P, C, and T.

Since left-handed u and d quarks have identical couplings to QCD, the matrices U_u and U_d commute with the QCD couplings in the covariant derivative. However, u_L and d_L are mixed by the weak interactions, and so we must investigate the effect of (20.140) on the $SU(2) \times U(1)$ couplings more carefully. This is most easily done by referring to the Lagrangian (20.79). The matrices U_u and U_d cancel out of the pure kinetic terms in the first line of (20.79). They also cancel out of the electromagnetic current J^μ_{EM}; for example,

$$\bar{u}_L^i \gamma^\mu u_L^i \to \bar{u}_L^i U_u^{\dagger ij} \gamma^\mu U_u{}^{jk} u_L^k = \bar{u}_L^i \gamma^\mu u_L^i. \tag{20.143}$$

By the same logic, U_u and U_d cancel out of the Z^0 boson current.

However, in the current that couples to the W boson field, we find

$$J^{\mu+} = \frac{1}{\sqrt{2}} \bar{u}_L^i \gamma^\mu d_L^i \to \frac{1}{\sqrt{2}} \bar{u}_L^i \gamma^\mu \left(U_u^\dagger U_d\right)^{ij} d_L^j. \tag{20.144}$$

That is, the charge-changing weak interactions link the three u_L^i quarks with a unitary rotation of the triplet of d_L^i quarks, with this rotation given by the unitary matrix

$$V = U_u^\dagger U_d. \tag{20.145}$$

The matrix V is known as the *Cabibbo-Kobayashi-Maskawa* (CKM) mixing matrix.

The matrix V can have complex elements, but we can remove phases from V by performing phase rotations of the various quark fields. Before analyzing the case of three generations, it is useful to consider the case of two generations—u, d, c, and s. In this case, V is a 2×2 unitary matrix. Such a matrix has 4 parameters; we can write its most general form as

$$V = \begin{pmatrix} \cos\theta_c e^{i\alpha} & \sin\theta_c e^{i\beta} \\ -\sin\theta_c e^{i(\alpha+\gamma)} & \cos\theta_c e^{i(\beta+\gamma)} \end{pmatrix}. \tag{20.146}$$

One parameter of V is a rotation angle, and the other three are phases. We can remove these phases by performing the change of variables on the quark fields

$$q_L^i \to \exp[i\alpha^i] q_L^i. \tag{20.147}$$

This global phase rotation has no effect on any term of the Lagrangian except for the weak charged current (20.144). A phase rotation that is equal for all four quark flavors cancels out of (20.144). However, the other three possible phase transformations are just what we need to eliminate α, β, and γ.

When we have chosen the phases of the quark fields in this way, V takes the form

$$V = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix}. \tag{20.148}$$

Then the quark terms in the weak charged current can be written

$$J^{\mu+} = \frac{1}{\sqrt{2}}\big(\cos\theta_c \bar{u}_L\gamma^\mu d_L + \sin\theta_c \bar{u}_L\gamma^\mu s_L - \sin\theta_c \bar{c}_L\gamma^\mu d_L + \cos\theta_c \bar{c}_L\gamma^\mu s_L\big). \tag{20.149}$$

We have already seen, in Eqs. (18.31) and (18.32), that this is the way the s quark enters the weak interactions. The angle θ_c is the Cabibbo angle, as defined in Eq. (18.30).

The same set of arguments can be made for the theory with three generations. Here V is a general unitary 3×3 matrix. Such a matrix has 9 parameters. Of these, 3 are rotation angles; this is the number of parameters of an $O(3)$ rotation. The remaining 6 parameters are phases. We can remove these phases by making phase rotations of quark fields as in (20.147), but the overall phase is redundant, so we can remove only 5 of these phases. The final form of V contains 3 angles, of which one is the Cabibbo angle, and one phase. After all the transformations we have made, this one phase that makes some couplings of the W^+ to quarks complex is the only remaining parameter that violates CP.

We began this argument from a Lagrangian for the quark-Higgs boson coupling that seemed to violate all possible flavor symmetries and all discrete spacetime symmetries. However, by making changes of variables on the fermion fields, we have been able to dramatically simplify the form of the Lagrangian. If we keep only those terms involving the massless gauge bosons, the photon and the gluons, plus the mass terms and interactions written in (20.142), we see that this set of terms conserves P, C, T, and all flavor symmetries. This dramatic simplification occurs because the unbroken gauge symmetry of Nature, the gauge symmetry of QCD and QED, is nonchiral and can be written as acting on Dirac fermions. Since we have omitted only the

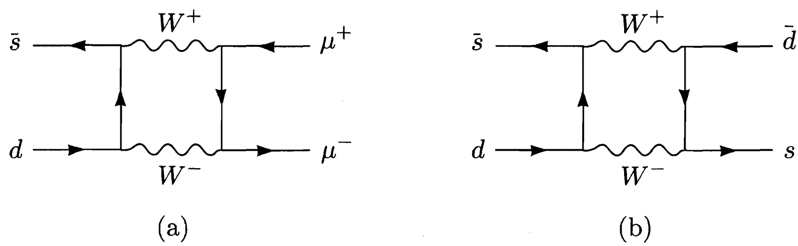


Figure 20.7. Higher-order diagrams that seem to give the leading contributions to flavor-changing weak neutral current processes: (a) $K^0 \to \mu^+\mu^-$; (b) $K^0 \leftrightarrow \overline{K}^0$.

effects mediated by the massive W and Z bosons, this much of the analysis already guarantees that Nature will appear, to a high degree of approximation, to respect the three separate discrete symmetries and all quark flavor conservation laws. Notice that we did not assume any fundamental global symmetries, but depended only on the assignment of gauge quantum numbers in the $SU(3) \times SU(2) \times U(1)$ gauge theory.

If we include the Z boson and the weak neutral current, we have a theory that violates P and C through Z exchange but that respects CP. In addition, this theory respects all flavor conservation laws. We describe this situation by saying that there is no flavor-changing weak neutral current. The experimental evidence for this statement is quite impressive. The best tests come from the study of the neutral K^0 meson, which is an $s\bar{d}$ bound state and so could decay by Z^0 exchange if this boson coupled to a flavor-changing current. In fact, the decay $K^0 \to \mu^+\mu^-$ is highly suppressed, to the level of the one-loop weak interaction correction shown in Fig. 20.7(a). Similarly, the interconversion of K^0 and $\overline{K}^0$, which could proceed directly if the Z^0 could change flavor, is suppressed to the level of the contribution shown in Fig. 20.7(b).

On the other hand, W bosons couple to currents that can change quark flavor, in a pattern parametrized by the Cabibbo angle and the other angles in the CKM matrix. Thus, heavy quark flavors decay by W boson exchange processes. Since the W couples to a current that contains only left-handed quarks, it mediates an interaction that violates P and C maximally. This violation of discrete symmetries is concealed from our ordinary experience because the amplitude for W exchange is small. However, this P and C violation is a dramatic qualitative feature of weak decays.

Since the coupling of the W to quarks contains an irreducible phase, these couplings in principle can violate CP. However, we have seen that this phase can be removed in a theory with only two generations. This means that the phase of the CKM matrix can have physical consequences only in a process that involves all three generations. Typically, this means that the CKM phase can contribute only to weak interaction loop corrections or to complicated exclusive decay processes. Thus the $SU(3) \times SU(2) \times U(1)$ theory can account for CP violation, and also explains why this effect is much weaker even than the weak interactions. It is interesting to note that Kobayashi and Maskawa originally proposed the existence of the third generation in order to provide a mechanism for CP violation.†

On the other hand, at this moment there is no conclusive evidence that the origin of CP violation is indeed the phase of the CKM matrix. All of the arguments we have given in this section have used the simplest model of the Higgs sector, in which this sector consists of a single scalar field. More general models of the Higgs sector may leave behind a more complicated set of quark-Higgs couplings than appear in (20.142), and some of these may violate CP. In addition, there may be terms in the Higgs sector itself that lead to CP violation. The origin of the observed CP violation is still an open problem that needs both theoretical and experimental exploration.

Before leaving this subject, we must discuss one more aspect of this argument that is still mysterious. To simplify the Lagrangian of the gauge theory of quarks to its final form, we needed to make chiral changes of variables in the functional integral. We saw in Section 19.2, and we reviewed at the beginning of this section, that such changes of variables produce the new P- and T-violating terms written in Eq. (20.127). It can be shown, using the fact that these terms are total derivatives, that the terms involving $SU(2)$ and $U(1)$ field strengths have no observable effects. However, the term involving QCD field strengths can induce an electric dipole moment for the neutron, a T-violating effect that has been searched for and excluded at an impressive level of accuracy. Thus the P- and T-violating combination of QCD field strengths cannnot be allowed to appear in the Lagrangian. On the other hand, if the original up and down quark Higgs coupling matrices were of the most general possible form, it seems that this cannot be avoided. This problem is known as the *strong CP problem*. To solve this problem, one must either constrain the Higgs coupling matrices, violating the spirit of the argument we have just concluded, or one must add additional structure to the Higgs sector.‡

Finally, let us discuss the general form and simplification of the Higgs boson couplings to leptons. When we wrote the Glashow-Weinberg-Salam Lagrangian in the previous section, we noted that no gauge field coupled to the right-handed neutrino. Thus, we chose to eliminate this particle from the theory. We might need right-handed components of the neutrinos to construct neutrino mass terms, but at the moment there is no evidence for nonzero neutrino masses. Thus, in the remainder of this section, we will assume that there are no right-handed neutrinos and work out the consequences of this assumption.*

---

†M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49**, 652 (1973).

‡The strong CP problem, its proposed solutions, and their unexpected implications are reviewed by R. D. Peccei in *CP Violation*, C. Jarlskog, ed. (World Scientific, 1989).

*In generalizations of the $SU(2) \times U(1)$ model, neutrinos can acquire Majorana

Generalizing Eq. (20.133), we can write the most general coupling of a Higgs boson to three generations of leptons. Since there are no right-handed neutrinos, the only possible coupling is

$$\mathcal{L}_m = -\lambda_\ell^{ij} \overline{E}_L^i \cdot \phi \, e_R^j + \text{h.c.} \tag{20.150}$$

To diagonalize this coupling, represent λ_ℓ in the form

$$\lambda_\ell = U_\ell D_\ell W_\ell^\dagger, \tag{20.151}$$

and eliminate the matrices U_ℓ and W_ℓ by the changes of variables

$$e_L^i \to U_\ell^{ij} e_L^j, \qquad \nu_L^i \to U_\ell^{ij} \nu_L^j, \qquad e_R^i \to W_\ell^{ij} e_R^j. \tag{20.152}$$

Since we are now making the same change of variables on the two components of the weak doublet E_L^i, this change of variables commutes with the $SU(2)$ interactions in the covariant derivative. Thus the unitary matrices U_ℓ and W_ℓ completely disappear from the theory. The result is a theory of leptons that conserves CP exactly and also conserves the lepton number of each generation. This last result is very accurately tested experimentally. For example, there is no evidence for the generation-changing muon decay processes $\mu^- \to e^-\gamma$ or $\mu^- \to e^-e^-e^+$; the branching ratios for these processes are known to be below 10^{-10}.

We have seen, then, that the $SU(3) \times SU(2) \times U(1)$ gauge theory of quarks and leptons does an excellent job of accounting for the symmetries and conservation laws that are observed in elementary particle phenomena. It predicts which symmetries should be exact in Nature and which should be approximate. For approximate symmetries, it gives an accurate estimate of the level of symmetry violation. Most remarkably (except for the one issue of the strong CP problem), none of these predictions depend on any underlying global discrete or flavor symmetries in the fundamental equations. The global symmetries that we observe in Nature follow only from gauge invariance and the specific representation assignments that we made in constructing our gauge theory description.

---

mass terms that are naturally very small. These models also respect the constraints on lepton flavor mixing described in the next paragraph. For an introduction to these ideas on neutrino mass, see P. Ramond, in *Perspectives in the Standard Model,* R. K. Ellis, C. T. Hill, and J. D. Lykken, eds. (World Scientific, 1992).