4
Infra-red safety and non-safety

In this chapter we examine the simplest measurable quantity that can be computed purely perturbatively in QCD: the total cross section for e^+e^- annihilation at high energy Q to hadrons. This is the paradigm of physical single-scale problems: when the renormalization scale μ is of order Q, low-order perturbation theory in $\alpha_s(Q)$ gives a valid estimate of the cross section.

Since the calculation involves quark and gluon final states in a confining theory, we will examine how to justify the use of perturbation theory with apparently incorrect states. There are divergences in individual terms in the calculation. But, in the total cross section, the divergences cancel after a sum over all terms of a given order of α_s. This property is called "infra-red (IR) safety", and in this case is a version of the theorem of Kinoshita (1962) and Lee and Nauenberg (1964) (KLN theorem).

More general situations need a systematic analysis of non-IR-safe situations, and are the primary concern of the rest of this book.

4.1 e^+e^- total cross section

We consider the process $e^+e^- \rightarrow$ hadrons, to lowest order in electromagnetism.[1] The amplitude, Fig. 4.1, involves an s-channel exchange of a photon of momentum q^μ, and an incoming electron and positron of momenta l_1 and l_2, with the center-of-mass energy being $Q = \sqrt{q^2}$. The leptonic and hadronic parts of the cross section factorize, as in DIS:

$$\sigma = \frac{e^4}{2Q^6} L_{\mu\nu} W^{\mu\nu}, \tag{4.1}$$

where (with neglect of the electron mass, and with unpolarized beams)

$$L^{\mu\nu} = l_1^\mu l_2^\nu + l_2^\mu l_1^\nu - g^{\mu\nu} l_1 \cdot l_2, \tag{4.2}$$

and

$$W^{\mu\nu}(q) = \int d^4x \; e^{iq \cdot x} \langle 0| \, j^\mu(x) \, j^\nu(0) \, |0 \rangle . \tag{4.3}$$

---

[1] There can be large IR-dominated higher-order electromagnetic corrections when the cross section is rapidly varying, e.g., near a narrow resonance. The techniques for unfolding such radiative corrections are standard, and we will not treat them here. A full treatment needs the addition to the amplitude of the Z exchange graph. This does not change the principles, so the reader is referred elsewhere, e.g., Ellis, Stirling, and Webber (1996), for details.
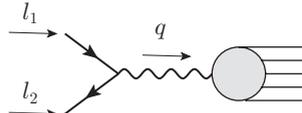
Fig. 4.1. Amplitude for $e^+e^- \longrightarrow$ hadrons.

Conservation of j^μ gives $q_\mu W^{\mu\nu} = 0$, so that we can decompose $W^{\mu\nu}$ in terms of a scalar structure function $R(Q^2)$ as

$$W^{\mu\nu} = \left(-g^{\mu\nu}q^2 + q^\mu q^\nu\right) \frac{1}{6\pi} R(Q^2)\theta(q^0). \tag{4.4}$$

Hence the cross section is

$$\sigma = \frac{4\pi\alpha^2}{3Q^2} R(Q^2). \tag{4.5}$$

The normalization coefficient in (4.4) is chosen so that R is the ratio of σ to the lowest-order cross section $e^+e^- \longrightarrow \mu^+\mu^-$:

$$R = \frac{\sigma(e^+e^- \longrightarrow \text{hadrons})}{\sigma(e^+e^- \longrightarrow \mu^+\mu^-, \text{LO, em})}. \tag{4.6}$$

Some authors define the denominator to be the complete cross section for $e^+e^- \longrightarrow \mu^+\mu^-$; the definition here is the PDG one.

A compilation of the data is shown in Fig. 4.2. At low energies, there are several large peaks, resonances corresponding to mesons made of light quarks. After that, the cross section generally decreases with energy, approximately as $1/Q^2$ as is generic for processes involving a large virtuality like the photon in Fig. 4.1. The trends are easier to see in the plot of R, whose basically constant value is interrupted at around 4 GeV and 10 GeV by jumps that correspond to the thresholds for production of charm and bottom quarks, preceded by sharp peaks for the bound states of these quarks with their antiquarks. Finally, the addition of a graph with the exchange of a Z instead of a photon in Fig. 4.1 gives rise to the prominent peak at $Q = m_Z \simeq 91$ GeV that interrupts the fall of the cross section.

4.1.1 Short-distance dominance in averaged cross section

When Q is large, the high virtuality of the photon in Fig. 4.1 suggests that it has a short lifetime, of order $1/Q$ in its rest frame, and hence that the process occurs over a short scale in time and distance. This makes it suitable for exploiting asymptotic freedom, so that a first approximation is obtained from the lowest-order graph, Fig. 4.3, for $e^+e^- \to q\bar{q}$.

However, the two currents in (4.3) need not actually have a small space-time separation. Consider a semi-classical approximation in which a quark and antiquark are assigned trajectories after their creation at a particular time and position. Suppose that the quark-antiquark force were such that they repeatedly bounce back to their creation position, as in Fig. 4.4. Now the incoming electron and positron have almost definite momenta
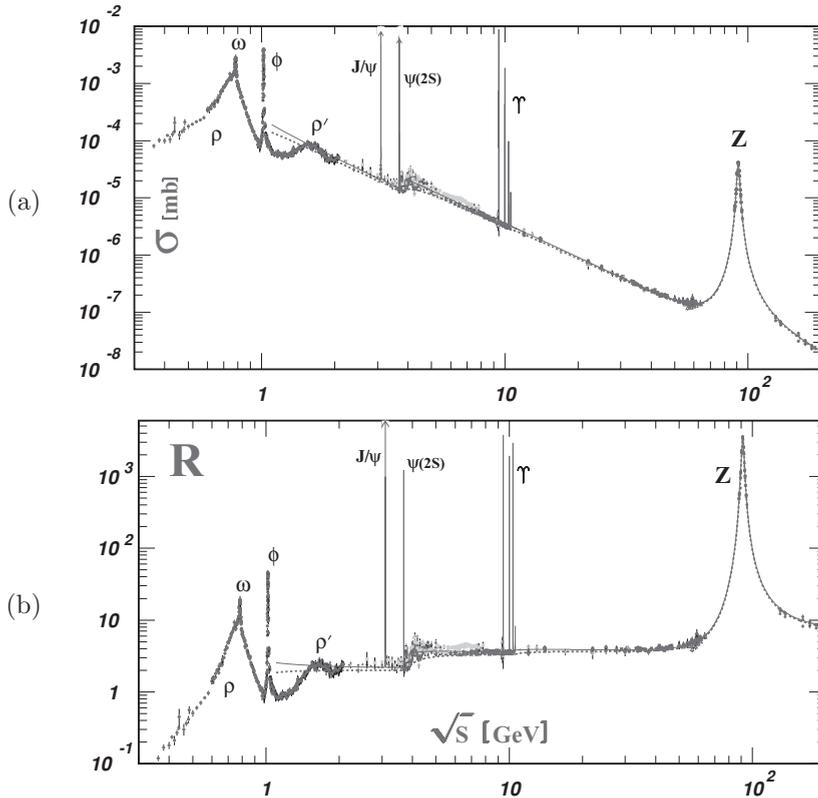
Fig. 4.2. (a) Total cross section and (b) R, for $e^+e^- \longrightarrow$ hadrons. Reprinted from Amsler *et al.* (2008), with permission from Elsevier. The dashed line is the lowest-order "parton-model" prediction, and the solid line is the 3-loop pQCD prediction from equations (1)–(3) of Chetyrkin, Harlander, and Kuhn (2000).

(in a normal experiment), so that their states can be represented by long wave packets, Fig. 4.5. Therefore their collision and the production of quark-antiquark pairs occurs over an extended time.

Now a pair produced late in the collision is in the same spatial position as a pair that is produced early but that has bounced back, and we get interference when we add the quantum-mechanical amplitudes for pair production at different times. At certain energies the phases of the interfering terms could all be the same, giving constructive interference, and a resonance peak. Off-resonance, the phases vary, giving destructive interference. Thus we can get sharp resonances, as seen at certain energies in the data in Fig. 4.2. These correspond to interference between quark-antiquark pairs produced at very different values of space-time positions.

For example, the sharp J/ψ and Υ peaks occur just below the thresholds for the production of c and b quark pairs, respectively; there, the heavy quarks are slowly moving and are easy to bring back to the production point. However, resonances are not present
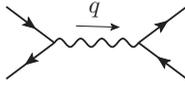
Fig. 4.3. Lowest-order diagram for
$e^+e^- \to q\bar{q}$ (or $e^+e^- \to \mu^+\mu^-$).
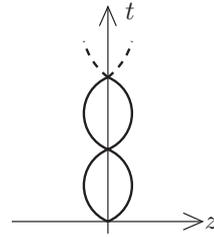
Fig. 4.4. Space-time evolution of semi-classical trajectory of a $q\bar{q}$ pair created at the origin, if the quark-antiquark force caused them to bounce back.

Fig. 4.5. Representation of wave packets for incoming electron and positron.

much above the heavy-quark thresholds, so that we deduce from the data that fast-moving quarks and antiquarks do not bounce back.

Unfortunately, the relevant long-distance phenomena in QCD are non-perturbative, and not readily susceptible to a first-principles analysis. So we ask what properties of the cross section are predicted purely perturbatively without any need to understand long-distance phenomena. A solution (Poggio, Quinn, and Weinberg, 1976) is to use a local average of the cross section in energy.

To understand this idea, we investigate the relation between the space-time structure of the scattering and the momentum spread in a physical initial state. This exemplifies a general issue that intuition and understanding can be obtained by studying the evolution of states in coordinate space, even though actual calculations are typically performed in momentum space.

Now a physical incoming e^+e^- state cannot be exactly a state with particles of definite momenta. We must use a superposition of momentum eigenstates corresponding to coordinate-space wave packets, as in Fig. 4.5:

$$|\psi\rangle = \sum_{l'_1, l'_2, \lambda_1, \lambda_2} \left|l'_1, \lambda_1, l'_2, \lambda_2; \text{in}\right\rangle \ \psi_1(l'_1, \lambda_1) \, \psi_2(l'_2, \lambda_2). \tag{4.7}$$

Here l'_1 and l'_2 are the momenta of the incoming electron and positron, and λ_1 and λ_2 label their spin states. The momentum-space wave functions $\psi_1(l'_1, \lambda_1)$ and $\psi_2(l'_2, \lambda_2)$ are narrowly peaked around central values of momentum l_1 and l_2. We let $q = l_1 + l_2$ be the corresponding central value of total momentum. The notation $\sum_{l'_1, l'_2}$ is the usual Lorentz-invariant integral over a particle's momentum, (A.15).

As in (4.1) and (4.3), we treat electroweak interactions perturbatively. The initial state $|\psi\rangle$ evolves to a slightly depleted version of $|\psi\rangle$ plus a hadronic component $|\phi\rangle$, plus
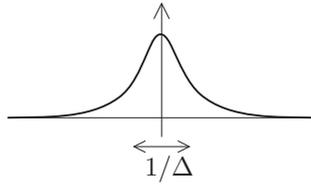
Fig. 4.6. A component of the wave function ϕ_μ for the state in (4.9), as a function of a component of position x in the overall center-of-mass.

components with scattered leptons:

$$|\text{final state}\rangle = |\phi\rangle + |\psi\rangle\,(1 - \ldots) + |\text{leptonic part}\rangle. \tag{4.8}$$

In the graph Fig. 4.1, the hadronic factor is a vacuum-to-X matrix element of the electromagnetic current, $\langle X,\,\text{out}|\,j^\mu(x)|0\rangle$, for a general hadronic out-state X.

Hence the hadronic final state $|\phi\rangle$ is $j^\mu(x)\,|0\rangle$ integrated with an x-dependent factor to be computed from the Feynman rules for Fig. 4.1, and the wave packet state (4.7). Now the creation of the hadronic final state occurs in the space-time region where the beams collide. So at later times, the QCD part $|\phi\rangle$ of the state in $e^+e^- \to$ hadrons is a time-independent Heisenberg state. We write this as

$$|\phi\rangle \stackrel{\text{def}}{=} \int \mathrm{d}^4x\; j^\mu(x)\,|0\rangle\, e^{-iq\cdot x}\phi_\mu(x), \tag{4.9}$$

where we have extracted a factor $e^{-iq\cdot x}$, anticipating that it is the dominant oscillatory factor in the coefficient, with q being the central value of the total momentum. Lowest-order electromagnetic perturbation theory gives

$$\phi_\mu(x) = \sum_{l'_1,l'_2,\lambda_1,\lambda_2} \frac{-ie^2\,\bar{v}_{l'_2\lambda_2}\gamma_\mu u_{l'_1\lambda_1}}{(l'_1 + l'_2)^2 + i0}\,\psi_1(l'_1,\lambda_1)\,\psi_2(l'_2,\lambda_2)\, e^{i(q-l'_1-l'_2)\cdot x}. \tag{4.10}$$

Here the $e^{-i(l'_1+l'_2)\cdot x}$ factor arises from Fourier-transforming the leptonic part of the Feynman graph, and the $e^{iq\cdot x}$ factor compensates the corresponding factor in (4.9).

The beams have approximately definite momenta, centered at $l_1 + l_2 = q$, so the oscillatory factor mostly cancels. Let the states be localized to within Δ in momentum. Then each component ϕ_μ is a smooth function with little oscillation, as in Fig. 4.6. Correspondingly the position x in (4.9) is localized to about $1/\Delta$.

Once the hadronic state $|\phi\rangle$ has been created by the current and the current is no longer acting because the coefficients $\phi_\mu(x)$ have become zero, the state cannot be destroyed, to lowest order in electroweak interactions. Thus the probability of the transition $e^+e^- \to$ hadrons is just $\langle\phi|\phi\rangle$. This is genuinely a scattering *probability*, not a cross section. The concept of a cross section arises when one observes that experiments are done with beams of particles which are distributed over an area that is large compared with the scattering region. The relative transverse separation of the beam particles has a broad distribution. The cross section is obtained by displacing one beam transversely with respect to the other,

and then integrating over the displacement $\boldsymbol{b}_{\mathrm{T}}$. Let the hadronic state with a displaced beam be $|\phi_{\boldsymbol{b}_{\mathrm{T}}}\rangle$, with wave function $\phi_\mu(x; \boldsymbol{b}_{\mathrm{T}})$. The cross section is

$$\sigma = \int \mathrm{d}^2 \boldsymbol{b}_{\mathrm{T}} \langle \phi_{\boldsymbol{b}_{\mathrm{T}}} | \phi_{\boldsymbol{b}_{\mathrm{T}}} \rangle = \int \mathrm{d}^4 x \, e^{iq \cdot x} \langle 0 | \, j^\mu(x) \, j^\nu(0) \, | 0 \rangle \, \tilde{t}_{\mu\nu}(x), \tag{4.11}$$

where

$$\tilde{t}_{\mu\nu}(x) = \int \mathrm{d}^2 \boldsymbol{b}_{\mathrm{T}} \int \mathrm{d}^4 w \, \phi_\mu \left(w + \frac{1}{2} x; \boldsymbol{b}_{\mathrm{T}} \right)^* \phi_\nu \left(w - \frac{1}{2} x; \boldsymbol{b}_{\mathrm{T}} \right), \tag{4.12}$$

which is localized in x to within $1/\Delta$. After a Fourier transformation

$$\tilde{t}_{\mu\nu}(x) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4} t_{\mu\nu}(k) e^{i(k-q) \cdot x}, \tag{4.13}$$

we find that the cross section is a weighted average in momentum space:

$$\sigma = \int \frac{\mathrm{d}^4 k}{(2\pi)^4} t_{\mu\nu}(k) W^{\mu\nu}(k) = \frac{1}{6\pi} \int_0^\infty \mathrm{d}M^2 \, R(M^2) \, f(M^2), \tag{4.14}$$

where

$$f(M^2) = - \int \frac{\mathrm{d}^4 k}{(2\pi)^4} t^\mu_\mu(k) \theta(k^0) \delta(k^2 - M^2). \tag{4.15}$$

We now see one result of the wave-packet construction: that a local average of $R(Q)$ over a range of Q of width Δ corresponds to a localization of the positions of the current operators to $x \sim 1/\Delta$. Of course, real particle beams are very narrow in momenta. But a broader averaging applied to the measured cross section gives a quantity with better localization in position and therefore with better perturbative calculability.

The standard momentum-space analysis gives the cross section in terms of R, from which we deduce the correct normalization of the averaging function f without needing the detailed wave-packet analysis:

$$\int_0^\infty \mathrm{d}M^2 \, f(M^2) = \frac{8\pi^2 \alpha^2}{Q^2}, \tag{4.16}$$

up to terms that vanish when $\Delta/Q \to 0$.

It is convenient to consider as a standardized quantity, one particular *normalized* local average of R:

$$\bar{R}(Q^2, \Delta^2) \stackrel{\text{def}}{=} \int \mathrm{d}s \, F(s - Q^2, \Delta^2) \, R(s). \tag{4.17}$$

Here $F(s - Q^2, \Delta^2)$ is one particular averaging function, of unit integral, centered at $s = Q^2$, and of width Δ^2. I choose

$$F(s - Q^2, \Delta^2) = \frac{\Delta^2}{\pi \left[(s - Q^2)^2 + \Delta^4 \right]}. \tag{4.18}$$

We assume Δ is somewhat less than Q^2, but not enormously so. If R is smooth in a region of Q, as is the case experimentally for most large values of Q, then the local average $\bar{R}$ is

almost equal to R; the averaging does nothing. But where R has sharp features, e.g., near the thresholds for c and b quark production, the average smooths out the sharp peaks and the thresholds.

4.1.2 When is perturbation theory good?

To put the concept of perturbative calculability in a general context, current ideas can be summarized in the following assertion:

Consider a situation where all vertices in a perturbative calculation are dominantly separated by small distances, of order $1/M$, and that we set $\mu \sim M$. Then QCD perturbation theory in powers of the weak coupling $\alpha_s(M)$ provides a good approximation.

That is, short-distance-dominated quantities are perturbatively computable. The integration over the positions of vertices is, of course, unrestricted. What matters for the above assertion is whether the vertices are *dominantly* close to some external vertices determined by the problem.

A similar assertion could be made about momentum-space Green functions, where the premise would be about the lines of the graph being dominated by high virtualities, of order M^2. However, this assertion is not so general. This can be seen from the matrix element (4.3) defining $W^{\mu\nu}$. The currents have fixed ordering and perturbation theory gives final states with on-shell quarks and gluons, so that not all propagators are far off-shell, even when the positions of the current operators are arbitrarily close, a situation that is different for the time-ordered product of operators.

4.2 Explicit calculations

Since the locally averaged quantity $\bar{R}$ is short-distance-dominated, we can use perturbation theory to predict it reliably. Therefore, to the extent we are away from resonances, we predict the unaveraged $R(Q^2)$, in both cases at large Q. The electromagnetic current has zero anomalous dimension within pure QCD.[2] So we change the renormalization scale μ to be of order Q, without changing R, and then expand in powers of the small coupling. We also approximate light-quark masses by zero. Thus:

$$
\begin{aligned}
R(Q^2, \mu, g(\mu), \underline{m}(\mu)) &= R(Q^2, cQ, g(cQ), \underline{m}(cQ)) \\
&\simeq R(Q^2, cQ, g(cQ), \underline{0}) \\
&= \sum_{n \geq 0} \alpha_s(cQ)^n R^{[n]}(c).
\end{aligned}
\tag{4.19}
$$

If we truncate the series at order N, then the error is of order α_s^{N+1}, so that we have an effective method of calculation given that $\alpha_s(cQ)$ is small. From Sec. 3.4, we expect (in the $\overline{\text{MS}}$ scheme) optimal applicability of perturbation theory when μ^2 is of order a

---

[2] This is *not* true beyond QCD (Collins, Manohar, and Wise, 2006), contrary to many statements in the literature.

Fig. 4.7. Lowest-order graph for amplitude used in R.

typical internal virtuality. This could be governed by the width of the smoothing function, so in (4.19) μ is a constant c times Q, and a good value could be $c = \frac{1}{2}$ or $\frac{1}{4}$. RG invariance implies that the value of c is irrelevant in an exact calculation, while in a truncated perturbation calculation the effect of a modest change in c is of order the expected truncation error.

In the remainder of this section, we perform the perturbative calculation of R to order α_s.

4.2.1 Lowest order

The single graph for the lowest-order calculation, Fig. 4.7, is the same as for $\mu^+\mu^-$ production, with the replacement of a muon line by a quark line. So with the neglect of quark particle masses, the lowest-order value of R is

$$R^{[0]} = 3 \sum_f e_f^2. \tag{4.20}$$

The factor 3 is for the sum over quark colors, and the sum is over the accessible flavors of quark, which depends on the value of Q relative to the quark masses.

Some complications now occur because of the non-negligible masses of the c, b and t quarks. Any quark that is not accessible kinematically, i.e., for which $m_f > Q/2$, should certainly be dropped from the sum. The remaining quarks we term "accessible". Provided that Q is much larger than the other quark masses, these masses may be neglected, as in the calculation giving (4.20).

The remaining case is when Q is comparable to $2m_f$ for one of the quarks. As regards perturbation theory, there is a threshold at $Q = 2m_f$ for production of quark f. Since there are sharp resonances just below threshold (Fig. 4.2) we should apply the averaging procedure in Q before using the elementary perturbative prediction.

Hence we deduce that it is a good first approximation to restrict the sum in (4.20) to those quarks with $2m_f < Q$, and to otherwise ignore the effects of quark masses. The known quark charges and masses then give a first prediction of R:

$$R^{[0]} = \begin{cases} 2 & \text{if } Q \lesssim 3 \text{ GeV}, \\ 3\frac{1}{3} & \text{if } 3 \text{ GeV} \lesssim Q \lesssim 10 \text{ GeV}, \\ 3\frac{2}{3} & \text{if } 10 \text{ GeV} \lesssim Q. \end{cases} \tag{4.21}$$

Once Z exchange effects become important, this prediction needs changing, so we do not include a possible last line, to include the t quark. For the inclusion of masses at lowest
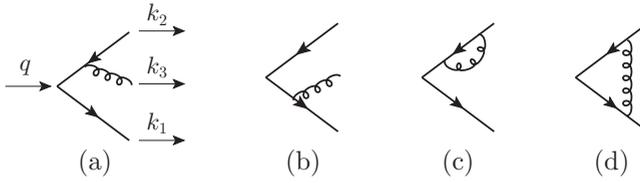
Fig. 4.8. NLO graphs for amplitudes for R.
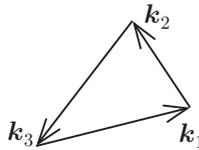
Fig. 4.9. Momentum configuration for 3-body final state.

order, see problem 4.2. For masses and the effects of Z at higher order, see Chetyrkin, Harlander, and Kuhn (2000).

4.2.2 Next-to-leading order: real gluon

The next-to-leading order (NLO) terms arise from the graphs of Fig. 4.8. One contribution is from the real-gluon emission graphs, (a) and (b), with a $q\bar{q}g$ final state:

$$\int \mathrm{dfsps} \; |(a) + (b)|^2 \,, \tag{4.22}$$

with dfsps given by (A.17). The other contribution is from the virtual corrections, (c) and (d), with a $q\bar{q}$ final state.

All of the terms individually have divergences which we regulate by using a space-time dimension $4 - 2\epsilon$.

Provided that the integrand involves only Lorentz scalars, the $(5 - 4\epsilon)$-dimensional integral for real-gluon emission can be simplified to a two-dimensional integral, so that angular averages can be performed to give (A.44). So we calculate the trace of $W_{\mu\nu}$:

$$W \overset{\mathrm{def}}{=} -g^{\mu\nu} W_{\mu\nu} = (3 - 2\epsilon)\frac{Q^2}{6\pi}R(Q^2). \tag{4.23}$$

In the overall center-of-mass, the 3-momenta of the final state form a triangle – Fig. 4.9, whose perimeter is $\sum_i |\boldsymbol{k}_i| = Q$, from energy conservation. The integration variables in (A.44) are the relative deficits of the spatial momenta relative to their maximum $Q/2$:

$$y_i = 1 - \frac{2|\boldsymbol{k}_i|}{Q}. \tag{4.24}$$

The integral is over positive values subject to $\sum_i y_i = 1$. We have

$$(k_1 + k_2)^2 = y_3 Q^2, \qquad (k_2 + k_3)^2 = y_1 Q^2, \qquad (k_3 + k_1)^2 = y_2 Q^2. \tag{4.25}$$

It is also convenient to factor out the lowest-order calculation in $4 - 2\epsilon$ dimensions, derived from (A.43) for the 2-body phase space:

$$W^{[0]} = R_0 \frac{Q^{-2\epsilon}}{2^{4-4\epsilon}\pi^{1/2-\epsilon}\Gamma(\frac{3}{2}-\epsilon)}(-g_{\mu\nu})\operatorname{Tr}\gamma^\mu \not{k}_1 \gamma^\nu \not{k}_2$$

$$= R_0 \frac{Q^{2-2\epsilon}(1-\epsilon)}{2^{2-4\epsilon}\pi^{1/2-\epsilon}\Gamma(\frac{3}{2}-\epsilon)}. \tag{4.26}$$

After a standard application of Feynman rules, etc., we find that the contribution of graphs (a) and (b) to W is

$$W^{[1]}(q\bar{q}g) = W^{[0]} \frac{\alpha_s C_F}{4\pi\,\Gamma(1-\epsilon)}\left(\frac{Q^2}{4\pi\mu^2}\right)^{-\epsilon}\int_0^1 \mathrm{d}y_1 \int_0^{1-y_1}\mathrm{d}y_2\,(y_1 y_2 y_3)^{-\epsilon}$$

$$\times\ \frac{4(y_3 + y_1 y_2 \epsilon) + 2(1-\epsilon)(y_1^2 + y_2^2)}{y_1 y_2}. \tag{4.27}$$

The sum over flavors and a factor e_f^2 are the same as in lowest order, and are in the factor $W^{[0]}$. Thus, the order-α_s correction factor is the same for all flavors of massless quark.

The integrand is singular when y_1 and/or y_2 is zero, and gives a divergence in the integral for space-time dimension 4 or less, i.e., when $\epsilon > 0$. In Ch. 5, we will analyze the physics of these and other divergences more generally. But for calculational purposes, it suffices that the divergence can be regulated and hence quantified by using a space-time dimension above 4, i.e., $\epsilon < 0$. In the ultimate result, for R, we will find a cancellation against divergences from the virtual-gluon graphs. The configuration of momenta at the singularities is easily deduced from the geometry of Fig. 4.9:

$$\begin{aligned}y_1 &= 0 \quad \text{gluon parallel to } k_2,\\ y_2 &= 0 \quad \text{gluon parallel to } k_1,\\ y_1 = y_2 &= 0 \quad \text{gluon of zero momentum.}\end{aligned} \tag{4.28}$$

In Ch. 5, we will analyze such singularities. The first two give "collinear divergences", where two final-state massless particles are parallel, and the last one gives a "soft divergence", where the gluon has zero momentum.

The integral is readily computed in terms of Γ functions. Its expansion in powers of ϵ exhibits the divergence quantitatively:

$$W^{[1]}(q\bar{q}g) = W^{[0]}\frac{\alpha_s C_F}{\pi}\left(\frac{Q^2}{4\pi\mu^2}\right)^{-\epsilon}\frac{\Gamma(-\epsilon)^2}{\Gamma(2-3\epsilon)}\left[1 - \frac{\epsilon(3-5\epsilon)}{2-3\epsilon}\right]$$

$$= W^{[0]}\frac{\alpha_s C_F}{4\pi}\left(4\pi e^{-\gamma_E}\right)^\epsilon\left[\frac{4}{\epsilon^2} + \frac{1}{\epsilon}\left(-4\ln\frac{Q^2}{\mu^2} + 6\right) + 2\ln^2\frac{Q^2}{\mu^2}\right.$$

$$\left. - 6\ln\frac{Q^2}{\mu^2} + 19 - \frac{7\pi^2}{3} + O(\epsilon)\right]. \tag{4.29}$$

That we obtain a relatively simple analytic result is associated with the masslessness of the quarks and gluons in the calculation.

4.2.3 Next-to-leading order: virtual gluon

For the virtual-gluon corrections, from Fig. 4.8(c) and (d), it is convenient to compute the matrix elements from Green functions with bare fields, rather than using the full counterterm structure of (3.13)–(3.15). One reason is that the electromagnetic current is simplest in terms of bare fields: $j_{\mathrm{em}}^{\mu} = \sum_f e_f \bar{\psi}_{f,0} \gamma^{\mu} \psi_{f,0}$. Another is that the implementation of LSZ reduction for massless theories is trivial.

The LSZ reduction formula tells us that to get an on-shell matrix element, we amputate complete external propagators and replace each by the square root of the residue of the particle pole. Let $z_2 = 1 + g^2 z_2^{[1]} + \dots$ be the residue of the pole in the propagator of a *bare* quark field. The one-loop term is

$$
\begin{aligned}
g^2 z_2^{[1]} &= \lim_{p^2 \to 0} \frac{i \not{p}}{p^2} \frac{g^2 C_F (2\pi\mu)^{2\epsilon}}{16\pi^4} \int \mathrm{d}^{4-2\epsilon} k \, \frac{-\gamma^{\nu} \not{k} \gamma_{\nu}}{(k^2 + i0)\,[(p-k)^2 + i0]} \\
&= \text{coefficient} \times \lim_{p^2 \to 0} (-p^2)^{-\epsilon} \\
&= 0.
\end{aligned}
\tag{4.30}
$$

Since dimensional regularization is used here to regulate *infra-red*-related divergences, we take ϵ negative, which gives the zero result in the last line. No UV counterterm is applied, since we work with bare fields.

The result (4.30) generalizes to all orders: an N-loop calculation gives a factor of $g^{2N} \mu^{2N\epsilon}$, and hence dimensional analysis shows that the power of p^2 is $(-p^2)^{-N\epsilon}$. Thus to all orders in perturbation theory the residue of the pole in the bare propagator is exactly $z_2 = 1$.

The only non-zero one-loop virtual contribution to R is therefore from the vertex graph (d). To get its contribution at order α_s to W, we multiply the graph by the complex conjugated LO graphs, then we add the complex conjugate, we take the trace of $W_{\mu\nu}$ with $-g^{\mu\nu}$ and perform the angular integral in $3 - 2\epsilon$ spatial dimensions. This gives

$$
W^{[1]}(q\bar{q}) = W^{[0]} \Re \frac{i g^2 C_F (2\pi\mu)^{2\epsilon}}{32\pi^4 Q^2 (1-\epsilon)} \int \mathrm{d}^{4-2\epsilon} k \, \frac{\operatorname{Tr} \not{k}_1 \gamma^{\kappa} (\not{k}_1 - \not{k}) \gamma_{\mu} (\not{k}_2 + \not{k}) \gamma_{\kappa} \not{k}_2 \gamma^{\mu}}{(k^2 + i0)\,[(k_1 - k)^2 + i0]\,[(k_2 + k)^2 + i0]}.
\tag{4.31}
$$

There is an extra factor of $4Q^2 (1 - \epsilon)$ in the denominator of the prefactor because of the normalization to $W^{[0]}$. Standard manipulations give

$$
\begin{aligned}
W^{[1]}(q\bar{q}) &= W^{[0]} \Re \frac{\alpha_s C_F}{4\pi} \left(\frac{-Q^2 - i0}{4\pi\mu^2} \right)^{-\epsilon} \frac{\Gamma(1+\epsilon)\Gamma(1-\epsilon)^2}{\epsilon^2 \Gamma(1-2\epsilon)} \left[\frac{-4}{1-2\epsilon} + 2\epsilon \right] \\
&= W^{[0]} \frac{\alpha_s C_F}{4\pi} \left(4\pi e^{-\gamma_E} \right)^{\epsilon} \left[-\frac{4}{\epsilon^2} + \frac{1}{\epsilon} \left(4 \ln \frac{Q^2}{\mu^2} - 6 \right) - 2 \ln^2 \frac{Q^2}{\mu^2} \right. \\
&\qquad \left. + 6 \ln \frac{Q^2}{\mu^2} - 16 + \frac{7\pi^2}{3} + O(\epsilon) \right].
\end{aligned}
\tag{4.32}
$$

4.2.4 Leading and next-to-leading order: total

In the total for $R^{[1]}$, the divergences cancel, and at $\epsilon = 0$ we find

$$R = R^{[0]} \left[1 + \frac{3\alpha_s(\mu)}{4\pi} C_F + O(\alpha_s^2) \right] = R^{[0]} \left[1 + \frac{\alpha_s(\mu)}{\pi} + O(\alpha_s^2) \right], \qquad (4.33)$$

with the physical value of C_F. Notice that both logarithms of Q/μ have canceled. This follows from the RG invariance of R, which implies that a logarithm of Q/μ first appears in the coefficient of α_s^2 (problem 4.5). We have left the renormalization scale μ arbitrary, but, as explained earlier, a value for μ of order Q should be used to ensure that higher-order calculations do not get large logarithms. Thus $O(\alpha_s^2)$ correctly represents the expected size of the error due to omission of higher-order perturbation theory.

This result is both reassuring and disturbing. It is reassuring that the divergences cancel in a quantity that was supposed to have a valid perturbation expansion. But it is also disturbing: the intermediate steps involve totally unphysical states. In another arena, QED, there are somewhat similar IR divergences because of the masslessness of the photon; but at least electrons and photons are actual identifiable particles. See Sec. 4.3 for a detailed analysis.

The calculation evidently makes important predictions. Among these is that measuring the ratio R gives an estimate of the sum of the squared charges of the accessible quarks. When first obtained, this was a rather dramatic result, and the data (Fig. 4.2) confirm the charge assignments of the quarks. Deviations from this value can be used to measure the strong coupling and to test its evolution with scale.

4.2.5 Full result, and phenomenological implications

The currently most accurate calculations may be traced from Chetyrkin, Harlander, and Kuhn (2000), where the calculation is extended at order α_s^3 to include quartic mass corrections (i.e., of order m^4/Q^4). With massless quarks the current results are

$$\frac{R}{R^{[0]}} = 1 + \frac{\alpha_s}{\pi} + \left(\frac{\alpha_s}{\pi} \right)^2 \left[\frac{365}{24} - 11\zeta_3 + n_f \left(-\frac{11}{12} + \frac{2}{3}\zeta_3 \right) \right]$$

$$+ \left(\frac{\alpha_s}{\pi} \right)^3 \left[\frac{87\,029}{288} - \frac{121}{8}\zeta_2 - \frac{1103}{4}\zeta_3 + \frac{275}{6}\zeta_5 \right.$$

$$\left. + n_f \left(-\frac{7847}{216} + \frac{11}{6}\zeta_2 + \frac{262}{9}\zeta_3 - \frac{25}{9}\zeta_5 \right) + n_f^2 \left(\frac{151}{162} - \frac{1}{18}\zeta_2 - \frac{19}{27}\zeta_3 \right) \right]$$

$$+ O\left(\frac{\alpha_s}{\pi} \right)^4$$

$$\approx 1 + \frac{\alpha_s}{\pi} + \left(\frac{\alpha_s}{\pi} \right)^2 (1.985\,71 - 0.115\,295 n_f)$$

$$+ \left(\frac{\alpha_s}{\pi} \right)^3 (-6.636\,94 - 1.200\,13 n_f - 0.005\,178\,36 n_f^2) + O\left(\frac{\alpha_s}{\pi} \right)^4. \qquad (4.34)$$
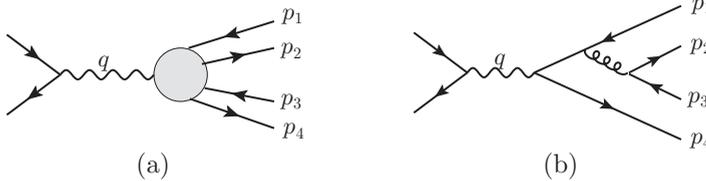
Fig. 4.10. (a) Matrix element of $q\bar{q}q\bar{q}$ fields used to obtain two-pion production. (b) An example of a perturbative graph for the matrix element.

Here, the $\overline{\text{MS}}$ unit of mass was set to $\mu = Q$, so that $\alpha_s = \alpha_s(Q)$. The logarithmic dependence on Q/μ can be restored with the aid of the renormalization group (problem 4.5).

4.3 Evolution of state

Individual terms in the perturbative calculation of R involve quarks and gluons rather than the hadrons that actually appear in the final state. To understand better why we nevertheless obtain a valid prediction of QCD, we examine the evolution of the hadronic final state, between the $j^\mu(x)$ and $j^\nu(0)$ operators in (4.11). The arguments in this section are not intended to be precise and rigorous.

Although the hadronic Heisenberg-picture state, $|\phi\rangle$, is time independent, its interpretation in terms of localized particle content does evolve. It can be analyzed by matrix elements of products of field operators between $|\phi\rangle$ and the vacuum. For example, consider [Fig. 4.10(a)]

$$\langle 0 \,|\, T\bar{u}(w_1)d(w_2)\bar{d}(w_3)u(w_4) \,|\, \phi \rangle, \tag{4.35}$$

where the fields annihilate $\bar{u}$, d, $\bar{d}$, and u quarks respectively. Fourier transformation gives a function of momenta p_1, p_2, p_3, p_4. Poles in this function correspond to particles in the asymptotic out-states. For example, a state with a π^- and a π^+ gives a pole in the exact matrix element at $(p_1 + p_2)^2 = m_\pi^2$ and at $(p_3 + p_4)^2 = m_\pi^2$. The poles are related to the coordinate-space asymptotics when the times of all the four fields are taken to $+\infty$; in the $\pi^-\pi^+$ example, the spatial components of w_1 and w_2 are close together and in the direction of the π^-, and similarly for w_3 and w_4 and the π^+.

In finite-order perturbation theory, we have diagrams like Fig. 4.10(b). This has no poles for pions, but only for quarks and gluons, for example at $p_1^2 = p_4^2 = m_u^2$ and $p_2^2 = p_3^2 = m_d^2$. Such poles give a large-time behavior for the individual graph that corresponds to a state that does not exist in a theory with color confinement; fixed-order perturbation theory gives an entirely incorrect approximation to asymptotic large-time Green functions and matrix elements. However, if the times are not too large, perturbation theory should approximate the true results. Thus the poles in fixed-order graphs imply that we do have, but only *approximately*, the propagation of the corresponding quarks and gluons.

Returning to the lowest-order graph for R, Fig. 4.7, we deduce in a rough fashion that at the earliest times we have predominantly an outward-moving q and $\bar{q}$, as at the lower
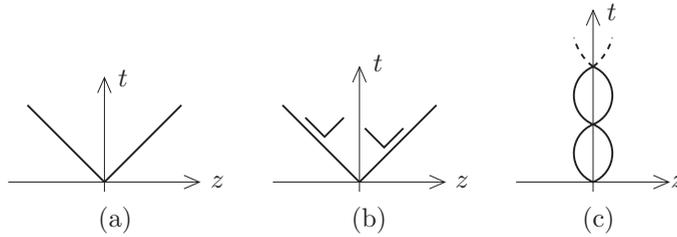
Fig. 4.11. Three semi-classical scenarios for evolution of $q\bar{q}$ system: (a) approximately free, (b) string, (c) spring. In case (b), extra $q\bar{q}$ pairs are produced in the middle.

end of Fig. 4.4. We can reasonably assign them a virtuality of some size M^2 that is much less than Q^2. The Lorentz boost to energy $Q/2$ implies that the lifetime of the $q\bar{q}$ state is of order Q/M^2. Perturbative corrections, like those in Fig. 4.8 or Fig. 4.10(b), alter the state, for example by changing the probability of the $q\bar{q}$ state and by adding a component with a gluon. At late times, QCD perturbation theory is entirely inapplicable, in the domain where, in the real world, the system non-perturbatively hadronizes into a set of isolated color-singlet hadrons.

4.3.1 String model for hadronization

If we ignored any knowledge of the real world we could imagine at least three scenarios for the time development, as illustrated in Fig. 4.11:

- The quarks and gluons continue basically unhindered into the observed final state, as in QED, where there is no confinement. Let us call this the "unconfined" or "free-quark-and-gluon" picture.
- In the gluon field between the quark and antiquark, extra $q\bar{q}$ pairs are made. We call this the "breakable string" picture. The $q\bar{q}$ pairs combine into color-singlet hadrons, mostly pions. Nothing returns to the production point, and the general momentum flow corresponds to the system at short times, which is little deflected. But the space between the ends of the kinematic range is filled in with particles, and the detected particles are hadrons, not quarks and gluons.
- A confining potential exists, which brings the quarks and gluons back. We can call this the "unbreakable elastic spring" picture. The final states form a sequence of bound states. After multiple bounces, it may be that the states decay, perhaps in the style of the string picture, but the directions of the decay products need not be very correlated with the initial $q\bar{q}$ direction.

Purely perturbative calculations in QCD cannot decide between these scenarios. But we can appeal to experiment, semi-classical intuition, modeling, and lattice gauge theory calculations, at least. The unconfined scenario is ruled out experimentally. The increase in α_s in the infra-red is a precondition for a rising potential. But the bound states or resonances in the spring picture do not appear to be relevant except close to quark thresholds, where

there is little energy for producing extra particles, and where the initial q and $\bar{q}$ are moving slowly.

It is the string picture that seems to be approximately correct. Embodied quantitatively in the semi-classical Lund string model (Andersson, 1998), it rather successfully describes the hadronization of quarks and gluons. In this model, when the $q\bar{q}$ separation is large enough, the gluon field collapses to a flux tube ("string") with a fixed cross-sectional area, and with a constant energy per unit length. Without $q\bar{q}$ production, this would correspond to a linearly rising potential, which has significant phenomenological support from quark models of hadrons, etc.

The Lund model postulates that creation of light $q\bar{q}$ pairs occurs in the string with a constant rate per unit length and unit time; this is the only Lorentz-invariant possibility. In a strong coupling, strong field situation such as we have here, the string therefore breaks, we have inelastic scattering, and the description in terms of a genuine potential breaks down.

A more detailed investigation shows that the string breaking and the hadronization occur along a hyperbolic region $t^2 - z^2 \sim 1/\Lambda^2$. The fastest particles, with energies of order Q, are generated at the ends of the string in a time of order Q/Λ^2, while the slowest particles are generated in the middle in a time of order $1/\Lambda$.

The Lund model is plausible and natural as a first approximation to real QCD dynamics in situations such as e^+e^- annihilation at high energies. For each outgoing parton, the model leads to the production of a jet of hadrons with approximately the 4-momentum of the parton. This can be seen in event pictures like Fig. 2.3 for a similar situation in DIS.

The validity of the string model depends on specific dynamical properties of real QCD, with its light u and d quarks. In contrast, there is the solvable model of 't Hooft (1974), pedagogically reviewed in Manohar (1998). This model is QCD but in $1 + 1$ space-time dimensions with a gauge group U(N) taken in the limit $N \to \infty$. This model provides an example of the "elastic spring" scenario;[3] the large N limit suppresses the $q\bar{q}$ production that causes string breaking in the Lund model. In the 't Hooft model, the final states in $e^+e^- \to$ hadrons form an infinite sequence of meson bound states with no continuum, whereas a simple perturbative calculation gives a continuum. It is a local average of the true cross section that agrees with the perturbative calculation, as we saw earlier. Explicit calculations support the general result, as was particularly clearly shown by Einhorn (1976), where the result was also extended to other cases, like DIS.

4.3.2 Analysis in terms of final states

We decompose the averaged cross section (4.11) in terms of a basis for the hadronic final states:

$$\sigma = \sum_X \int d^4x \; e^{iq \cdot x} \langle 0 | j^\mu(x) | X \rangle \langle X | j^\nu(0) | 0 \rangle \, \tilde{t}_{\mu\nu}(x), \qquad (4.36)$$

---

[3] The 't Hooft model is normally said to give an example of a string model. But I use the name "elastic spring" to emphasize its unbreakability, to contrast with the fragility of the string in real QCD.

and analyze the states in three bases:

- a momentum basis for the true out-states (involving hadrons);
- a spatially localized basis obtained using quark and gluon fields not too long after the creation of $|\phi\rangle$;
- a momentum basis for quark and gluon out-states, as seen in dimensionally regularized weak coupling perturbation theory. Here we must go to a space-time dimension above 4, i.e., to $\epsilon < 0$, so that the IR behavior is mild enough that the ordinary S-matrix exists.

The first basis gives the true distribution of observed final-state particles, but its use in calculations requires an unavailable non-perturbative solution of QCD. The second basis is most fundamentally suited to perturbative calculations, by working only with objects involving short distances. It completely justifies short-distance dominance for averaged cross section, but there is no known formulation explicit enough for actual calculations. The third basis, involving a momentum-space decomposition, is the easiest calculationally, but it involves a basis constructed from the $t \to \infty$ behavior of Green functions when a regulator is applied.

The low-order calculation of the individual terms in the ratio R in the third basis is only appropriate when the coupling is small enough that higher-order terms are not larger than lower-order terms. Given the *double* poles in ϵ that occur per loop, this implies that we should only apply the calculation when $\alpha_s \lesssim \epsilon^2$ (with ϵ *negative*). When the IR regulator is removed, the range of validity of the calculation shrinks to zero. So the cancellation of divergences in R is not sufficient *by itself* to justify the use of the result for R at non-zero $\alpha_s(Q)$.

But the result for R is independent of the basis for the completeness sum $\sum_X |X\rangle \langle X|$, so we can use the short-distance quark-gluon basis to justify the validity of perturbation theory for R for non-zero $\alpha_s(Q)$.

4.4 Dispersion relation and effective virtuality of final-state quarks and gluons

A perturbative calculation of $R(Q)$ involves cut graphs with an on-shell final state. In this section, I show that after a local average in Q, $R(Q)$ is given by an integral over an *uncut* graph, in which the final-state quarks and gluons are effectively off-shell by order Q^2. The derivation provides general principles that we will frequently generalize to other situations.

We consider a Green function $\Pi^{\mu\nu}$ that is defined like $W^{\mu\nu}$, but with *time-ordered* current operators. It has an associated scalar function $\Pi(q^2)$:

$$\Pi^{\mu\nu}(q) = \left(-g^{\mu\nu}q^2 + q^\mu q^\nu\right) \Pi(Q^2) = i \int \mathrm{d}^4x \; e^{iq\cdot x} \langle 0| \, T \, j^\mu(x) \, j^\nu(0) \, |0\rangle . \quad (4.37)$$

Note the factor i in the last part. Diagrammatically, $\Pi^{\mu\nu}$ and $W^{\mu\nu}$ are notated in Fig. 4.12 .
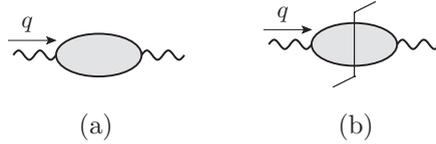
Fig. 4.12. (a) Uncut and (b) cut diagrams for the hadronic part of the photon self-energy, i.e., for $\Pi^{\mu\nu}$ and $W^{\mu\nu}$.
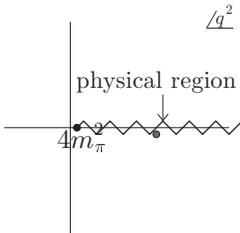
Fig. 4.13. Analyticity in Q^2 for $\Pi(Q^2)$.
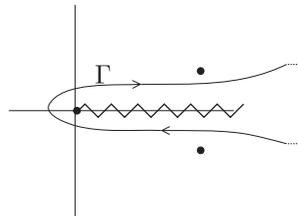
Fig. 4.14. Contour to relate $\bar{R}$ to Π. The off-real-axis singularities are those in the averaging function $F(s - Q^2, \Delta^2)$. The dots represent the singularities of F.

Now, $\Pi(Q^2)$ is an analytic function of Q^2 with a cut and singularities along the positive real axis, as in Fig. 4.13. When Q^2 is below the threshold for physical final states, Π is real, in particular for space-like q^μ. When Q^2 is above threshold, the physical region is on the upper side of the cut. Moreover, the cut amplitude is twice the imaginary part, as is provable from a dispersion relation. Hence

$$R(Q^2) = \frac{2\Pi(Q^2 + i0)}{i} = \frac{\Pi(Q^2 + i0) - \Pi(Q^2 - i0)}{i}. \qquad (4.38)$$

Hence we can relate the averaged R to the uncut amplitude Π:

$$\bar{R}(Q^2, \Delta^2) = \frac{6\pi}{i} \int_\Gamma \mathrm{d}s \; F(s - Q^2, \Delta^2)\, \Pi(s) \qquad \text{[general } F\text{]}$$

$$= \frac{6\pi}{i} \left[\Pi(Q^2 + i\Delta^2) - \Pi(Q^2 - i\Delta^2) \right] \qquad \text{[standard } F\text{]}. \qquad (4.39)$$

In the first line, the contour Γ loops around the positive real axis inside of the singularities of the averaging function F, as in Fig. 4.14. In the second line, F is chosen to have the standard form in (4.18) and the contour is closed on the two poles of F.

Thus we have expressed $\bar{R}$ in terms of Π evaluated at non-physical values of the momentum. If the averaging range Δ is large, then the momentum is correspondingly far from the physical region. The Landau analysis of the singularities of Feynman graphs (Ch. 5) shows that the contour for the loop-momentum integrals in Π can then be chosen to avoid the poles of the internal propagators. If Δ is of order Q, the internal lines have (typically complex) virtualities of order Q^2.

Therefore in the calculation of $\bar{R}$, we can treat the internal quark and gluon lines as far off-shell, thereby justifying its perturbative treatment in an asymptotically free theory.

4.5 Generalizations

One simple generalization of the work in this chapter is to allow for Z exchange as well as photon exchange. This is needed for fits to the high-energy parts of the data in Fig. 4.2. Another generalization (Baikov, Chetyrkin, and Kuhn, 2008) is to the hadronic part of the decay rate of the τ lepton, where the initiating boson is the W.

For these cases, the same principles apply as to the case we treated: There is a cancellation of IR-sensitive regions, leaving a quantity for which perturbation theory is applicable. Such quantities we call "IR-safe".

To analyze more general situations, we use the Libby-Sterman argument to be explained in Ch. 5; this determines both the nature and power-counting of the IR-sensitive regions. The most interesting cases are where the cancellations of divergences fail to occur. For many of these, we will be able to derive factorization theorems, where only part of an amplitude or cross section is IR-safe.

One outcome will be a discussion of IR-safe jet cross sections in Sec. 12.13.4.

Exercises

4.1 Compute the contribution of a scalar quark to R, to lowest order. What is the angular distribution of the $q\bar{q}$ final state in both the spin-0 and spin-$\frac{1}{2}$ cases?

4.2 Compute the value of $R^{[0]}$ with quark masses taken into account. You should get

$$R^{[0]} = \sum_{i:\, m_i < Q/2} 3e_i^2 \left(1 - \frac{2m_i^2}{Q^2}\right) \sqrt{1 - \frac{4m_i^2}{Q^2}}. \tag{4.40}$$

4.3 (**) Compute the order α_s correction to R for scalar quarks.

4.4 (*) I wrote that near a resonance in e^+e^- annihilation, the outgoing state is obtained by quantum-mechanical interference involving sources at a large range of time scales. Despite a large value of Q, the separation x of the currents in (4.3) is not small, of order $1/Q$. Verify these statements explicitly. You could use the following approximation to the cross section near the resonance of mass M and width Γ:

$$\frac{C}{(Q^2 - M^2)^2 + \Gamma^2 M^2}. \tag{4.41}$$

4.5 (*) Using the RG β function for the effective coupling, find the Q/μ dependence of the coefficients in the formula for R, equation (4.34).