Quantum ChromoDynamics

M. H. Seymour
University of Manchester, UK, and CERN, Geneva, Switzerland

Abstract
These lectures on QCD stress the theoretical elements that underlie a wide range of phenomenological studies, particularly gauge invariance, renormalization, factorization and infrared safety. The three parts cover the basics of QCD, QCD at tree level, and higher order corrections.

1 Basics of QCD

1.1 Introduction

QCD is the theory of the strong nuclear force, one of the four fundamental forces of nature. It describes the interactions of quarks, via their colour quantum numbers. It is an unbroken gauge theory. The gauge bosons are gluons. It has a similar structure to QED, but with one important difference: the gauge group is non-Abelian, SU(3), and hence the gluons are self-interacting. This results in a negative β-function and hence asymptotic freedom at high energies and strong interactions at low energies.

These strong interactions are confining: only colour-singlet states can propagate over macroscopic distances. The only stable colour singlets are quark–antiquark pairs, mesons, and three-quark states, baryons. In high energy reactions, like deep inelastic scattering, the quark and gluon constituents of hadrons act as quasi-free particles, partons. Such reactions can be factorized into the convolution of non-perturbative functions that describe the distribution of partons in the hadron, which cannot be calculated from first principles (at present) but are universal (process-independent), with process-dependent functions, which can be calculated as perturbative expansions in the coupling constant α_{S}.

Beyond leading order in α_{S}, the parton distribution functions and coefficient functions become intermixed. They can still be factorized, but the parton distribution functions become energy-dependent. Although the input distributions at some fixed energy scale still cannot be calculated, the energy dependence is given by perturbative evolution equations.

In sufficiently inclusive cross sections, called infrared safe, the non-perturbative distributions cancel and distributions can directly be calculated in perturbation theory. Non-perturbative corrections are then suppressed by powers of the high energy scale. The most important examples are jet cross sections, where jets of hadrons have a direct connection to the perturbatively-calculable quarks and gluons.

This course will attempt to give a brief overview of the subject. The approach will be pretty phenomenological, with most results stated rather than derived. I will however attempt to sketch in most cases roughly how they would be derived. One thing I will not have time to go into in much detail will be heavy quarks: in most cases we will treat the d, u, s, c and b quarks as massless and neglect the top quark, an approximation that I will motivate in Section 1.9.

It is hard to give a better introduction to the subject than the book 'QCD and Collider Physics', by Keith Ellis, James Stirling and Bryan Webber [1]. So I will follow the ESW approach and notation pretty closely. In most cases they will be able to give you a few more details and references to much more detailed treatments if you want to go further. For a much more detailed treatment of the formulation of QCD and its renormalization in particular Peskin and Schroeder [2] is also unbeatable.

As there are many parallels with QED I will have to assume prior knowledge of the basics of QED and that you can calculate a few simple cross sections. However we start by recapitulating a few features.

1.2 Basics of QED

QED is a gauge theory with gauge group U(1). It can be derived using the gauge principle. The classical Lagrangian density for n types of non-interacting fermion is

$$\mathcal{L}_{\text{ferm}} = \sum_i^n \bar{f}_i(i\partial\!\!\!/ - m_i)f_i, \tag{1.1}$$

where f_i is a spinor-valued wave function describing plane waves of momentum p_i, $\bar{f}_i$ its Dirac conjugate $f_i^\dagger\gamma^0$, $a\!\!\!/$ is shorthand for $\gamma^\mu a_\mu$ and γ^μ are Dirac spinor matrices with anticommutation relation

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{1.2}$$

The Lagrangian density (1.1) is invariant under global changes of gauge,

$$f_i \to f_i' = \exp(ie_i\theta)f_i, \tag{1.3}$$

where e_i is an arbitrary flavour-dependent parameter, which will turn out to be proportional to electric charge. We can derive QED by asking how we would need to modify (1.1) to make it also invariant under local changes of gauge,

$$f_i(x) \to f_i'(x) = \exp(ie_i\theta(x))f_i(x). \tag{1.4}$$

This can be done by introducing a new vector-valued field A_μ, which transforms under the same change of gauge like

$$A_\mu(x) \to A_\mu'(x) = A_\mu(x) + \frac{i}{e}\Big(\partial_\mu \exp(i\theta(x))\Big)\exp(-i\theta(x)), \tag{1.5}$$

and by replacing the derivative ∂_μ by the covariant derivative,

$$D_\mu = \partial_\mu + ie\hat{Q}\,A_\mu, \tag{1.6}$$

where $\hat{Q}$ is the charge operator, defined by

$$\hat{Q}\,f_i = e_i f_i. \tag{1.7}$$

Since A_μ is a new field that we have introduced, we must make it physical by adding a kinetic term,

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{1.8}$$

where the field strength tensor $F_{\mu\nu}$ is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.9}$$

The classical QED Lagrangian density is therefore given by

$$\mathcal{L}_{\text{classical}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \sum_i^n \bar{f}_i(iD\!\!\!\!/ - m_i)f_i. \tag{1.10}$$

This is now invariant under local changes of gauge.

Perturbative calculations are made according to the Feynman rules. These can be read off from the action, defined by

$$S = i\int d^4x\,\mathcal{L}. \tag{1.11}$$

There is however one complication. The photon propagator $\Delta_{\gamma,\mu\nu}(p)$ is derived from the inverse of the bilinear term in A_μ:

$$\Delta_{\gamma,\mu\nu}(p) \times i\Big[p^2 g^{\nu\sigma} - p^\nu p^\sigma\Big] = \delta_\mu^\sigma. \tag{1.12}$$

This does not have an inverse. However, we can exploit the gauge invariance of the theory to rewrite it in a physically equivalent form that is invertible. Since the Lagrangian density is gauge invariant, we can choose some convenient gauge to work in and the final answer should be independent of which we chose. For example, in the covariant gauge, we have the condition

$$\partial^\mu A_\mu = 0 \tag{1.13}$$

at every space-time point. We can therefore add an extra term to the Lagrangian density

$$\mathcal{L}_{\text{gauge-fixing}} = -\frac{1}{2\lambda}\left(\partial^\mu A_\mu\right)^2, \tag{1.14}$$

where λ is an arbitrary parameter, and provided we work in a covariant gauge we cannot have changed the physics, since we have only added zero. (This is essentially just the method of undetermined Lagrange multipliers for minimizing an action subject to a constraint: the constraint is (1.13) and the multiplier is $1/2\lambda$.) The final results must clearly be independent of λ, although it will appear at intermediate steps of calculations. Common choices are $\lambda = 1$ (Feynman gauge) and $\lambda \to 0$ (Landau gauge). For arbitrary λ, we must now solve

$$\Delta_{\gamma,\mu\nu}(p) \times i\Big[p^2 g^{\nu\sigma} - (1-\frac{1}{\lambda})p^\nu p^\sigma\Big] = \delta^\sigma_\mu, \tag{1.15}$$

which yields

$$\Delta_{\gamma,\mu\nu} = \frac{i}{p^2}\left(-g_{\mu\nu} + (1-\lambda)\frac{p_\mu p_\nu}{p^2}\right). \tag{1.16}$$

Clearly the Feynman gauge offers significant calculational advantages, so we use it for most of the rest of this course.

Another popular class of gauges are the axial (or physical) gauges, defined in terms of an arbitrary vector n, by

$$\mathcal{L}_{\text{gauge-fixing}} = -\frac{1}{2\lambda}\left(n^\mu A_\mu\right)^2. \tag{1.17}$$

These have the result that an on-shell photon has two polarization states, which, in the $(n+p)$ rest-frame, are purely transverse to its direction. The penalty is that the propagator becomes more complicated,

$$\Delta_{\gamma,\mu\nu} = \frac{i}{p^2}\left(-g_{\mu\nu} + \frac{n_\mu p_\nu + p_\mu n_\nu}{n\cdot p} - \frac{(n^2+\lambda\, p^2)p_\mu p_\nu}{(n\cdot p)^2}\right). \tag{1.18}$$

Obviously some simplification is obtained by setting $n^2 = 0$ and $\lambda \to 0$ (the 'lightcone' gauge), but practical calculations are still considerably more complicated than in covariant gauges. In particular, if making a numerical calculation, it is difficult to guarantee that the spurious singularities $n\cdot p \to 0$ cancel as they should.

We therefore have the Feynman rules (in Feynman gauge):

$$\Delta_i = \frac{i}{\not p - m_i} = i\frac{\not p + m_i}{p^2 - m_i^2}, \tag{1.19}$$

$$\Delta_{\gamma,\mu\nu} = i\frac{-g_{\mu\nu}}{p^2}, \tag{1.20}$$

$$\Gamma^\mu_{\gamma f_i \bar f_i} = -i\, e_i e \gamma^\mu. \tag{1.21}$$

To calculate the cross section for a given process, we must write down all possible diagrams, use the Feynman rules to give us the amplitude $i\mathcal{M}$, use Dirac algebra and trace theorems to calculate $\sum|\mathcal{M}|^2$, where the sum is over all unobserved quantum numbers for example spin, divide by the overcounting of incoming states, and integrate over phase space:

$$\sigma = \frac{1}{S}\,\frac{1}{2s}\int d\Gamma\,\sum|\mathcal{M}|^2. \tag{1.22}$$

An element of n-body phase space is given by

$$d\Gamma = \prod_{i=1}^{n} \left(\frac{d^4 p_i}{(2\pi)^4} \, (2\pi)\delta(p_i^2 - m_i^2) \right) (2\pi)^4 \delta^4 (p_{tot} - \textstyle\sum_i^n p_i) \tag{1.23}$$

$$= \prod_{i=1}^{n} \left(\frac{d^3 p_i}{(2\pi)^3 2E_i} \right) (2\pi)^4 \delta^4 (p_{tot} - \textstyle\sum_i^n p_i). \tag{1.24}$$

For example, the cross section for $e^+e^- \to \mu^+\mu^-$ is calculated as follows. The amplitude is

$$i\mathcal{M} = \bar{v}(p_{e^+})(ie)\gamma^\mu u(p_{e^-}) \, i\frac{-g_{\mu\nu}}{(p_{e^+} + p_{e^-})^2} \, \bar{u}(p_{\mu^-})(ie)\gamma^\nu v(p_{\mu^+}) \tag{1.25}$$

$$= \frac{-ie^2}{(p_{e^+} + p_{e^-})^2} \bar{v}(p_{e^+})\gamma^\mu u(p_{e^-}) \, \bar{u}(p_{\mu^-})\gamma_\mu v(p_{\mu^+}) \tag{1.26}$$

and hence

$$\sum |\mathcal{M}|^2 = \frac{(4\pi\alpha)^2}{s^2} \, \mathrm{Tr}\left\{ \not{p}_{e^+}\gamma^\mu \not{p}_{e^-}\gamma^\nu \right\} \, \mathrm{Tr}\left\{ \not{p}_{\mu^-}\gamma_\mu \not{p}_{\mu^+}\gamma_\nu \right\}, \tag{1.27}$$

where $\alpha = e^2/4\pi$ and $s = (p_{e^+} + p_{e^-})^2$, or

$$\sum |\mathcal{M}|^2 = \frac{16(4\pi\alpha)^2}{s^2} \, \left(p_{e^+}^\mu p_{e^-}^\nu + p_{e^-}^\mu p_{e^+}^\nu - p_{e^+}\cdot p_{e^-} g^{\mu\nu} \right) \, \left(p_{\mu^-,\mu} p_{\mu^+,\nu} + p_{\mu^+,\mu} p_{\mu^-,\nu} - p_{\mu^+}\cdot p_{\mu^-} g_{\mu\nu} \right) \tag{1.28}$$

$$= 8(4\pi\alpha)^2 \frac{t^2 + u^2}{s^2}, \tag{1.29}$$

where $t = (p_{e^-} - p_{\mu^-})^2$ and $u = (p_{e^-} - p_{\mu^+})^2 = -s - t$. The cross section is therefore

$$\sigma = \frac{1}{4} \, \frac{1}{2s} \int_{-s}^{0} \frac{dt}{8\pi s} \, 8(4\pi\alpha)^2 \frac{t^2 + u^2}{s^2} \tag{1.30}$$

$$= \frac{4\pi\alpha^2}{3s}. \tag{1.31}$$

1.3 SU(3) and colour

QCD can be derived in exactly the same way as QED: we start from the Lagrangian density for a set of non-interacting quarks and modify it in just such a way that it is invariant under changes of gauge. The only difference is that instead of the gauge transformation being a simple phase (U(1) group), we consider a non-Abelian group SU(N_c). This has several important consequences. Fermion charges will come in N_c different types, called colours, they will be quantized (in contrast to the electric charges e_i, which could take any value) and, most importantly, the gauge bosons will be self-interacting.

It has been well-known since the early days of QCD that there are three colours, for example from baryon wave functions, the total e^+e^- cross section (which is proportional to N_c) and π^0 decay rate (which is proportional to N_c^2). However, in most calculations it is useful to keep the number of colours N_c arbitrary until the very last step when it is set equal to three. The N_c-dependent coefficients are a useful diagnostic tool in understanding the physical origins of different terms, comparing different calculations and tracking down errors.

We start by restating briefly some features of SU(N), the group of $N \times N$ unitary matrices ($U^\dagger U = 1$) with determinant +1. Let U be an element of SU(N) that is infinitesimally close to the identity and write it as

$$U = 1 + iG, \tag{1.32}$$

where G has infinitesimal entries. It must be hermitian ($G^\dagger = G$) and traceless. One can choose a basis set of N^2-1 matrices, t^A, $A = 1, \ldots, N^2-1$, such that any G can be written as

$$G = \sum_A^{N^2-1} \epsilon^A t^A, \tag{1.33}$$

where ϵ_A are infinitesimal numbers. Note that I will always denote colour indices that run from 1 to N by a and from 1 to N^2-1 by A. The t^A are called the generators of the group and define its fundamental representation. You can show that $[t^A, t^B]$ is antihermitian and traceless and hence can be written as a linear combination of other t^Cs,

$$[t^A, t^B] \equiv i\, f^{ABC} t^C, \tag{1.34}$$

where f^{ABC} are a set of real constants, called the structure constants of the group. It is straightforward to see that f^{ABC} is antisymmetric in A, B, and with a little more work, one can prove that it is antisymmetric in all its indices. Equation (1.34) defines the Lie algebra of the group.

We can also define a set of $(N^2-1) \times (N^2-1)$ matrices that obey the same algebra:

$$\left(T^A\right)_{BC} \equiv -i f^{ABC}, \tag{1.35}$$

$$[T^A, T^B] = i\, f^{ABC} T^C. \tag{1.36}$$

These define the group's adjoint representation.

Although we started with elements infinitesimally close to the identity matrix, we can calculate an arbitrary element U by stringing together an infinite number of infinitesimal elements,

$$U = \lim_{N\to\infty} (1 + i\theta^A t^A/N)^N = \exp(i\theta^A t^A) \equiv \exp(it\cdot\theta). \tag{1.37}$$

Since U is unitary and t^A hermitian, we have

$$U^{-1} = \exp(-it\cdot\theta). \tag{1.38}$$

There are several identities we will require time and time again:

$$\mathrm{Tr}(t^A t^B) = \tfrac{1}{2}\delta^{AB} \equiv T_R \delta^{AB} \tag{1.39}$$

$$\sum_A t^A_{ab} t^A_{bc} = \frac{N^2-1}{2N}\delta_{ac} \equiv C_F \delta_{ac} \tag{1.40}$$

$$\mathrm{Tr}(T^C T^D) = \sum_{A,B} f^{ABC} f^{ABD} = N\delta^{CD} \equiv C_A \delta^{CD}, \tag{1.41}$$

where the constants C_F and C_A are the Casimir operators of the fundamental and adjoint representations of the group respectively. Although we know the numerical values of these constants:

$$T_R = \frac{1}{2}, \tag{1.42}$$

$$C_F = \frac{4}{3}, \tag{1.43}$$

$$C_A = 3, \tag{1.44}$$

it is good practice, as I said, to leave them unexpanded in all algebraic results.

In fact for practical calculations one only requires these, and other similar, identities and never an explicit representation for t^A or f^{ABC}.

1.4 The QCD Lagrangian

The classical Lagrangian density for n non-interacting quarks with masses m_i is

$$\mathcal{L}_{\text{quarks}} = \sum_i^n \bar{q}_i^a (i\partial\!\!\!/ - m_i)_{ab} q_i^b, \tag{1.45}$$

where the factor $(i\partial\!\!\!/ - m_i)_{ab}$ is proportional to the identity matrix in colour space. This is invariant under global SU(N_c) transformations,

$$q_a \to q'_a = \exp(it\cdot\theta)_{ab} q_b. \tag{1.46}$$

To make it invariant under local transformations,

$$q_a(x) \to q'_a(x) = \exp(it\cdot\theta(x))_{ab} q_b(x), \tag{1.47}$$

we have to introduce the covariant derivative,

$$D_{\mu,ab} = \partial_\mu 1_{ab} + ig_s\,(t\cdot A_\mu)_{ab}, \tag{1.48}$$

where A_μ^A are coloured vector fields that transform in just the right way that we have

$$D'_{\mu,ab} q'_b(x) = \exp(it\cdot\theta(x))_{ab} D_{\mu,bc} q_c(x), \tag{1.49}$$

giving

$$t\cdot A'_\mu = \exp(it\cdot\theta(x))\, t\cdot A_\mu \,\exp(-it\cdot\theta(x)) + \frac{i}{g_s}\Big(\partial_\mu \exp(it\cdot\theta(x))\Big)\exp(-it\cdot\theta(x)). \tag{1.50}$$

We again have to introduce a kinetic term for this new field,

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4} F_{\mu\nu}^A F_A^{\mu\nu}, \tag{1.51}$$

where $F_{\mu\nu}^A$ is the non-Abelian field strength tensor. However, the definition we used in QED (1.9) does not result in an invariant Lagrangian density under transformation (1.50). One must add an extra term,

$$F_{\mu\nu}^A = \partial_\mu A_\nu^A - \partial_\nu A_\mu^A - g_s f^{ABC} A_\mu^B A_\nu^C, \tag{1.52}$$

and only then is (1.51) invariant under gauge transformations.

This extra term has profound consequences for the theory: it means that gluons are self-interacting, through three- and four-point vertices. This will turn out to give rise to asymptotic freedom at high energies and strong interactions at low energies, among the most fundamental properties of QCD. We therefore see that these are absolute requirements of the SU(N_c) gauge symmetry.

Before reading off the Feynman rules we again have to fix the gauge. This proceeds in exactly the same way as in QED, leading to, in covariant gauges,

$$\mathcal{L}_{\text{gauge-fixing}} = -\frac{1}{2\lambda}\left(\partial^\mu A_\mu^A\right)^2. \tag{1.53}$$

Finally, it turns out that in a non-Abelian gauge theory, it is necessary to add one extra term to the Lagrangian density, related to the need for ghost particles. These are beyond the scope of this course, but basically they arise because when a non-Abelian gauge theory is renormalized it is possible for unphysical degrees of freedom to propagate freely. These are cancelled off by introducing into the theory an unphysical set of fields, the ghosts, which are scalars but have Fermi statistics. For practical purposes it is enough to know that there exist Feynman rules for ghosts and that in every diagram with a closed loop of internal gluons containing only triple-gluon vertices, we must add a diagram with the gluons in the loop replaced by ghosts. It is worth noting that in physical gauges, as the name suggests, ghost contributions always vanish and they can be ignored.

The final Lagrangian is therefore

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} F_{\mu\nu}^A F_A^{\mu\nu} + \sum_i^n \bar{q}_i^a (iD\!\!\!\!/ - m_i)_{ab} q_i^b - \frac{1}{2\lambda}\left(\partial^\mu A_\mu^A\right)^2 + \mathcal{L}_{\text{ghost}}. \tag{1.54}$$

1.5 Feynman rules

Just as in QED it is straightforward to read off the Feynman rules from the action. We obtain in Feynman gauge (only the gluon propagator is gauge dependent)

$$
\begin{aligned}
\Delta_i^{ab} &= \delta^{ab}\frac{i}{\not{p}-m_i} = \delta^{ab} i \frac{\not{p}+m_i}{p^2-m_i^2}, &(1.55)\\
\Delta_{g,\mu\nu}^{AB} &= \delta^{AB} i \frac{-g_{\mu\nu}}{p^2}, &(1.56)\\
\Gamma_{gq\bar{q}}^{\mu} &= -i\, g_s\, t^A\, \gamma^\mu, &(1.57)\\
\Gamma_{ggg} &= -g_s f^{ABC}\left[(p-q)^\lambda g^{\mu\nu} + (q-r)^\mu g^{\nu\lambda} + (r-p)^\nu g^{\lambda\mu}\right]. &(1.58)
\end{aligned}
$$

Note that, apart from the triple-gluon vertex, the only difference relative to QED is in the colour structure: propagators are diagonal in colour and the vertex for a gluon of colour A to scatter a quark of colour b to a quark of colour c contains $(t^A)_{cb}$. Note also that unlike QED the quark–gluon vertex is flavour-independent (it is straightforward to check that, unlike in QED, we cannot introduce a flavour-dependence into the gauge transformation, Eq. (1.47) and retain gauge invariance). In the triple-gluon vertex, the three gluons have momenta p, q, r, Lorentz indices μ, ν, λ and colour indices A, B, C respectively. The momenta are all ingoing: $p + q + r = 0$.

The Feynman rules for ghosts and for the four-gluon vertex can be found in ESW [1] (p. 10). They will not be needed for this course.

Note also that in analogy with QED the strong charge g_s is usually substituted by α_{S},

$$\alpha_{\mathrm{S}} \equiv \frac{g_s^2}{4\pi}. \tag{1.59}$$

1.6 $e^+e^- \to q\bar{q}$

One of the most fundamental quantities in QCD is the total e^+e^- annihilation cross section to hadrons. We will see in a later lecture that to leading order in α_{S} this is equal to the total $e^+e^- \to q\bar{q}$ cross section. The calculation is very similar to that for $e^+e^- \to \mu^+\mu^-$, the only difference being in the colour structure. The photon is colour blind, so the Feynman rule for a photon to couple to a quark contains a trivial colour matrix, δ^{ab}. Summing over colours and dividing by the number of incoming colour states (1 in this case since electrons are not coloured), we therefore obtain

$$\sigma(e^+e^- \to q\bar{q}) = \sigma(e^+e^- \to \mu^+\mu^-) \times e_q^2 \times \sum_{a,b} \delta^{ab}\delta^{ba}. \tag{1.60}$$

We obtain

$$\sum_{a,b} \delta^{ab}\delta^{ba} = \sum_a \delta^{aa} = N_c, \tag{1.61}$$

and hence

$$R_{\mathrm{had}} \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = \sum_q e_q^2 N_c. \tag{1.62}$$

1.7 $e^+e^- \to q\bar{q}g$

This process will be important for the higher order corrections to $\sigma(e^+e^- \to \text{hadrons})$ and particularly for the study of three-jet final states in e^+e^- annihilation, among the most important test-beds for QCD.

There are two Feynman diagrams, shown in Fig. 1.1. We label the momenta and colours $e^-(p_-)+$

Fig. 1.1: Feynman diagrams for the process $e^+e^- \to q\bar{q}g$

$e^+(p_+) \to q_a(p_1) + \bar{q}_b(p_2) + g_A(p_3)$. For the matrix element we obtain

$$\begin{aligned} i\mathcal{M} &= \bar{v}(p_+)(ie)\gamma^\mu u(p_-)\; i\frac{-g_{\mu\nu}}{s}\; \varepsilon_A^{*\lambda} \qquad (1.63)\\ &\quad \bar{u}_a(p_1)\left\{(-ig_s)t^A_{ab}\gamma^\lambda \frac{\not{p}_1+\not{p}_3}{(p_1+p_3)^2}(-iee_q)\gamma^\nu + (-iee_q)\gamma^\nu \frac{-\not{p}_2-\not{p}_3}{(p_2+p_3)^2}(-ig_s)t^A_{ab}\gamma^\lambda\right\} v_b(p_2). \end{aligned}$$

We will evaluate the cross section from this matrix element later. Here we are interested in the colour algebra. Using the fact that the spin sum of a massless vector particle is proportional to the colour identity matrix,

$$\sum_{\text{spin}} \varepsilon_A^{*\mu}\varepsilon_B^\nu = -g^{\mu\nu}\delta_{AB}, \qquad (1.64)$$

we obtain

$$\sum|\mathcal{M}|^2 \propto \sum_{a,b,A} t^A_{ab}\left(t^A_{ab}\right)^* = \sum_{a,b,A} t^A_{ab}t^A_{ba} = \sum_A \mathrm{Tr}(t^At^A) = C_F\mathrm{Tr}(1) = C_FN_c, \qquad (1.65)$$

where the first step uses the fact that t^A are hermitian, the second is simply a trivial rewrite, switching to matrix notation, the third uses Eq. (1.40) and the fourth uses the fact that the matrix being traced is the identity matrix of the fundamental representation, i.e. the $N_c \times N_c$ identity matrix. Note that since the colour factor of the lowest order process is N_c, we can associate C_F with the emission of the additional gluon. Since the emission probability of a gluon from a quark is proportional to C_F, and we will later see that that from a gluon is proportional to C_A, C_F and C_A are sometimes referred to as the squares of the colour charges of the quark and gluon respectively.

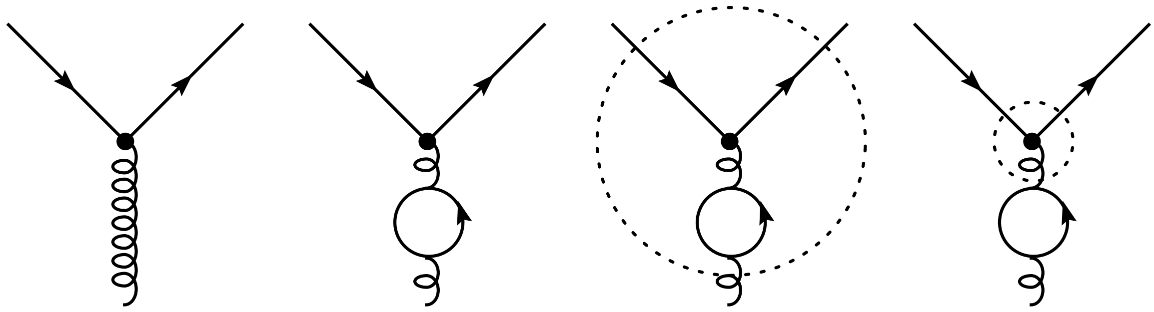
Performing the trace Dirac algebra on the matrix element, we finally obtain

$$\sum|\mathcal{M}|^2 = \frac{16C_FN_ce^4e_q^2g_s^2}{s\,p_1\cdot p_3\,p_2\cdot p_3}\left((p_1\cdot p_+)^2+(p_2\cdot p_+)^2+(p_1\cdot p_-)^2+(p_2\cdot p_-)^2\right). \qquad (1.66)$$

(Note the misprint in ESW [1] — their result is a factor of 4 too large.)

1.8 The coupling constant α_S and renormalization

As we mentioned above, in practical calculations, α_S is usually used rather than g_s. Besides the quark masses, which we will neglect in most of this course, g_s is the only parameter in the QCD Lagrangian and therefore assumes a central role in our study of QCD. However, it is not *a priori* clear that parameters in the Lagrangian are physically observable quantities — any physical observable can be calculated as a function of them (at least in perturbation theory) and their values can be extracted from measured values of physical observables, but they are not necessarily themselves physical. It is worthwhile therefore to consider whether we can reformulate our theory in such a way that one physical observable can be written as a function of another. This reformulation is known as renormalization.



(a) (b) (c) (d)

Fig. 1.2: When a quark–gluon vertex (a) is corrected by a loop (b), one must decide whether to describe it as a correction to the vertex (c), or to the rest of the diagram (d)

In this section I give a very handwaving description of renormalization, which I believe conveys the important physical point. Of course for practical calculations one needs a much more precise definition of the renormalization prescription, which I describe at the end.

We redefine g_s to be the strength of the quark–gluon coupling, as in Fig. 1.2a. At first sight, this seems like a trivial statement and at the lowest order of perturbation theory it is — the two definitions are identical. However, when we calculate higher orders of perturbation theory, we encounter loop corrections like the one in Fig. 1.2b, which correct the vertex. To avoid double-counting, we must uniquely decide whether these corrections are part of the vertex, as in Fig. 1.2c, or the rest of the diagram, as in Fig. 1.2d. One way to decide is to introduce a *renormalization scale* μ_R and say that physics at high scales (therefore short distances) above μ_R is part of the vertex and physics at lower scales (longer distances) below μ_R is part of the rest of the diagram. Of course, this is simply a book-keeping device, which does not change the physics, it simply ensures that each physical contribution to the process is counted once and only once. Since μ_R is a completely arbitrary book-keeping scale, introduced by hand, its value should not affect the physical prediction — changing it simply moves contributions between what we call the vertex and what we call the rest of the diagram. Since the amount of physics that we include in the vertex depends on μ_R, and we defined g_s to be the strength of the vertex, it is clear that g_s must now be a function of μ_R.

It is worth mentioning that, although I defined g_s as the strength of the quark–gluon vertex, I could equally well have defined it as the strength of the triple-gluon vertex. It is one of the remarkable features of gauge theories that, as a direct result of the gauge symmetry, I would get exactly the same result for the renormalized coupling $g_s(\mu_R)$. That is, the equality of the strengths of the quark–gluon and triple-gluon vertices is true even after renormalization.

When it is clear that I am talking about the renormalization scale, I will henceforth drop the subscript $_R$.

1.8.1 Renormalization group equation

As I said, varying μ moves physical contributions (loop corrections) around within a calculation, but it should not change the result of the physical calculation. We can use this fact to derive an equation for how g_s varies as a function of μ. This is one of a set of equations that together describe how the whole theory varies with renormalization scale (and scheme), which formally form a group.

We study this by considering a dimensionless physical observable R that is a function of only one physical scale Q^2 (think of $R_{\rm had}$ at $\sqrt{s} = Q$ for example). Assume that this observable is not sensitive to quark masses (we will return to this point shortly). After renormalization, R can only be a function of Q^2, μ^2 and $\alpha_{\rm S}(\mu^2)$. By dimensional analysis, the only way the dimensionless function R can depend on the dimensionful variables Q^2 and μ^2 is through their ratio. We can therefore write

$$R = R(Q^2/\mu^2, \alpha_S(\mu^2)). \tag{1.67}$$

We can use the fact that R, as a physical quantity, must be independent of the value of μ, and the chain rule for partial derivatives, to write

$$\mu^2 \frac{d}{d\mu^2} R(Q^2/\mu^2, \alpha_S) = 0 \quad = \quad \left[\mu^2 \frac{\partial}{\partial\mu^2} + \mu^2 \frac{\partial\alpha_S}{\partial\mu^2} \frac{\partial}{\partial\alpha_S}\right] R \tag{1.68}$$

$$\equiv \quad \left[\mu^2 \frac{\partial}{\partial\mu^2} + \beta(\alpha_S) \frac{\partial}{\partial\alpha_S}\right] R\,, \tag{1.69}$$

i.e., $\beta(\alpha_S) \equiv \mu^2 \frac{\partial\alpha_S}{\partial\mu^2}$. There are several points to note about this.

- A physical solution is provided by $R(1, \alpha_S(Q))$, i.e., by setting the renormalization scale equal to the physical scale in the problem.
- Q-dependence of the physical quantity R comes about only because of the renormalization of the theory and would not be present in the classical theory. Thus measuring the Q-dependence of R directly probes the quantum structure of the theory.
- By rearranging Eq. (1.69), one can derive the μ^2 dependence of α_S from a calculation of R,

$$\beta(\alpha_S) = -\frac{\mu^2 \frac{\partial R}{\partial\mu^2}}{\frac{\partial R}{\partial\alpha_S}}. \tag{1.70}$$

- If α_S is small, R is perturbatively calculable and hence $\beta(\alpha_S)$ is too.

The β function of QCD is now known to four-loop accuracy,

$$\beta(\alpha_S) = -\alpha_S^2(\beta_0 + \beta_1\alpha_S + \beta_2\alpha_S^2 + \beta_3\alpha_S^2 + \ldots). \tag{1.71}$$

Although the higher orders are essential for quantitative calculation, they are not for qualitative understanding: almost all QCD phenomenology can be understood using the one loop result,

$$\beta_0 = \frac{11C_A - 4T_R N_f}{12\pi}, \tag{1.72}$$

where N_f is the number of quark flavours that can appear in loops, to be discussed further shortly.

Note that β_0 is positive and hence that the β function is negative, at least when α_S is small. This results in asymptotic freedom: the fact that the interactions become weak at high energies (short distances) and infrared slavery: the fact that they become strong at low energy.

If we neglect the higher orders, we can solve the renormalization group equation exactly, to obtain α_S at some scale Q as a function of its value at the renormalization scale μ,

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\beta_0 \ln\frac{Q^2}{\mu^2}}. \tag{1.73}$$

1.8.2 Choosing μ^2

Although physical quantities do not depend on μ, a calculation truncated at a finite order of perturbation theory does. We must therefore choose some value for μ. To illustrate this, suppose that our dimensionless physical quantity R has a perturbative expansion that starts at $\mathcal{O}(\alpha_S)$,

$$R = R_1\alpha_S + \ldots, \tag{1.74}$$

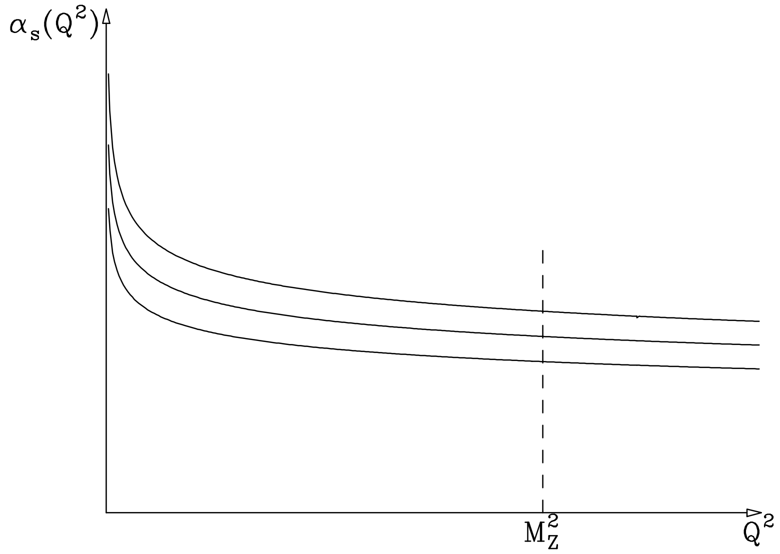


Fig. 1.3: A measurement of α_S at any scale Q fixes which curve our universe lies on, but to compare measurements at different scales we have to agree to label the curves in a standard way, for example using $\alpha_S(M_z)$

then if we truncate at leading order,

$$R \approx R_1 \alpha_S, \tag{1.75}$$

our truncated expression for $R(1, \alpha_S(Q))$ can be expanded as a power series in $\alpha_S(\mu^2)$

$$\begin{aligned} R(1, \alpha_S(Q^2)) &\approx R_1 \alpha_S(Q^2) &(1.76)\\ &= R_1 \alpha_S(\mu^2) \left[1 - \beta_0 \alpha_S(\mu^2) \ln \frac{Q^2}{\mu^2} + \beta_0^2 \alpha_S^2(\mu^2) \ln^2 \frac{Q^2}{\mu^2} + \dots \right]. &(1.77) \end{aligned}$$

The leading order result in renormalized perturbation theory is the first term of this series, i.e., $R_1 \alpha_S(\mu^2)$. It is therefore clear that although μ is completely arbitrary, choosing it far from Q guarantees a large truncation error (note that the converse is not true). One should therefore choose μ^2 'close' to Q^2, but how close is close?

The conventional approach is to set $\mu = Q$ and to use the μ variation in a reasonable range, e.g., $Q/2$ to $2Q$ as an estimate of the truncation uncertainty. It should be clear from the foregoing discussion that this is an extremely arbitrary procedure. However, the folklore is that in almost all cases where higher order corrections have been calculated, they have fallen within the band given by this procedure.

1.8.3 Measuring α_S

The β function tells us how α_S varies with scale, but it does not tell us the value of α_S at any particular scale: we need an experimental measurement to do that. Effectively $\beta(\alpha_S)$ defines a family of curves, as illustrated in Fig. 1.3, and one measurement at any scale is sufficient to tell us which curve our universe lies on. However, in order to compare and combine measurements of α_S at different scales, we have to agree on some convenient labeling of the curves. The measurement at any given scale can then be converted into a measurement of the label. Historically, this was often done using the 'QCD scale', Λ_{QCD}, described in the next section, but more recently it has been realized that the value of α_S at some fixed scale at which it is relatively small is a lot more convenient. Since some of the best measurements come from Z^0 decays, it has become universal to use $\alpha_S(M_z)$ as the label. We will discuss the measurements of α_S further in Section 3.1.4.

1.8.4 The 'QCD Scale', Λ

As I just mentioned, this is another way to label the running coupling, which is to construct a renormalization group invariant scale from $\alpha_{\mathrm{S}}(\mu)$. Although the Lagrangian of massless QCD has no scale, the renormalization process introduces a dimensionful parameter,

$$\Lambda^2 = \mu^2 \exp - \int^{\alpha_{\mathrm{S}}(\mu^2)} \frac{dx}{\beta(x)} \approx \mu^2 \mathrm{e}^{-1/\beta_0 \alpha_{\mathrm{S}}(\mu^2)}, \tag{1.78}$$

where the approximation uses only the one-loop term in the β function[1.1]. This process by which a scaleless theory gets a physically observable scale by the introduction of the unphysical renormalization scale is known as dimensional transmutation.

At leading order, Λ has a simple interpretation, it is the scale at which the coupling becomes infinite. However, this interpretation is not self-consistent, since it relies on a truncation of the perturbation series in a region in which the coupling is large, ultimately divergent. More generally, Λ can be viewed as a renormalization group invariant parameterization of the scale at which the theory becomes non-perturbative. All non-perturbative quantities, for example the hadron masses, would be expected to be of order Λ.

However, Λ is not a very practically useful label for the value of α_{S}. This is because its precise value, for a given measured value of α_{S}, depends strongly on the theoretical input used in the calculation, for example which order of perturbation theory we truncate β at, which renormalization scheme we use, the number of flavours we assume, the way we match the running coupling at the flavour thresholds, etc.

In principle any labeling suffers from these problems, but by using the value of α_{S} in a region where it is small, and where the scale is not too different from that at which the measurements are made, the impact on $\alpha_{\mathrm{S}}(M_z)$ is small, whereas Λ is related to the region where α_{S} is large, far away from where the measurements are made, and these effects are large.

1.8.5 Renormalization in practice

To give a simple physical picture of renormalization, I have described it in terms of a cutoff on the scale of the physical effects that are included in different components of a Feynman diagram calculation. However, in practice, this definition is extremely unattractive, because it breaks Lorentz and gauge invariance, two of the fundamental symmetries of our theory. If calculating in this scheme, these symmetries will get violated by a truncation at any finite order of perturbation theory and only restored in an all orders calculation. There are other simple schemes that work well in certain cases, for example the Pauli–Villars regularization, but the only known scheme consistent with all the symmetries of QCD, and hence guaranteed to work at any order of perturbation theory, is *dimensional regularization*. In this section I give a very brief description of how this works in practice. The difference between μ and μ_R will be (slightly) relevant here, so I temporarily reinstate the subscript.

The basic observation is that the loop corrections that we have been discussing are divergent in four or more space-time dimensions, but are finite in less than four dimensions. We therefore choose to calculate our Feynman diagrams in $d < 4$ dimensions (we always work in Minkowski space, with one time dimension and $d-1$ space dimensions). With a little thought, we can analytically continue

[1.1] Note that the definition in the first equality of Eq. (1.78), while formally renormalization group invariant, is not practically useful, since the lower limit of the integration is not defined (corresponding to the fact that any definition of Λ that differs by a multiplicative constant is equally renormalization group invariant). For perturbative calculations, various definitions, equivalent to Eq. (1.78) to the order to which they are defined, can be used. For non-perturbative calculations, for example in lattice QCD, the precise definition is more critical. A commonly-used convention (see for example [3]) is $\Lambda^2 = \mu^2 \exp\left\{ -\frac{1}{\beta_0 \alpha_{\mathrm{S}}(\mu^2)} - \frac{\beta_1}{\beta_0^2} \log \alpha_{\mathrm{S}}(\mu^2) - \int_0^{\alpha_{\mathrm{S}}(\mu^2)} dx \left(\frac{1}{\beta(x)} + \frac{1 - \frac{\beta_1}{\beta_0} x}{\beta_0 x^2} \right) \right\}$. In contrast to the definition given in [1], for example, this can be seen to depend only on the β function at $\alpha_{\mathrm{S}}(\mu^2)$ and at smaller values, so is well-defined perturbatively and, as can be easily checked, is exactly renormalization scheme invariant.

the number of dimensions to be a complex number such that at the end of the calculation, after the renormalization prescription has been followed, we can let it smoothly tend back to 4 and obtain finite results. We therefore define $d = 4 - 2\epsilon$ and consider the $\epsilon \to 0_+$ limit.

By counting the dimensionality of terms in the Lagrangian, we discover that the coupling constant becomes dimensionful in $d \neq 4$ dimensions. This is not very convenient, so we define a dimensionless parameter α_{S}, by introducing a completely arbitrary scale μ,

$$\alpha_{\mathrm{S}}^{(d)} = \alpha_{\mathrm{S}}\,\mu^{2\epsilon}, \tag{1.79}$$

where $\alpha_{\mathrm{S}}^{(d)}$ is the dimensionful d-dimensional coupling. μ is called the regularization scale. It is often set equal to the renormalization scale μ_R, but I consider this confusing since we have not yet renormalized the theory, so, for now, I keep them distinct and only set them equal again at the end of this section.

When calculating loop corrections, we then find terms that have $1/\epsilon$ singularities for small ϵ. These have the right form to be absorbed by a redefinition (i.e. a renormalization) of the coupling. Since we also want the renormalized coupling to be dimensionless, we have to introduce a dimensionful scale at which the renormalization is performed, μ_R. To make this concrete, at one-loop order, the prescription is straightforward: after calculating all the one-loop diagrams, rewrite all occurrences of α_{S} in terms of the renormalized coupling,

$$\alpha_{\mathrm{S}}(\mu_R) = \alpha_{\mathrm{S}} + \beta_0\,F(\epsilon)\,\left(\frac{\mu^2}{\mu_R^2}\right)^{\epsilon}\,\frac{1}{\epsilon}\,\alpha_{\mathrm{S}}^2\,. \tag{1.80}$$

Provided $F(0) = 1$, once this substitution has been made, the amplitude is finite. That is, the ϵ poles that this expression produces exactly cancel those from the one-loop calculation. Moreover, the arbitrary scale μ cancels from the amplitude at this point. One is left with a finite amplitude that depends only on μ_R and $\alpha_{\mathrm{S}}(\mu_R)$, in exactly the same way as discussed earlier.

The arbitrary function $F(\epsilon) = 1 + \mathcal{O}(\epsilon)$ defines the renormalization scheme. More precisely, it defines what finite parts of the loop amplitude are subtracted into the renormalized coupling, in addition to the divergent part. The MS, or minimal subtraction, scheme, is defined by subtracting nothing else,

$$F_{\mathrm{MS}}(\epsilon) = 1. \tag{1.81}$$

The most commonly used scheme is the $\overline{\mathrm{MS}}$, or modified minimal subtraction, scheme, in which one identifies some additional overall factors coming from the analytical continuation of the angular integrations in the one-loop calculation. Since they are universal it is convenient to subtract them into the coupling,

$$F_{\overline{\mathrm{MS}}}(\epsilon) = \frac{(4\pi)^{\epsilon}}{\Gamma(1-\epsilon)} = 1 + (\ln 4\pi - \gamma_E)\epsilon, \tag{1.82}$$

where Γ is the Euler gamma function and γ_E the Euler gamma constant, $\gamma_E \approx 0.577216$. Note that the two expressions on the right-hand side of Eq. (1.82) differ at order ϵ^2. Different practitioners use either of the two definitions, resulting in a finite difference at two loops that is straightforward to keep track of.

1.9 Quark masses and decoupling

The quark masses m_q are also parameters of the Lagrangian and face the same issues: for a physical calculation we should redefine them in a physical way. For the electron mass, we have a simple definition: we can isolate a single electron and 'weigh' it in the laboratory. That is, we can define its mass through the classical limit. We cannot use the same procedure for quarks, because confinement means that we can never take a single quark off to our laboratory to weigh it individually. We must therefore define some other renormalization procedure.

It is possible to proceed in close analogy with the coupling strength. We renormalize our theory at the same scale μ. We encounter gluon loop corrections to the quark propagator and absorb the part of them at scales above μ into the definition of the mass. We therefore obtain a 'running' (i.e. scale-dependent) mass. Just like for the coupling, we can obtain a renormalization group equation with perturbatively-calculable coefficients,

$$\frac{\mu^2}{m}\frac{dm}{d\mu^2} = -\frac{1}{\pi}\,\alpha_{\rm S}(\mu^2) + \dots\,. \tag{1.83}$$

At leading order it can be solved exactly, to give

$$m(\mu^2) = M\left[\alpha_{\rm S}(\mu^2)\right]^{\frac{1}{\pi\beta_0}}, \tag{1.84}$$

where M is a renormalization group invariant constant (c.f. Λ_{QCD}). Note that increasing μ^2 decreases m^2. Thus quarks appear to get lighter as they are probed at scales further and further above their masses.

An alternative scheme, which is often used in electroweak physics, and in the physics of heavy mesons, is the *pole mass*. Here one defines m_q to be the pole of the propagator $i(\not{p}+m_q)/(p^2-m_q^2)$ to all orders. This is very useful for $Q \sim m_q$, but it turns out that it is similar to a running mass scheme with μ of order m_q and hence generates large logarithms and a large truncation error for $Q \gg m_q$.

If our dimensionless observable R is finite for massless quarks then the quark mass effects must vanish smoothly as the mass goes to zero. Therefore the mass effects must be suppressed by $(m_q/Q)^n$, with $n \geq 1$. If there are quarks with mass much greater than Q, they can only affect our observable through loop corrections. A dimensional argument shows that such corrections must be suppressed by $(Q/m_q)^n$, with $n \geq 2$.

These observations form the basis of the decoupling theorem, in which quarks heavier than our physical scale can be ignored, and quarks lighter than it can be treated as massless. Thus, for most QCD calculations, we work with N_f flavours of massless quark (recall the N_f that appeared in β_0). Care must be taken when Q is close to a quark mass, or we study a range of processes at scales that span a quark mass, but in fact for most of the phenomenology considered in this course we can simply take N_f to be fixed, $N_f = 5$.

1.10 Summary

We have seen that QCD is a gauge theory. The fact that the gauge symmetry is non-Abelian predicts that the gluon is self-interacting. This leads to the fact that the theory becomes strongly interacting at low energies, and hence non-perturbative, and weakly interacting at high energies so that perturbation theory can be used.

The main tools that we will use to study QCD are the *factorization* of non-perturbative effects and the *renormalization* and *decoupling* of high-energy physics. These allow us to use perturbation theory and, in particular, the Feynman rules, to study the phenomenology of QCD.

2 QCD phenomenology at tree level

Leading order perturbation theory, together with the one-loop renormalization group equation is enough to understand a wide variety of QCD phenomenology. In this section, we briefly review the phenomenology of QCD before introducing the complications of loop corrections to it in the following section. Most of the salient ideas are introduced in the context of e^+e^- annihilation and deep inelastic scattering, but apply equally well to hadron collisions and photoproduction, which we discuss more briefly at the end.

2.1 The cross section for $e^+e^- \to$ hadrons

One of the most striking features of e^+e^- annihilation events is the fact that many of them produce many hadrons. In trying to calculate the cross section for this process, however, we are immediately faced with a problem: the Lagrangian does not contain any information about hadrons, so there are no Feynman rules involving them. Even if there were, calculating all the diagrams for events involving thirty or forty particles would be prohibitively complicated, let alone integrating them over the corresponding phase space to produce a total cross section. Fortunately a simple application of the Feynman rules of QED, together with some simple symmetry arguments, allows us to make a surprisingly strong statement about the cross section for e^+e^- annihilation to hadrons.

We postulate that the matrix element for the sum of all diagrams in which a virtual photon with Lorentz index ν and momentum q produces a particular set of n hadrons with momenta $\{p_1 \dots p_n\}$ is known and parameterize it by a function $T_\nu(n, q, \{p_1 \dots p_n\})$. Using this function, it is straightforward to write down the matrix element for the full process,

$$\mathcal{M} = \{\bar{v}(q_2) e \gamma_\mu u(q_1)\} \; \frac{-g^{\mu\nu}}{q^2} \; T_\nu(n, q, \{p_1 \dots p_n\}) \tag{2.1}$$

and hence the phase-space integral for its total cross section. The total cross section to produce any number of any type of hadrons is then simply given by the sum of this integral over hadron type and multiplicity (both generically represented by $\sum_n$),

$$\begin{aligned} \sigma \;\; &= \;\; \frac{1}{2s} \, \frac{1}{4} \, \frac{e^2}{s^2} \mathrm{Tr}(\not{q}_2 \gamma^\mu \not{q}_1 \gamma^\nu) & (2.2) \\ & \times \sum_n \int dPS_n \; T_\mu(n, q, \{p_1 \dots p_n\}) \; T^*_\nu(n, q, \{p_1 \dots p_n\}). & (2.3) \end{aligned}$$

We then define a new two-index tensor, $H_{\mu\nu}$, to represent this sum of integrals,

$$H_{\mu\nu}(q) \equiv \sum_n \int dPS_n \; T_\mu \; T^*_\nu, \tag{2.4}$$

which after the integration and summation can only be a function of q[2.1]. Now, there are only two possible Lorentz covariant two-index tensor functions of one four-vector, $g_{\mu\nu}$ and $q_\mu q_\nu$. We therefore parameterize $H_{\mu\nu}$ as a linear combination of these, with coefficients that are functions of the only available Lorentz scalar, q^2,

$$H_{\mu\nu} = A(q^2) g_{\mu\nu} + B(q^2) q_\mu q_\nu. \tag{2.5}$$

Finally, since the theory is gauge invariant (in practice boiling down to invariance under the change $\varepsilon^\mu \to \varepsilon^\mu + q^\mu$ for the polarization vector of a photon of momentum q), $H_{\mu\nu}$ must be perpendicular to both q^μ and q^ν,

$$q^\mu H_{\mu\nu} = q^\nu H_{\mu\nu} = 0, \tag{2.6}$$

giving a relation between the two functions,

$$A = -q^2 B. \tag{2.7}$$

The final step is to realize that $B(s)$ has to be dimensionless. Since it is a function of only one dimensionful parameter, it must therefore be constant. We therefore have the fundamental prediction that (for energies well above all hadron masses) the cross section to produce any number of hadrons is proportional to that to produce a muon–antimuon pair,

$$R(e^+e^-) \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = \text{constant}, \tag{2.8}$$

[2.1] Can you spot the flaw in this argument? It assumes that all information about the hadron momenta is washed out by the integration, which is only true if they are massless. In general since p_h^2 is fixed at m_h^2 during the integration, H also depends in a complicated way on the masses of all possible hadrons. In fact we will shortly justify, on the basis of a space-time picture, neglecting these, in the limit that q^2 is much greater than all m_h^2. It also ignores any other masses in the problem, like the Z mass, which we remedy later on.

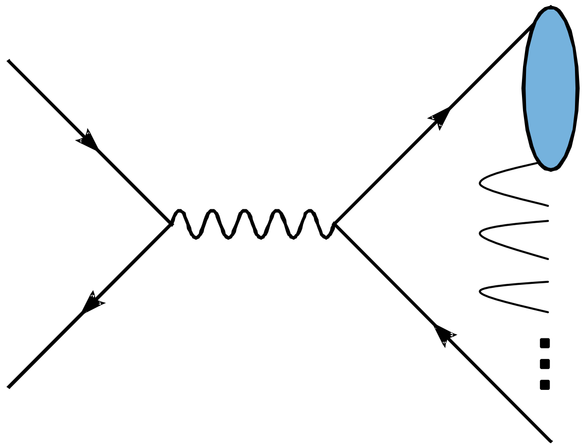


Fig. 2.1: Space-time sketch of the production of a hadron in e^+e^- annihilation

without knowing anything about the interactions of hadrons!

In order to go further than this and try to predict this constant, or learn something from its measurement, we need a specific model of the production of hadrons. This is provided by the *quark parton model*. Of course this can be more rigorously derived, but I find it more useful to illustrate the physics with a space-time argument, see Fig. 2.1. Since the photon is highly virtual, it is produced and decays to quarks in a small space-time volume, $t \sim 1/\sqrt{s}$. On the other hand, the wavefunction of a hadron with mass $\sim m_{\mathrm{had}}$ has spatial extent $\sim 1/m_{\mathrm{had}}$ and hence the confinement of a quark pair into the hadron takes $t \sim 1/m_{\mathrm{had}}$. Thus there is no time for the confinement to affect the annihilation cross section and we expect

$$\sigma(e^+e^- \to \text{hadrons}) \approx \sigma(e^+e^- \to \text{quarks}), \tag{2.9}$$

and the Feynman rules do tell us how to calculate that.

In fact, we can go further than that and use an argument from quantum mechanics to postulate the form of the corrections to this approximation. Over a region of size $\sim 1/\sqrt{s}$, the amount by which the wave function of a hadron with spatial extent $\sim 1/m_{\mathrm{had}}$, could vary is $\sim m_{\mathrm{had}}/\sqrt{s}$ and the corrections should be at least this to some positive power,

$$\sigma(e^+e^- \to \text{hadrons}) = \sigma(e^+e^- \to \text{quarks}) \times \left(1 + \mathcal{O}\left(\frac{m_{\mathrm{had}}}{\sqrt{s}}\right)^n\right). \tag{2.10}$$

On the basis of the space-time picture, we can only justify that the corrections to the quark parton model are suppressed by some (positive) power of the ratio of scales. In practice, n is believed to be 6 for e^+e^- annihilation, making these corrections so small as to be almost impossible to measure. For most cross sections however, n is 2, and for jet cross sections, 1.

We calculated the cross section for $e^+e^- \to q\bar{q}$ in Section 1.6 and obtained

$$R_{e^+e^-} \equiv \frac{\sigma(\text{hadrons})}{\sigma(\text{muons})} = N_c \sum_q e_q^2, \tag{2.11}$$

where the sum over q is over all quark flavours that are kinematically allowed, i.e. for which $\sqrt{s} > 2m_q$. If we ignore effects close to threshold, such as the formation of bound states, we can expect a plot of $R_{e^+e^-}$ against $\sqrt{s}$ to present a series of steps at twice the quark masses and be flat in between. In principle one can read off the quark masses and charges from this plot.

Looking at the data in Fig. 2.2, we see that the general trend is as expected, but there are clearly corrections that are not accounted for by the quark parton model. One of these is the effect of higher

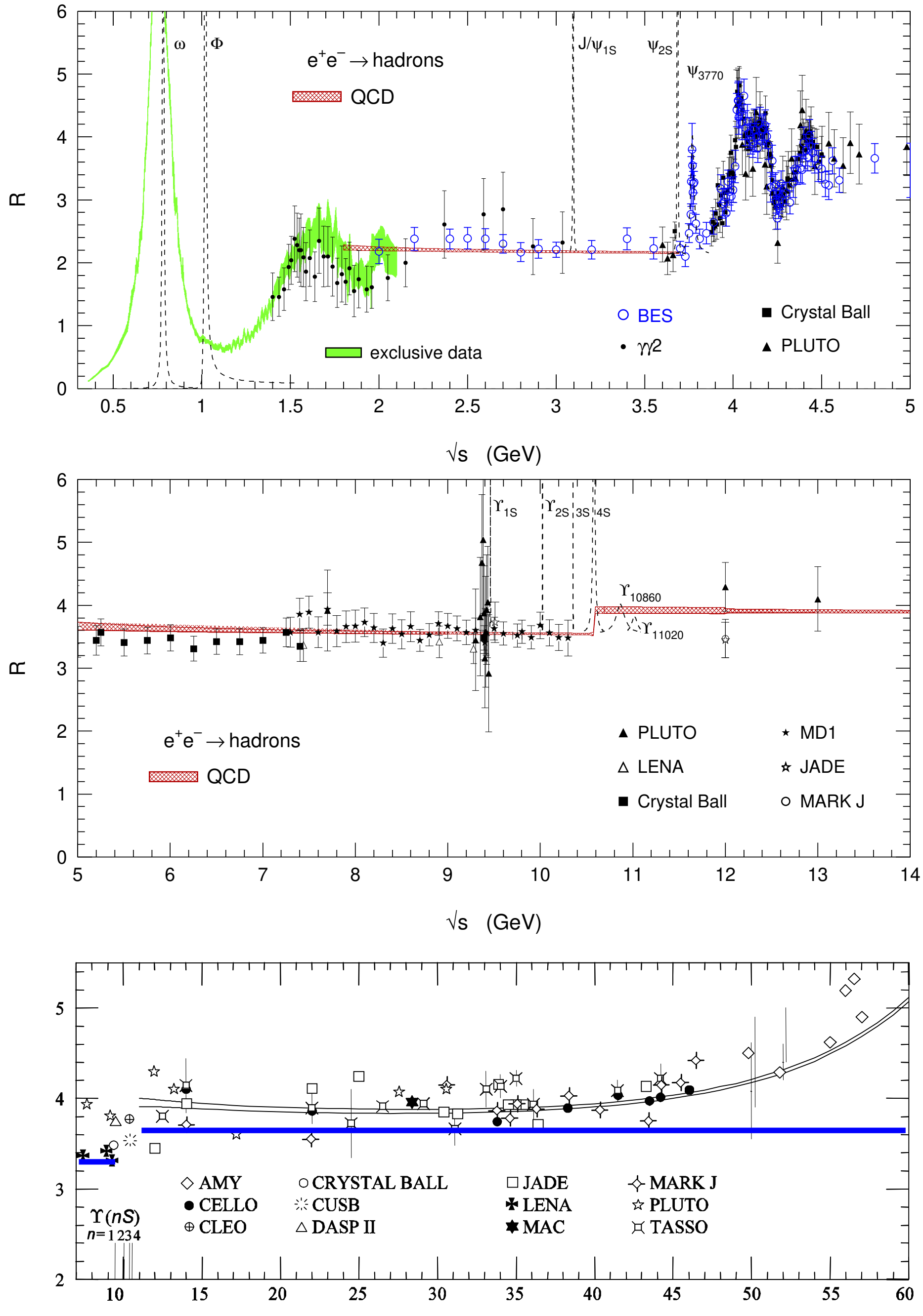


Fig. 2.2: Data on $R_{e^+e^-}$ as a function of centre-of-mass energy. Upper two panels taken from [4], lower from ESW [1]. The bands (red above, white below) show the QCD prediction, while the horizontal lines in the lower panel show the quark parton model expectations.

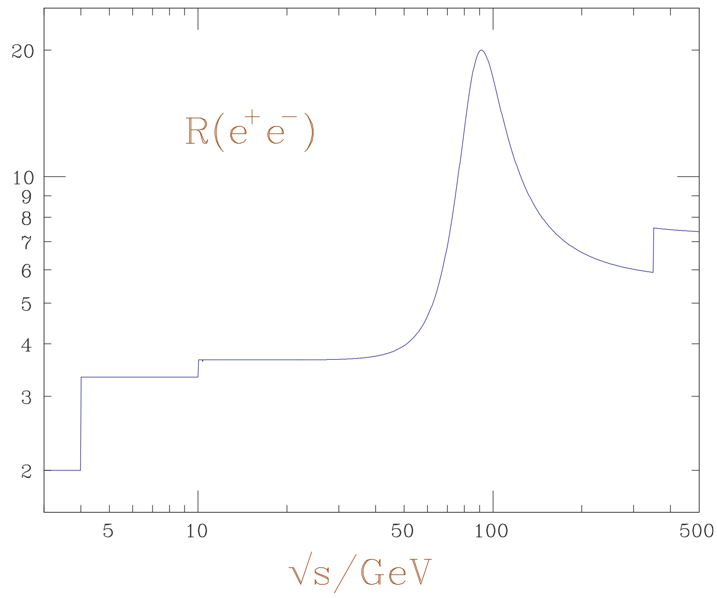


Fig. 2.3: Calculation of R_{had} as a function of centre-of-mass energy

order QCD corrections, which we include in the next lecture. Another is the effect of the Z^0 boson, which is clearly seen at the high energy end of Fig. 2.2, which we include shortly.

Before including the Z^0 contribution, it is worth remarking on a historical ambiguity that affects this figure. Although people wrote

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}$$

they often didn't actually use that formula to show their experimental results, but rather

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\frac{4\pi\alpha^2}{3s}},$$

using the leading order QED result for the denominator. Clearly many of the experimental and theoretical systematic errors would be smaller if the former was used, although of course the statistical errors would be larger, by around a factor of 2. More recent measurements, for example from LEP, have used the more honest notation in which the numerator and denominator are calculated or measured in the same way. This is sometimes called R_{had} to differentiate it from R.

In Fig. 2.3 I show the calculation of R_{had} in the quark parton model, including the Z^0 contribution. It is clear that γ–Z interference is important, even far from the Z peak. However, exactly on the peak the interference is zero (you might like to think about a simple explanation for why) and R_{had} is given to a good approximation by the Z contribution alone,

$$R_{had} = N_c \frac{\sum_q v_q^2 + a_q^2}{v_\mu^2 + a_\mu^2} = 20.095, \tag{2.12}$$

where v_i and a_i are the vector and axial couplings of the Z^0 to fermion type i. I note for future reference that the value including the photon contribution is 19.984. This number compares well with the LEP average measured value of 20.767 ± 0.025. However, the difference is still large on the scale of the experimental uncertainty, again indicating a clear need for the QCD corrections.

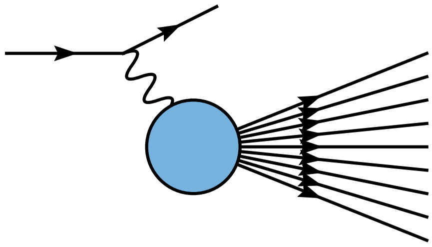


Fig. 2.4: Decay of τ lepton to hadrons

2.1.1 τ decays

We conclude this section by mentioning the closely-related process of τ decay to hadrons, depicted in Fig. 2.4. One can apply exactly the same arguments to the blob in this diagram as to annihilation of e^+e^- to hadrons. The only differences are that we have a virtual W boson producing hadrons instead of a virtual photon, and that we have an integral over all virtualities of the W between the τ mass and zero, rather than a single virtuality fixed by the beam energies. Nevertheless exactly the same arguments follow through and one obtains

$$R_\tau \equiv \frac{\mathcal{B}(\tau \to \text{hadrons})}{\mathcal{B}(\tau \to \mu)} = N_c \sum_{i,j} |V_{ij}|^2 \approx N_c, \tag{2.13}$$

where the sum is over the flavours of quark and antiquark that can appear in the W decay and V is the CKM matrix. Since a τ^- can decay to $\bar{u}d$ or $\bar{u}s$, to a good approximation this sum is $\cos^2\theta_C + \sin^2\theta_C$ and the final result follows.

We will see later that this process provides an excellent measurement of α_S.

2.2 Deep inelastic scattering

Historically, the quark model developed as a way of rationalizing the vast array of strongly-interacting particles that had been found by the 1960s. However, it was not clear whether quarks were really physical constituents of hadrons, or merely a convenient mathematical language to describe the hadrons' wave functions. The decisive evidence came from deep inelastic scattering experiments at SLAC. Today, deep inelastic scattering experiments give us by far the best information about the internal structure of the proton.

2.2.1 Quarks as partons in hadronic scattering

The classic probe of nuclear structure is electron–nucleus scattering. Assuming the scattering takes place by exchanging a single photon, measuring the kinematics of the scattered electron uniquely constrains that of the photon. The scattered electron has two non-trivial kinematic variables, its energy and scattering angle. These can more conveniently be converted into the photon virtuality ($Q^2 \equiv -q \cdot q$) and energy in the nucleus rest frame ν. Q^2 controls the resolving power of the photon, $Q^2 \sim 1/\lambda^2$. For fixed small $Q^2 \ll 1/R^2$, where R is the nuclear radius, the photon is absorbed elastically by the nucleus, giving a narrow peak in the ν distribution at $\nu = Q^2/2M_N$. For increased $Q^2 \sim 1/R^2$ one begins to resolve nuclear resonances as additional peaks at higher ν. Finally, for large $Q^2 \gg 1/R^2$, one resolves the proton constituents of the nucleus, with the photon being absorbed elastically by individual protons. These show up as a peak at $\nu = Q^2/2M_p$, broadened by the internal motion of the proton within the nucleus.

The scattering of electrons off hadrons, protons for example, is exactly analogous: at low Q^2 one sees only elastic proton scattering, but as Q^2 is increased, the photon can be elastically absorbed by the

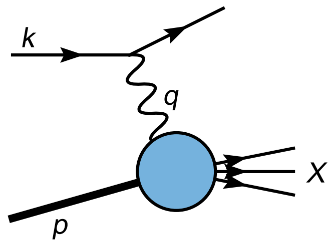


Fig. 2.5: Deep inelastic scattering

(charged) quark constituents of the proton. (Eventually at very large Q^2 and ν something new happens relative to the nuclear case, but we will not discuss that until the next lecture.)

We are interested in the region of Deep ($Q^2 \gg M_p^2$) Inelastic ($W^2 \gg M_p^2$, where W is the invariant mass of the photon–proton system) Scattering, DIS. We are therefore justified in neglecting the proton mass throughout, provided we do not work in the proton rest-frame, which is not well defined in that case. This is most conveniently done by working in terms of Lorentz-invariant variables.

2.2.2 Lorentz-invariant variables

It is convenient to describe this in terms of Lorentz-invariant variables. We label the momenta as shown in Fig. 2.5. For an electron of momentum k to scatter to one of momentum k' by exchanging a photon of momentum q with a proton of momentum p we again have, for fixed centre-of-mass energy s, only two independent kinematic variables,

$$s = (k+p)^2, \tag{2.14}$$

$$Q^2 = -q^2, \tag{2.15}$$

$$x = \frac{Q^2}{2p\cdot q}, \tag{2.16}$$

in terms of which we can calculate two other commonly-used variables

$$W^2 = (p+q)^2 = Q^2\frac{1-x}{x}, \tag{2.17}$$

$$y = \frac{p\cdot q}{p\cdot k} = \frac{Q^2}{xs}. \tag{2.18}$$

The kinematic limits are

$$Q^2 < s, \tag{2.19}$$

$$x > \frac{Q^2}{s}. \tag{2.20}$$

The coverage of the (x, Q^2) plane by the HERA, and earlier fixed target, DIS experiments is shown in Fig. 2.6

2.2.3 Structure functions

Since we do not yet know anything about the internal structure of protons, we cannot calculate the matrix element for the interaction of a photon with the proton to produce some arbitrary state X. However, like in the case of e^+e^- to hadrons we can get a surprisingly long way just by considering the properties that that matrix element must satisfy.

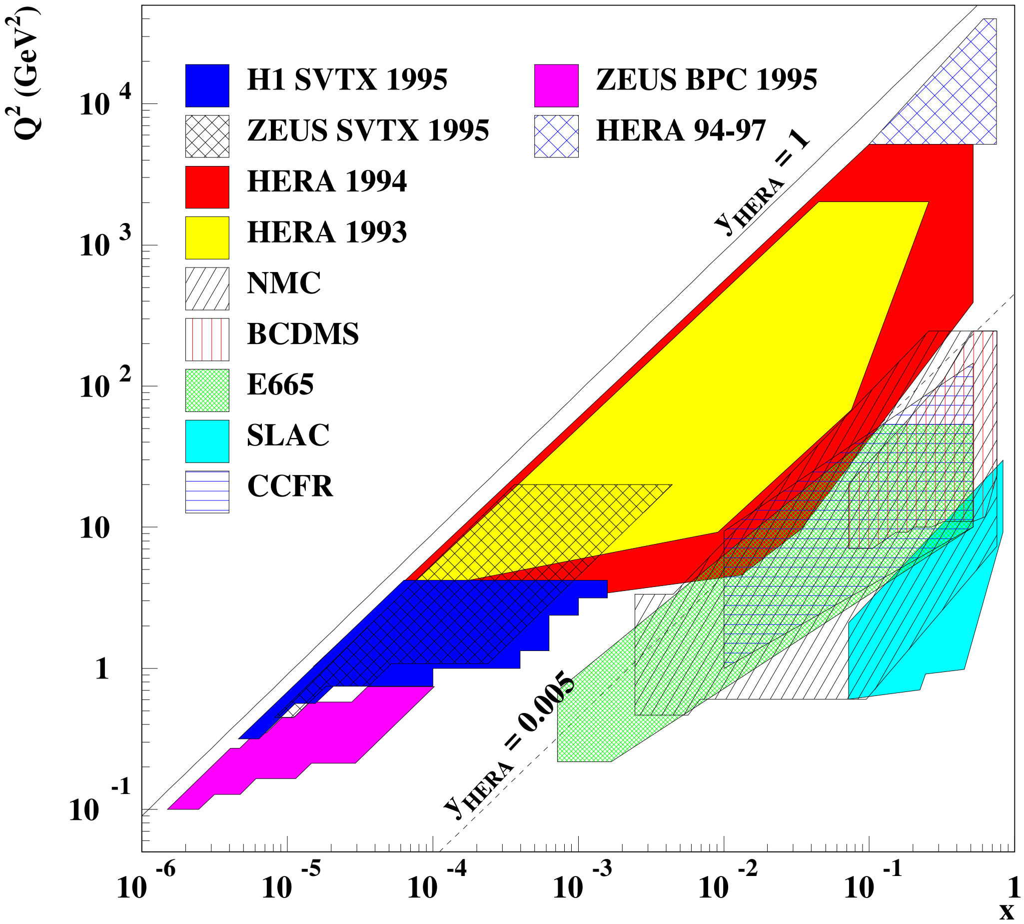


Fig. 2.6: The x–Q^2 plane, showing the coverage of measurements by various experiments

We parameterize the matrix element for a proton of momentum p to absorb a photon of momentum q and Lorentz index μ to produce an arbitrary set of hadrons X with fixed momenta $\{p_X\}$ as

$$e\,T_\mu(p,q;\{p_X\}). \tag{2.21}$$

We therefore have the matrix element squared for the whole process

$$\frac{1}{4}|\mathcal{M}|^2 = \frac{1}{4}\frac{e^4}{Q^4}\,\mathrm{Tr}\left\{\not{k}\gamma^\mu\not{k}'\gamma^\nu\right\}\,T_\mu(p,q;\{p_X\})\,T_\nu^*(p,q;\{p_X\}). \tag{2.22}$$

For convenience we define the Lorentz tensor

$$L^{\mu\nu} = \mathrm{Tr}\left\{\not{k}\gamma^\mu\not{k}'\gamma^\nu\right\}. \tag{2.23}$$

If the state X consists of n hadrons, then the $n{+}1$-body phase space for the whole process can be factorized into a part describing the electron kinematics times the n-body phase space for X,

$$dPS = \frac{Q^2}{16\pi^2 sx^2}\,dQ^2\,dx\,dPS_X. \tag{2.24}$$

This is as far as we can go for a specific state X, but we can get further by integrating over the phase space of X and summing over all possible states X. We define

$$\sum_X \int dPS_X\,\frac{1}{4}|\mathcal{M}|^2 \equiv \frac{e^4}{Q^4}\,L^{\mu\nu}H_{\mu\nu}, \tag{2.25}$$

or

$$\sum_X \int dPS_X \, T_\mu(p,q;\{p_X\}) \, T^*_\nu(p,q;\{p_X\}) = H_{\mu\nu}. \tag{2.26}$$

Since we have summed and integrated out all dependence on X, $H_{\mu\nu}$ can only depend on the vectors p and q. Since the electromagnetic and strong interactions conserve parity, it must be symmetric in μ and ν. There are only four possible symmetric two-index tensors that can be constructed from two vectors, so we can parameterize the hadronic tensor as a linear combination of them:

$$H_{\mu\nu} = -H_1 g_{\mu\nu} + H_2 \frac{p_\mu p_\nu}{Q^2} + H_4 \frac{q_\mu q_\nu}{Q^2} + H_5 \frac{p_\mu q_\nu + q_\mu p_\nu}{Q^2}, \tag{2.27}$$

where the Hs are scalar functions of the only two Lorentz scalars available $q{\cdot}q = -Q^2$ and $p{\cdot}q = Q^2/2x$, i.e., of x and Q^2 only (not s). (Note that we neglect $p{\cdot}p = M_p^2$ since we work in the limit $|q{\cdot}q|, p{\cdot}q \gg p{\cdot}p$.)

If we include Z^0 exchange (or charged current scattering) we can construct one further tensor, which is antisymmetric in μ and ν, $H_3 \, \epsilon_{\mu\nu\lambda\sigma} p^\lambda q^\sigma$, where $\epsilon_{\mu\nu\lambda\sigma}$ is the totally antisymmetric Lorentz tensor.

Contracting with $L^{\mu\nu}$ we find that H_4 and H_5 cannot contribute to physical cross sections (think about a simple explanation why not) and we have

$$L^{\mu\nu} H_{\mu\nu} = 4k{\cdot}k' \, H_1 + 4\frac{p{\cdot}k \; p{\cdot}k'}{Q^2} \, H_2. \tag{2.28}$$

Redefining (just a matter of convention) $H_1 = 4\pi F_1$ and $H_2 = 8\pi x F_2$, we obtain the final result for the scattering cross section

$$\frac{d^2\sigma}{dx\, dQ^2} = \frac{4\pi\alpha^2}{xQ^4} \left[y^2 x F_1(x,Q^2) + (1-y) F_2(x,Q^2) \right]. \tag{2.29}$$

Without knowing anything about the interactions of hadrons, we have been able to derive the s dependence of the scattering cross section for fixed x and Q^2 (which enters through the y dependence: recall $y = Q^2/xs$).

The Fs are called the structure functions of the proton. It is common to see other linear combinations of the structure functions,

$$F_T(x,Q^2) = 2xF_1(x,Q^2), \tag{2.30}$$
$$F_L(x,Q^2) = F_2(x,Q^2) - 2xF_1(x,Q^2), \tag{2.31}$$

which correspond to scattering of transverse and longitudinally polarized photons respectively. We therefore have

$$\frac{d^2\sigma}{dx\, dQ^2} = \frac{2\pi\alpha^2}{xQ^4} \left[(1+(1-y)^2) F_T(x,Q^2) + 2(1-y) F_L(x,Q^2) \right]. \tag{2.32}$$

In fact the most common form you will see this in nowadays is

$$\frac{d^2\sigma}{dx\, dQ^2} = \frac{2\pi\alpha^2}{xQ^4} \left[(1+(1-y)^2) F_2(x,Q^2) - y^2 F_L(x,Q^2) \right]. \tag{2.33}$$

For the majority of current data, y^2 is small and F_L can be neglected: only close to the kinematic limit, or for very precise data, need it be considered.

We have isolated all the non-trivial x and Q^2 dependence into the two functions $F_2(x,Q^2)$ and $F_L(x,Q^2)$, but we still have no idea how those functions behave. If we make the assumption that the

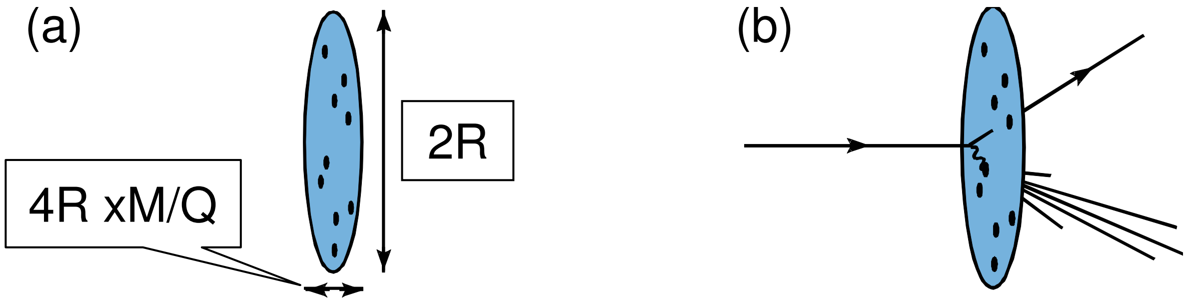


Fig. 2.7: In the Breit frame, the proton of diameter $2R$ is contracted to a pancake of thickness $4RxM_p/Q$ (a) so that a photon of high virtuality Q interacts incoherently with a single parton within it (b)

interaction of the photon with the innards of the proton does not involve any dimensionful scale, then we immediately get the result that the dimensionless Fs cannot depend on the dimensionful Q^2 and we get

$$\frac{d^2\sigma}{dx\, dQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[(1+(1-y)^2)F_2(x) - y^2F_L(x)\right], \tag{2.34}$$

known as Bjorken scaling. Experimentally this is true to a pretty good approximation, but given that the proton is supposed to consist of quarks, bound together with a distance scale $\sim 1/M_p$, how can the interaction possibly be M_p-independent? The answer to this lies in the parton model.

2.2.4 Parton distribution functions and Bjorken scaling

Although it is of course Lorentz-invariant, the parton model is most easily formulated in a frame in which the proton is fast moving. Most convenient is the so-called Breit frame, in which the photon has zero energy and collides head-on with the proton. In this frame, the proton energy is $Q/2x$. Assuming that in its own restframe it is a sphere of radius R, in the Breit frame it is massively Lorentz contracted to a flat pancake, still with transverse diameter $2R$, but with length $4RxM_p/Q \ll 2R$, as illustrated in Fig. 2.7a. The transverse size of the photon is $\sim 1/Q \ll 2R$. The photon therefore interacts with a tiny fraction of a thin disk, so provided that the quarks are sufficiently dilute the photon is not able to resolve the quarks' interactions and they act as if they were free. That is, the photon effectively collides with a single free quark, as illustrated in Fig. 2.7b.

Since they act as if they do not interact, their interactions do not introduce a dimensionful scale, and so the structure functions will obey Bjorken scaling.

More precisely, we suppose that the proton consists of a bundle of comoving partons, which carry a range of the proton's momentum. We posit probability distribution functions (more often called parton distribution functions, pdfs), such that partons of type q carry a fraction of the proton's momentum between η and $\eta+d\eta$ a fraction $f_q(\eta)d\eta$ of the time. Provided that these partons are pointlike $r^2 \ll 1/Q^2$ and dilute $f_q(\eta) \ll Q^2R^2$, the photons will scatter incoherently off individual partons. The cross section can then be factorized as the convolution of the pdfs with the cross section for parton scattering,

$$\frac{d^2\sigma(e+p(p))}{dx\, dQ^2} = \sum_q \int_0^1 d\eta\, f_q(\eta) \frac{d^2\sigma(e+q(\eta p))}{dx\, dQ^2}. \tag{2.35}$$

We will calculate the partonic cross section shortly, but first let me point out a couple of features it must have.

Firstly if we assume that the scattering is elastic, then the outgoing parton must be on mass-shell. Since we are then considering a two-to-two collision, which has only one nontrivial kinematic variable, the double-differential cross section in x and Q^2 must be proportional to a δ function fixing one of the variables. Specifically, if we assume that the partons are massless, then we obtain the relation

$$(q+\eta p)^2 = 2\eta\, p\cdot q - Q^2 = 0, \tag{2.36}$$

or

$$\eta = x. \tag{2.37}$$

Secondly if we assume that the struck partons are the quarks of the quark model, they must be fermions. Simply from helicity conservation, we can then show that $F_L = 0$. This is known as the Callan–Gross relation and was one of the first proofs that the quarks of the quark model really were the partons of the parton model. (If the partons were instead scalars we would have $F_T = 0$ and hence completely different y-dependence of the cross section.)

2.2.5 Scattering cross sections

To calculate the parton model prediction for the structure functions, we need the matrix elements for $eq \to eq$. These can be obtained by crossing symmetry from those for $e^+e^- \to q\bar{q}$. That is,

$$\sum |\mathcal{M}|^2 = 8(4\pi\alpha)^2\, e_q^2 N_c\, \frac{(p_e\cdot p_q)^2 + (p_e\cdot p_q')^2}{(p_e\cdot p_e')^2}. \tag{2.38}$$

Converting to the kinematic variables we defined earlier, we have

$$\sum |\mathcal{M}|^2 = 8(4\pi\alpha)^2\, e_q^2 N_c\, \frac{1+(1-y)^2}{y^2}. \tag{2.39}$$

Using (2.24), we have

$$dPS = \frac{Q^2}{16\pi^2 s x^2} dQ^2\, dx\, dPS_X. \tag{2.40}$$

Since X consists only of one massless parton, we have

$$\begin{aligned} dPS_X &= \frac{d^4p_X}{(2\pi)^3}\delta(p_X^2)\,(2\pi)^4\delta^4(\eta p + q - p_X) && (2.41)\\ &= (2\pi)\delta((\eta p + q)^2) && (2.42)\\ &= \frac{2\pi x}{Q^2}\delta(\eta - x). && (2.43)\end{aligned}$$

The full cross section is therefore

$$\begin{aligned} \frac{d\sigma}{dx\, dQ^2} &= \frac{1}{4N_c}\,\frac{1}{2\hat{s}}\,\frac{Q^2}{16\pi^2 s x^2}\,\frac{2\pi x}{Q^2}\,\delta(x-\eta)\,\sum|\mathcal{M}|^2 && (2.44)\\ &= \frac{1}{4N_c}\,\frac{y^2}{16\pi Q^4}\,\delta(x-\eta)\,\sum|\mathcal{M}|^2, && (2.45)\end{aligned}$$

where the factor of $1/N_c$ is the average over incoming colours. We therefore have

$$\frac{d\sigma(e+q)}{dx\, dQ^2} = \frac{2\pi\alpha^2}{Q^4}\,\delta(x-\eta)\; e_q^2\, \left(1+(1-y)^2\right) \tag{2.46}$$

and hence

$$\frac{d\sigma(e+p)}{dx\, dQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left(1+(1-y)^2\right)\sum_q e_q^2\, x f_q(x). \tag{2.47}$$

Comparing (2.47) with (2.33) we therefore have

$$F_2(x,Q^2) = \sum_q e_q^2\, x f_q(x), \tag{2.48}$$

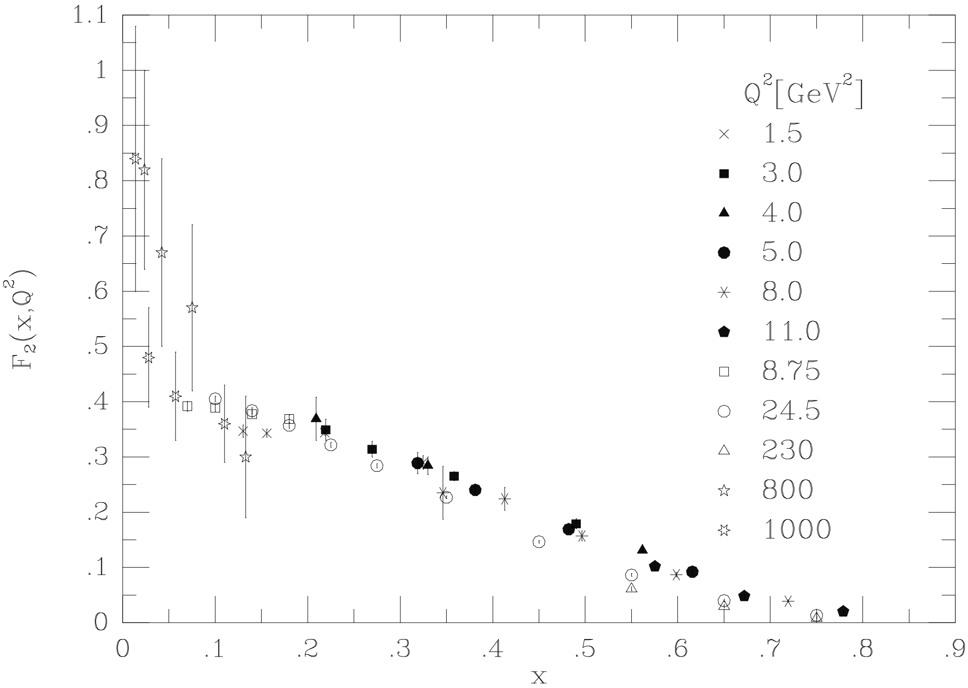


Fig. 2.8: The structure function F_2 as a function of x for various Q^2 values, exhibiting Bjorken scaling, taken from ESW [1]

$$F_L(x,Q^2) \quad = \quad 0. \tag{2.49}$$

Note that F_2 is Q^2-independent, showing Bjorken scaling.

Although we will see that QCD corrections do violate Bjorken scaling, it is satisfied pretty well by the data, as can be seen in Fig. 2.8.

2.2.6 Charged current neutrino DIS

We can consider charged current neutrino scattering in exactly the same way. Since the scattering takes place by the weak interaction, parity is violated, allowing one additional Lorentz structure,

$$L^{\substack{\nu \\ \bar\nu}}_{\mu\nu} \quad = \quad L^e_{\mu\nu} \pm 2i\epsilon_{\mu\nu\rho\sigma}k^\rho k'^\sigma, \tag{2.50}$$

$$H^{\mu\nu} \quad = \quad -H_1 g^{\mu\nu} + H_2\frac{p^\mu p^\nu}{Q^2} - \frac{i}{Q^2}\epsilon^{\mu\nu\rho\sigma}p_\rho q_\sigma H_3, \tag{2.51}$$

$$\Rightarrow L^{\substack{\nu \\ \bar\nu}}_{\mu\nu}H^{\mu\nu} \quad = \quad 2Q^2H_1 + Q^2\frac{1-y}{x^2y^2}H_2 \pm \frac{Q^2}{xy}H_3\;(1-y/2). \tag{2.52}$$

Thus, defining $H_3 = 8\pi x F_3$, we have a third structure function F_3:

$$\frac{d^2\sigma(\substack{\nu \\ \bar\nu}+p)}{dx\,dQ^2} = \frac{G_F^2}{4\pi x}\left(\frac{M_w^2}{Q^2+M_w^2}\right)^2\left[\left(1+(1-y)^2\right)F_2^{\substack{\nu \\ \bar\nu}} - y^2F_L^{\substack{\nu \\ \bar\nu}} \pm \left(1-(1-y)^2\right)xF_3^{\substack{\nu \\ \bar\nu}}\right], \tag{2.53}$$

where G_F is the Fermi constant and M_w the W boson mass. In the parton model we have

$$F_2^{\substack{\nu \\ \bar\nu}}(x,Q^2) \quad = \quad \sum_q 2xf_q(x) + \sum_{\bar q} 2xf_{\bar q}(x), \tag{2.54}$$

$$xF_3^{\substack{\nu \\ \bar\nu}}(x,Q^2) \quad = \quad \sum_q 2xf_q(x) - \sum_{\bar q} 2xf_{\bar q}(x), \tag{2.55}$$

where the sums for neutrino scattering are over all partons that can absorb a W^+, i.e., d, s, $\bar{\text{u}}$ and $\bar{\text{c}}$ and for antineutrino over those that can absorb a W^-, i.e., u, c, $\bar{\text{d}}$ and $\bar{\text{s}}$.

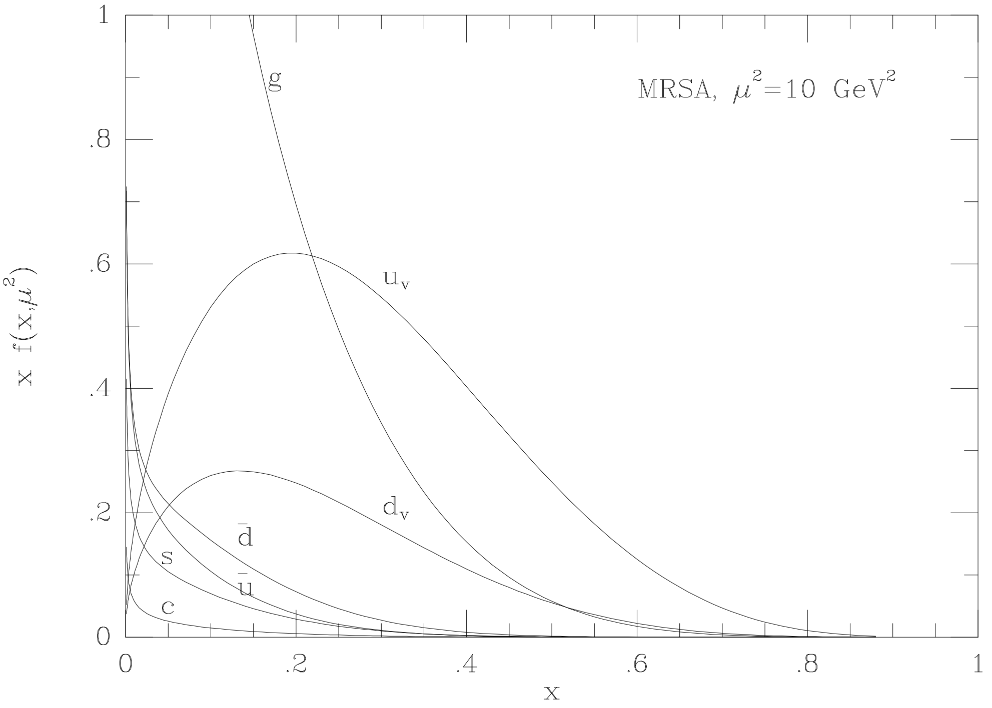


Fig. 2.9: Parton distribution function set A from the Martin-Roberts-Stirling group, taken from ESW [1]

2.2.7 Global fits

It is also possible to measure DIS on the neutron, or at least on deuterium from which the neutron structure functions can be derived. Using strong isospin symmetry, we have the relations

$$f_{u/n}(x) = f_{d/p}(x), \tag{2.56}$$
$$f_{\bar{u}/n}(x) = f_{\bar{d}/p}(x), \tag{2.57}$$
$$f_{d/n}(x) = f_{u/p}(x), \tag{2.58}$$
$$f_{s/n}(x) = f_{s/p}(x), \tag{2.59}$$

and so on. It is conventional to always refer to the proton case, dropping the "$/p$" subscript. We therefore have the slightly confusing result for F_2^{en} shown below, in which f_d is multiplied by $(2/3)^2$, and so on.

We therefore have

$$F_2^{ep} = \tfrac{1}{9}xf_d + \tfrac{4}{9}xf_u + \tfrac{1}{9}xf_{\bar{d}} + \tfrac{4}{9}xf_{\bar{u}} + \tfrac{1}{9}xf_s + \tfrac{1}{9}xf_{\bar{s}} + \tfrac{4}{9}xf_c + \tfrac{4}{9}xf_{\bar{c}}, \tag{2.60}$$
$$F_2^{en} = \tfrac{4}{9}xf_d + \tfrac{1}{9}xf_u + \tfrac{4}{9}xf_{\bar{d}} + \tfrac{1}{9}xf_{\bar{u}} + \tfrac{1}{9}xf_s + \tfrac{1}{9}xf_{\bar{s}} + \tfrac{4}{9}xf_c + \tfrac{4}{9}xf_{\bar{c}}, \tag{2.61}$$
$$F_2^{\nu p} = 2xf_d + 2xf_{\bar{u}} + 2xf_s + 2xf_{\bar{c}}, \tag{2.62}$$
$$xF_3^{\nu p} = 2xf_d - 2xf_{\bar{u}} + 2xf_s - 2xf_{\bar{c}}, \tag{2.63}$$
$$F_2^{\bar{\nu} p} = 2xf_u + 2xf_{\bar{d}} + 2xf_c + 2xf_{\bar{s}}, \tag{2.64}$$
$$xF_3^{\bar{\nu} p} = 2xf_u - 2xf_{\bar{d}} + 2xf_c - 2xf_{\bar{s}}. \tag{2.65}$$

If we make the assumption that $f_{\bar{s}} = f_s$ and $f_{\bar{c}} = f_c$, then we have six unknowns for six pieces of data so, given precise enough data, we could solve for all the pdfs exactly. In practice of course it is never so simple and one must make global fits to as wide a variety of data as possible.

One gets typical results like those shown in Fig. 2.9. Note that this uses the common notation of defining valence quark distributions,

$$f_{u_v} \equiv f_u - f_{\bar{u}}, \tag{2.66}$$

$$f_{d_v} \equiv f_d - f_{\bar{d}}. \tag{2.67}$$

Non-valence quarks are generically referred to as the sea.

2.2.8 Sum rules

Having results for the pdfs, one can form interesting integrals over them, for example,

$$\int_0^1 dx\, f_{u_v}(x) = 2, \tag{2.68}$$

$$\int_0^1 dx\, f_{d_v}(x) = 1. \tag{2.69}$$

Various such integrals can be constructed directly from the structure functions. It is worth checking that you can reproduce the physical interpretation of each.

2.2.8.1 The Gross–Llewellyn-Smith sum rule

$$\tfrac{1}{2}\int_0^1 dx \left(F_3^{\nu p} + F_3^{\bar{\nu}p}\right) = 3, \tag{2.70}$$

which counts the number of valence quarks in the proton. In QCD this provides a useful measurement of α_{S}, because the right-hand side is actually equal to $3\left(1 - \frac{\alpha_{\mathrm{S}}}{\pi} + \mathcal{O}(\alpha_{\mathrm{S}}^2)\right)$.

2.2.8.2 The Adler sum rule

$$\tfrac{1}{2}\int_0^1 \frac{dx}{x} \left(F_2^{\bar{\nu}p} - F_2^{\nu p}\right) = 1, \tag{2.71}$$

which counts the difference between the number of up and down valence quarks. This has the property that it is exact even in QCD, i.e., all higher order corrections vanish.

2.2.8.3 The Gottfried sum rule

$$\int_0^1 \frac{dx}{x} \left(F_2^{ep} - F_2^{en}\right) \approx 0.23, \tag{2.72}$$

where the result is experimental. This is sensitive to the difference between the number of up and down sea quarks: it would be 1/3 if they were equal.

2.2.8.4 The momentum sum rule

Finally, we have the particularly significant result

$$\tfrac{1}{2}\int_0^1 dx \left(F_2^{\nu p} + F_2^{\bar{\nu}p}\right) \approx 0.5, \tag{2.73}$$

where the result is again experimental. This tells us that only about half of the proton's momentum is carried by quarks and antiquarks.

2.3 Hadronic collisions

2.3.1 The Drell–Yan process

If the parton model is correct, the parton distribution functions should be universal. We should therefore be able to use the DIS measurements to make predictions for other hadronic scattering processes. The classic example is the so-called Drell–Yan process, of lepton pair production in hadron collisions,

$$h_1 + h_2 \to \mu^+ + \mu^- + X, \tag{2.74}$$

where the state X goes unmeasured. In the parton model this arises as the sum over all quark types of

$$q + \bar{q} \to \mu^+ + \mu^- . \tag{2.75}$$

The cross section can be written as the convolution of pdfs with a partonic cross section, exactly like in DIS:

$$\frac{d\sigma(h_1(p_1) + h_2(p_2) \to \mu^+\mu^-)}{dM^2} = \sum_q \int_0^1 d\eta_1 f_{q/h_1}(\eta_1) \int_0^1 d\eta_2 f_{\bar{q}/h_2}(\eta_2) \frac{d\sigma(q(\eta_1 p_1) + \bar{q}(\eta_2 p_2) \to \mu^+\mu^-)}{dM^2}, \tag{2.76}$$

where M is the mass of the $\mu^+\mu^-$ pair. Note that since the partonic cross section contains a $\delta(M^2 - \eta_1\eta_2 s)$ term, binning the data in M gives extra information about the pdfs. In fact, binning also in the rapidity of the lepton pair, defined by

$$y \equiv \frac{1}{2} \ln \frac{E_{\mu^+\mu^-} + p_{z,\mu^+\mu^-}}{E_{\mu^+\mu^-} - p_{z,\mu^+\mu^-}}, \tag{2.77}$$

both η values are fixed, providing a direct measurement of the parton distribution functions (the partonic cross section can easily be obtained by crossing the $e^+e^- \to q\bar{q}$ one we calculated in Section 1.6, divided by a factor of N_c^2 for the average over incoming colours):

$$\frac{d^2\sigma}{dM^2 dy} = \frac{4\pi\alpha^2}{3N_c M^2 s} \sum_q e_q^2 f_{q/h_1}(e^y M/\sqrt{s}) f_{\bar{q}/h_2}(e^{-y} M/\sqrt{s}). \tag{2.78}$$

Note that the case $h_1 = h_2 = p$ provides a particularly good measure of the sea quark distribution functions, which are hard to extract from DIS data.

2.3.2 *Prompt photon and jet production*

Although we have not yet mentioned gluons, we will see in the next lecture that there is also a non-zero pdf for the gluon, $f_g(\eta)$, as can also be inferred from the momentum sum rule mentioned earlier. As well as being important for higher order corrections to the processes given above, there are many processes in which they participate at tree level. The most important of these are prompt photon production,

$$h_1 + h_2 \to \gamma + X, \tag{2.79}$$

and jet production

$$h_1 + h_2 \to q + q + X, \tag{2.80}$$
$$h_1 + h_2 \to q + \bar{q} + X, \tag{2.81}$$
$$h_1 + h_2 \to q + g + X, \tag{2.82}$$
$$h_1 + h_2 \to g + g + X, \text{ etc.} \tag{2.83}$$

The gluon pdf is used in exactly the same way as the quark ones, and hadronic cross sections can still be calculated as the sum of convolutions of pdfs with partonic cross sections. Prompt photon production receives contributions from two partonic processes,

$$q + \bar{q} \to \gamma + g, \tag{2.84}$$
$$q + g \to \gamma + q. \tag{2.85}$$

In the case $h_1 = h_2 = p$, the latter dominates, providing a measure of the gluon pdf. However there is a slight complication, in that processes (2.84), (2.85) are proportional to α_{S}, which is less well-known than α, which controls the other processes we have studied. In fact this is always the case, that measurements of the gluon pdf actually measure $\alpha_{\mathrm{S}} \times f_g$ in general. The QCD corrections to this process turn out to be a lot larger than any of the others we have considered, further complicating this measurement.

2.4 Summary

We have considered the tree-level phenomenology of e^+e^- annihilation, deep inelastic scattering and, more briefly, hadron collisions. It is remarkable how much QCD phenomenology can be understood using tree level results. However, we have to worry that α_{S} is not so small, so higher order corrections must be important. Equally importantly, it would be nice to see whether, and if so how, the parton model emerges from QCD.

We discuss both these issues in the next lecture.

3 Higher order corrections

3.1 e^+e^- annihilation at one loop

In this section, I go through the calculation of the NLO correction to the $e^+e^- \to$ hadrons cross section in some detail. I will briefly describe some of the more technical aspects of the calculation, for those interested, in Section 3.1.2, but those who are not can safely skip this section, since I recap the important results at the start of Section 3.1.3.

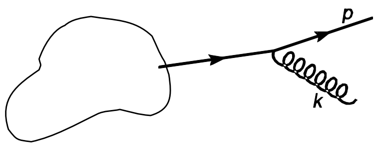
In discussing the $e^+e^- \to$ hadrons cross section at tree level, we assumed that quarks produce hadrons with probability 1. Therefore we calculated the $e^+e^- \to q\bar{q}$ cross section in Section 1.6. In discussing jet cross sections, we extended this to say that all partons produce hadrons with probability 1. Therefore we should calculate the total cross section to produce any number or type of partons. At leading order this makes no difference, since the only possible process is $e^+e^- \to q\bar{q}$, but at order α_{S} we have to calculate and sum the cross sections for $q\bar{q}$ and $q\bar{q}g$ final states. We start with the latter.

Recall that the total $q\bar{q}g$ cross section is divergent,

$$\sigma = \sigma_0 \, C_F \frac{\alpha_s}{2\pi} \int dx_1 \, dx_2 \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}, \tag{3.1}$$

where the region of integration is the upper right triangle of the unit square, bordered by the lines $x_1 = 1$ and $x_2 = 1$, which are the singular regions. This divergence must be regularized in some way, before we can make progress.

First though we discuss the origin of the divergences. They arise from propagator factors that diverge,



$$\frac{1}{(p+k)^2} = \frac{1}{2p\cdot k} = \frac{1}{2E\omega(1-\cos\theta)} \approx \frac{1}{E\omega\theta^2}, \tag{3.2}$$

where E and ω are the quark and gluon energies and θ is the angle between them.

In the collinear limit, $\theta \to 0$, one in principle obtains $1/\theta^4$ in the matrix element squared, but in fact the numerators always contribute a factor of θ^2, so one obtains

$$|\mathcal{M}|^2 \sim \frac{1}{\theta^2}. \tag{3.3}$$

In the soft limit, $\omega \to 0$, one has in the interference between diagrams in which the gluon is attached to quark 1 and quark 2,

$$|\mathcal{M}|^2 \sim \frac{p_1\cdot p_2}{p_1\cdot k \, p_2\cdot k} \sim \frac{1}{\omega^2}. \tag{3.4}$$

In terms of ω and θ the phase space is given by

$$\frac{d^3k}{2\omega} = \tfrac{1}{2}\,\omega d\omega \, \sin\theta \, d\theta \, d\phi \sim \omega d\omega \, \theta d\theta. \tag{3.5}$$

We therefore have logarithmic singularities in both the soft and collinear limits. We generically refer to both of these as the infrared limit.

3.1.1 Regularization

As in the discussion of renormalization, the simplest way we could regularize this cross section is with a cutoff, for example on the transverse momentum of the gluon, which would prevent the integration entering both the soft and collinear regions. However, we will see that infrared singularities cancel between different contributions, in this case $q\bar{q}$ and $q\bar{q}g$, so we must use a regularization that can be consistently applied in all contributions. It is not clear that this is the case for a cutoff, since it must be applied in both real and virtual contributions, which have very different structures. Instead, to ensure consistent application across all processes, it is better to modify the theory in such a way that some dimensionless parameter ϵ regulates the divergences. Then the complete calculation can be performed in this modified theory and at the end of the calculation, when all the divergences have cancelled, the limit $\epsilon \to 0$ can be smoothly taken. Remarkably, dimensional regularization, which we used for ultraviolet singularities, also provides a consistent regulator for infrared singularities, as we shall discuss in detail shortly.

Another regularization scheme, which actually works well in QED, and for simple processes in QCD, is the gluon (or photon) mass regularization. We introduce a non-zero gluon mass $m_g^2 = \epsilon Q^2$. This prevents the propagators from reaching zero and diverging: for massless quarks the minimum value is m_g^2 and for a quark of mass m_q it is $2m_q m_g$. With this modification one can recalculate the differential cross section and integrate it to give a finite result,

$$\sigma_{q\bar{q}g} = \sigma_0 \, C_F \frac{\alpha_s}{2\pi} \left(\log^2 \frac{1}{\epsilon} - 3 \log \frac{1}{\epsilon} + 7 - \frac{\pi^2}{3} + \mathcal{O}(\epsilon) \right). \tag{3.6}$$

However, since a non-zero gluon mass violates gauge invariance, this method is bound to fail in general. In particular, it is not suitable for any process in which any lowest order contributions have external gluons. As in the ultraviolet case, the only scheme that is known to be consistent with all the symmetries of QCD, and hence to work to arbitrary orders in arbitrary processes, is dimensional regularization.

The reason why I said that it is remarkable that dimensional regularization works in the infrared limit is the fact that the two limits have non-overlapping regions of applicability in the complex d plane. Ultraviolet-singular integrals are regularized by working in $d < 4$ dimensions, but infrared-singular integrals are only rendered finite by working in $d > 4$ dimensions. However, by carefully splitting contributions that are singular in both the infrared and ultraviolet one can consider the regularization schemes that are used in each as independent. In each region, one considers the appropriate dimensionality ($d = 4 - 2\epsilon$ with $\epsilon > 0$ in the ultraviolet and with $\epsilon < 0$ in the infrared) and then analytically continues to the whole complex ϵ plane. Since analytical continuation is unique, this gives a unique result for each, in the region of applicability of the other, and the two can be combined before the limit $\epsilon \to 0$ is taken. This subtlety leads to some surprising results, for example for the self-energy of a massless quark, discussed below.

As the calculation of cross sections in dimensional regularization is rather technical, it is rare to see it done in summer school lectures, but I think it brings out some interesting points, so I at least sketch how the calculation works in Section 3.1.2. As I said, those who disagree can safely skip ahead to Section 3.1.3.

3.1.2 Aside: Real and virtual corrections in dimensional regularization

It is straightforward to generalize the Feynman rules to d dimensions and fairly straightforward to generalize the Dirac algebra. The result is that d-dimensional matrix elements still have propagators $\sim 1/p^2$, but that the numerators become d dependent. (It is worth mentioning the closely-related dimensional reduction scheme, which is often used for supersymmetry calculations, since conventional dimensional regularization violates supersymmetry. In this scheme one works in d dimensions, but modifies the theory in such a way that fermions and massless vector bosons still have 2 spin states, instead of $d-2$ as in dimensional regularization. The result is that the matrix elements themselves are equal to the 4-

dimensional ones and it is only on performing the loop and phase space integrals that the d dimensionality gets introduced.)

3.1.2.1 Phase space integrals

We will have to integrate over d-dimensional phase space. We begin by considering integer values of d and then continue the results to real values. It is straightforward to write down the basic integration measure,

$$d^d k\, \delta_+(k^2) = \frac{d^{d-1}k}{2\omega} = \frac{1}{2}\omega^{d-3} d\omega\, d\Omega_{d-2}, \tag{3.7}$$

where ω is the energy of k and $d\Omega_{d-2}$ is an element of $d-2$-dimensional solid angle. The only difficulty concerns the evaluation of integrals over this solid angle. In four dimensions we have

$$k = \omega(1; \sin\phi\ \sin\theta, \cos\phi\ \sin\theta, \cos\theta) \qquad \text{(4 dimensions)}, \tag{3.8}$$

where θ and ϕ are the usual spherical polar coordinates with θ the polar angle and ϕ the azimuthal angle. In five dimensions we have

$$k = \omega(1; \sin\psi\ \sin\phi\ \sin\theta, \cos\psi\ \sin\phi\ \sin\theta, \cos\phi\ \sin\theta, \cos\theta) \qquad \text{(5 dimensions)}, \tag{3.9}$$

where ψ is an azimuthal angle in the additional dimension. Generalizing to d dimensions, we have $d-4$ additional azimuths and we write k generically as

$$k = \omega(1; \ldots, \cos\phi\ \sin\theta, \cos\theta) \qquad (d \text{ dimensions}), \tag{3.10}$$

where the ellipsis represents a $d-3$-vector of length $\sin\phi\ \sin\theta$ containing $d-4$ azimuths. Depending on the complexity of the calculation, more or less of these additional components have to be specified precisely. In fact in our case, since we only consider the relative orientations of three momenta that have zero total momentum, and therefore all lie in a plane, it is sufficient to specify

$$k = \omega(1; \ldots, \cos\theta) \qquad (d \text{ dimensions}), \tag{3.11}$$

where the ellipsis represents a $d-2$-vector of length $\sin\theta$ containing $d-3$ azimuths.

We can see how to integrate over the additional azimuths by again considering integer d and then generalizing,

$$\int d\Omega_1 \quad = \quad \int d\phi = 2\pi, \tag{3.12}$$

$$\int d\Omega_2 \quad = \quad \int d\phi\ \sin\theta\, d\theta = 4\pi, \tag{3.13}$$

$$\int d\Omega_3 \quad = \quad \int d\psi\ \sin\phi\, d\phi\ \sin^2\theta\, d\theta = 2\pi^2, \tag{3.14}$$

and so on. We have a recursion relation

$$\int d\Omega_n = \int d\Omega_{n-1}\ \sin^{n-1}\theta\, d\theta, \tag{3.15}$$

which is solved by

$$\Omega_n \equiv \int d\Omega_n = \frac{2\pi^{(n+1)/2}}{\Gamma[(n+1)/2]}. \tag{3.16}$$

We are now equipped to tackle the phase space integral, and see how the dimensional regularization succeeds in regularizing our integrals.

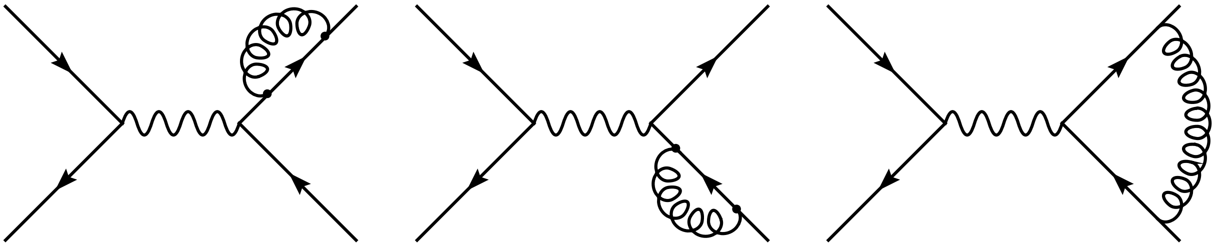


Fig. 3.1: One-loop diagrams for $e^+e^- \to q\bar{q}$

3.1.2.2 Regularization

Since the form of the propagator factors is unchanged in d dimensions, and it is these that dominate the singular region, it is straightforward to read off the behaviour in the regularized theory. In the soft region we have

$$\int_0 \omega^{1-2\epsilon} d\omega \, \frac{1}{\omega^2} = \int_0 \frac{d\omega}{\omega^{1+2\epsilon}} \sim -\frac{1}{2\epsilon}, \qquad \epsilon < 0, \tag{3.17}$$

and in the collinear region

$$\int_0 \sin^{1-2\epsilon}\theta \, d\theta \, \frac{1}{\theta^2} \sim \int_0 \frac{d\theta}{\theta^{1-2\epsilon}} \sim -\frac{1}{2\epsilon}, \qquad \epsilon < 0. \tag{3.18}$$

Since our cross section is divergent in both limits, and they can overlap, i.e., a radiated gluon can be both soft and collinear, we expect the total cross section to be of order $1/\epsilon^2$. Note, as a consistency check, that the integrands are positive definite and that, in the region in which they are well-defined, $\epsilon < 0$, the results are positive (and divergent as $\epsilon \to 0$).

3.1.2.3 Total $e^+e^- \to q\bar{q}g$ cross section

We now have all the ingredients we need to calculate the differential cross section for $e^+e^- \to q\bar{q}g$ and to integrate it over all phase space in dimensional regularization. We obtain

$$\sigma_{q\bar{q}g} = \sigma_0 \, C_F \frac{\alpha_s}{2\pi} H(\epsilon) \left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + \frac{19}{2} - \pi^2 + \mathcal{O}(\epsilon) \right), \tag{3.19}$$

where σ_0 is the lowest order cross section and $H(\epsilon)$ is a smooth function, with $H(0) = 1$, that we will not ultimately need to know. Note that, as we anticipated from Eqs. (3.17) and (3.18), this result is positive, and divergent like $1/\epsilon^2$ as $\epsilon \to 0$.

So far, the regularization scheme has succeeded in quantifying the degree of divergence of the total three-parton cross section, but it has not helped us solve the problem of the divergence, by recovering a finite result for a physical cross section. As we already anticipated above, this will come by calculating the loop correction to $e^+e^- \to q\bar{q}$.

3.1.2.4 $\sigma(e^+e^- \to q\bar{q})$ at one loop

We already made the point that to calculate the total cross section for $e^+e^- \to$ hadrons, we must sum over all $e^+e^- \to$ partons processes. At this order of perturbation theory $q\bar{q}$ is the only other process that contributes. There are three diagrams, shown in Fig. 3.1. They are down by one power of $\alpha_{\rm S}$ relative to the tree-level diagram,

$$\mathcal{M}_1 \propto \alpha_{\rm S} \mathcal{M}_0. \tag{3.20}$$

Therefore $|\mathcal{M}_1|^2$ is two powers down and hence negligible at the order to which we are working. However, since the final state is the same as that of the tree-level diagram, the two interfere, and their interference, $\mathrm{I\!Re}\{\mathcal{M}_0^* \mathcal{M}_1\}$ does contribute at order $\alpha_{\rm S}$.

In quantum mechanics, you know that we must sum over all unobserved quantum numbers at the amplitude level. Since the gluon momentum is unconstrained by the outgoing quark momenta, we must sum over *all* gluon momenta,

$$\int d^d k. \tag{3.21}$$

Note that there is no mass-shell-constraining delta-function: the virtual integral is over all arbitrary on- and off-shell momenta.

We begin with the first two diagrams, which are proportional to the self-energy of a massless quark. It is actually easy to see that these have to be zero in dimensional regularization: the value of the integral has dimensions E^{d-4}, but by Lorentz invariance the result of the integral can only be a function of the square of the quark's momentum, $p^2 = 0$, so there is nothing that can provide this dimensionality[3.1]. The only way these two facts can be reconciled is if the integral is zero. However, if we examine the integrand somewhat closer, this is very surprising, because it is positive definite. How can a positive definite quantity integrate to zero?

The answer to this question comes from a subtle use of dimensional regularization. In fact this integral is divergent in both the infrared and ultraviolet. If we split the integral into two parts by introducing an arbitrary separation scale Λ, then we obtain an ultraviolet contribution $\sim \Lambda^{-2\epsilon}/\epsilon$ and an infrared contribution $\sim -\Lambda^{-2\epsilon}/\epsilon$. Each is positive in its domain of applicability ($\epsilon > 0$ and $\epsilon < 0$ respectively), but after analytically continuing each to arbitrary ϵ, they are exactly equal and opposite, giving a zero result for these diagrams.

Turning to the third diagram, the vertex correction, we find that it is also divergent in the infrared and ultraviolet regions. However, its ultraviolet divergence is exactly equal and opposite to the one from the sum of the two self-energy diagrams. Therefore the sum of the three diagrams is ultraviolet finite and no renormalization is needed at this order. This actually follows directly from the Ward identity of QED. Thus, one simply has to evaluate the vertex correction diagram in dimensional regularization, to obtain the complete order α_{S} contribution to $e^+e^- \to q\bar{q}$. We find that the infrared divergences do not cancel, and we obtain

$$\sigma_{q\bar{q}} = \sigma_0\, C_F \frac{\alpha_s}{2\pi} H(\epsilon) \left(-\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \pi^2 + \mathcal{O}(\epsilon) \right). \tag{3.22}$$

Dimensional regularization has succeeded in regularizing the divergence of this contribution as well. This time, however, the result is negative and divergent as $\epsilon \to 0$. This should not surprise us, as we already noted that this is an interference term, so there is no requirement that it be positive, as there was for $\sigma_{q\bar{q}g}$. In fact a quick glance at Eqs. (3.19) and 3.22) shows us that the divergences are going to cancel between them.

3.1.3 The total cross section

In the previous section we discussed how dimensional regularization provides finite results for the total cross sections for the $e^+e^- \to q\bar{q}$ and $e^+e^- \to q\bar{q}g$ processes, which each diverge as $\epsilon \to 0$. For the benefit of those who slept through it, I restate them here:

$$\sigma_{q\bar{q}} = \sigma_0\, C_F \frac{\alpha_s}{2\pi} H(\epsilon) \left(-\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \pi^2 + \mathcal{O}(\epsilon) \right), \tag{3.23}$$

$$\sigma_{q\bar{q}g} = \sigma_0\, C_F \frac{\alpha_s}{2\pi} H(\epsilon) \left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + \frac{19}{2} - \pi^2 + \mathcal{O}(\epsilon) \right). \tag{3.24}$$

According to our earlier discussion, the total cross section for $e^+e^- \to$ hadrons is given by the sum of the two. It is finite, so the limit $\epsilon \to 0$ can be taken,

[3.1] In fact this statement relies on working in a covariant gauge. In a lightcone gauge for example, the self-energy can depend on $n{\cdot}p$. This diagram is not then zero, but of course the final answer for the sum of the three diagrams is gauge invariant.

$$\sigma_{e^+e^-\to\text{hadrons}} = \sigma_0 \left(1 + C_F \frac{\alpha_S}{2\pi}\frac{3}{2}\right) \tag{3.25}$$

$$= \sigma_0 \left(1 + \frac{\alpha_S}{\pi}\right). \tag{3.26}$$

Of course, this would be useless if it depended on the regularization procedure. The proof of its independence is beyond us here, but it is worth demonstrating it, by comparison with another scheme, the gluon mass regularization, in which we have

$$\sigma_{q\bar{q}} = \sigma_0 C_F \frac{\alpha_S}{2\pi}\left[-\log^2\frac{1}{\epsilon} + 3\log\frac{1}{\epsilon} - \frac{11}{2} + \frac{\pi^2}{3} + \mathcal{O}(\epsilon)\right], \tag{3.27}$$

$$\sigma_{q\bar{q}g} = \sigma_0 C_F \frac{\alpha_S}{2\pi}\left[\ \ \log^2\frac{1}{\epsilon} - 3\log\frac{1}{\epsilon} + 7 - \frac{\pi^2}{3} + \mathcal{O}(\epsilon)\right], \tag{3.28}$$

$$\sigma_{\text{had}} = \sigma_0 \left[1 + \frac{\alpha_S}{\pi}\right]. \tag{3.29}$$

Note that the individual cross sections have completely different forms in the different schemes, but that the sum of the two is scheme independent.

Equation (3.26) is one of the most fundamental quantities in QCD and is certainly one of the most well-calculated and measured. Despite the fact that it is a relative small correction to the total rate, experimental and theoretical systematic errors are so small that they can almost be neglected — even with the large statistics of τ decays and Z decays at LEP, the statistical errors dominate. This means that not only does it provide one of the most accurate measurements, but its quoted accuracy is rather easy to interpret and implement in global analyses for example, unlike measurements that are dominated by systematics.

Equation (3.26) is now known up to order α_S^3. As discussed in Section 1.8, renormalization introduces a renormalization scale dependence into α_s and the coefficient functions beyond the first one,

$$\sigma_{e^+e^-\to\text{hadrons}} = \sigma_0 \left(1 + \frac{\alpha_S(\mu)}{\pi} + C_2\left(\frac{\mu^2}{s}\right)\left(\frac{\alpha_S(\mu)}{\pi}\right)^2 + C_3\left(\frac{\mu^2}{s}\right)\left(\frac{\alpha_S(\mu)}{\pi}\right)^3\right). \tag{3.30}$$

Reducing this renormalization-scale dependence is one of the biggest reasons for going to higher orders. As can be seen in Fig. 3.2, the scale-dependence is indeed significantly smaller at each order, giving stability over a wider range of μ. It can also be seen that provided μ is of order Q, the higher order corrections are relatively small. We will see shortly that simply taking the leading order result with $\mu = \sqrt{s}$ does surprisingly well and is certainly sufficient to understand the phenomenology.

3.1.4 α_S measurements

As I mentioned above, the experimental measurement of $R_{e^+e^-}$ gives one of the best measurements of α_s. In fact the LEP combined value of R_{had} is

$$R(LEP) = 20.767 \pm 0.025, \tag{3.31}$$

while the tree-level prediction is

$$R_0(M_z) = 19.984. \tag{3.32}$$

Combining the two, and simply using the leading order result with $\mu = M_z$, we obtain our first measurement of α_S,

$$\alpha_S(M_z) = 0.124 \pm 0.004, \tag{3.33}$$

surprisingly close to the value using the four-loop result [5], 0.119 ± 0.003.

As we discussed in Section 1.8.3, since QCD predicts the scale dependence of α_S, one measurement at any scale is sufficient to give a prediction for all scales. We can therefore phrase measurements

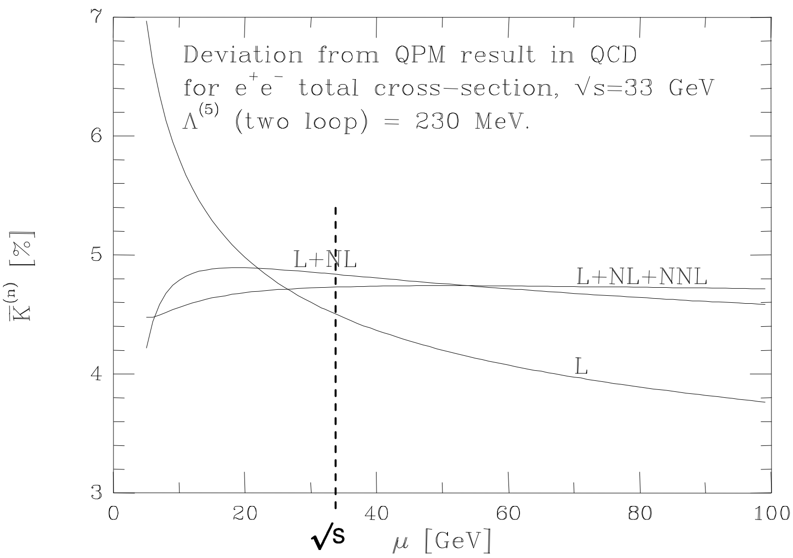


Fig. 3.2: The QCD prediction for the corrections to $R_{e^+e^-}$ at $\sqrt{s} = 33$ GeV as a function of renormalization scale at leading, next-to-leading, and next-to-next-to-leading order, taken from ESW [1]

at other scales either as tests of QCD throughout the intervening energy range or, by translating them all into measurements at a single scale, as different measurements of the same quantity that can be combined to give a better overall measurement.

As an example, the average measurement of R over several energy points around 34 GeV is

$$R(PETRA) = 3.88 \pm 0.03, \tag{3.34}$$

while the tree-level prediction is

$$R_0(34\text{ GeV}) = 3.69. \tag{3.35}$$

Again using the leading order result, we obtain

$$\alpha_\text{S}(34\text{ GeV}) = 0.162 \pm 0.026. \tag{3.36}$$

Finally, using the one-loop renormalization group equation, we can convert this into a measurement of $\alpha_\text{S}(M_z)$,

$$\alpha_\text{S}(M_z) = 0.134 \pm 0.018. \tag{3.37}$$

This is in good agreement with the value from LEP, although with much larger uncertainties, simply due to the fact that the statistics of the PETRA experiments were much lower.

As a final example, we consider τ decays. The QCD corrections to the hadronic decay rate actually have two effects: on the ratio of branching fractions, R_τ, as discussed earlier, and also directly on the total decay rate of the τ. These can form the basis for two analyses in which the experimental errors are largely independent. The combined result for the two is

$$\alpha_\text{S}(M_\tau = 1.77\text{ GeV}) = 0.34 \pm 0.01. \tag{3.38}$$

This time, because we are translating over such a wide energy range the one-loop renormalization group equation does not do quite such a good job,

$$\alpha_\text{S}^{\text{(one-loop)}}(M_z) = 0.1272 \pm 0.0014, \tag{3.39}$$

compared to the four-loop value [5]

$$\alpha_\text{S}(M_z) = 0.1212 \pm 0.0011, \tag{3.40}$$

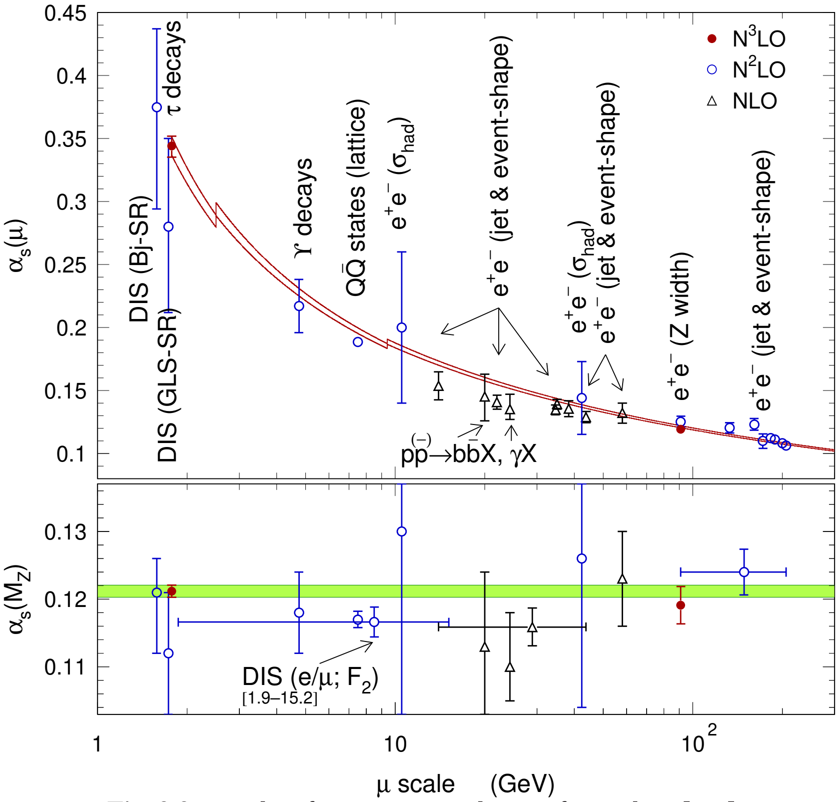


Fig. 3.3: Results of a recent compilation of α_{S} values [5, 6]

but it is not so far out. Note in this case the phenomenon of the 'incredible shrinking error'. Although the measurement at the τ mass scale has a precision of about 3%, after evolving it to M_z the relative uncertainty gets scaled down by the ratio of the two α_s values, and τ decays give the best single measurement of $\alpha_s(M_z)$.

The results of a recent compilation [5, 6] are shown in Fig. 3.3. The scale dependence shows excellent agreement with the predictions of perturbative QCD over a wide energy range. When translated into measurements of $\alpha_{\mathrm{S}}(M_z)$, the separate measurements cluster strongly around the average value,

$$\alpha_{\mathrm{S}}^{(\mathrm{average})}(M_z) = 0.1204 \pm 0.0009. \tag{3.41}$$

3.2 Deep inelastic scattering revisited

The parton model I described in the last lecture assumed that the partons are non-interacting. But we know that they do interact via QCD, so what will happen when we consider these interactions? We will discover that the structure functions do become slowly (logarithmically) varying with Q^2. We start by considering the next-to-leading order QCD corrections to quark scattering. We will find that these, if calculated naively, would be divergent, but that these divergences can be absorbed into the parton distribution functions. These will then become scale-dependent, giving rise to the Q^2-dependence of the structure functions.

3.2.1 NLO corrections to DIS

The next-to-leading order corrections come from three sources (recalling that we sum and integrate over all final states X, so we must sum over all contributions in which any kind of parton is scattered):

1. One-loop corrections to $eq \to eq$,

2. $eq \to eqg$,
3. $eg \to eq\bar{q}$.

The third contribution is completely new in QCD and is not present in the parton model. We come back to it in a later section. The other two can more genuinely be thought of as higher-order corrections to the parton model process. We start with the second.

There are two contributing diagrams. The matrix element squared can be obtained by crossing from $e^+e^- \to q\bar{q}g$ (1.66). Labeling the momenta as

$$e(k) + q(\eta p) \to e(k') + q(p_1) + g(p_2), \tag{3.42}$$

we obtain

$$\sum |\mathcal{M}|^2 = \frac{8C_F N_c e^4 e_q^2 g_s^2}{k\cdot k'\, p_1\cdot p_2\, \eta p\cdot p_2} \left((p_1\cdot k)^2 + (\eta p\cdot k)^2 + (p_1\cdot k')^2 + (\eta p\cdot k')^2\right). \tag{3.43}$$

As usual the phase space is given by (2.24),

$$dPS = \frac{Q^2}{16\pi^2 s x^2} dQ^2\, dx\, dPS_X. \tag{3.44}$$

This time X consists of two partons so is non-trivial,

$$dPS_X = \frac{d\cos\theta\, d\phi}{32\pi^2}, \tag{3.45}$$

where θ and ϕ refer to the direction of p_1 in the centre-of-mass system of $\eta p + q$. It is conventional to replace $\cos\theta$ by the manifestly Lorentz-invariant variable z,

$$z \equiv \frac{p_1\cdot p}{q\cdot p} = \tfrac{1}{2}(1 - \cos\theta), \tag{3.46}$$

with range $0 < z < 1$, giving

$$dPS_X = \frac{dz\, d\phi}{16\pi^2}. \tag{3.47}$$

It will later be instructive to know the transverse momentum of p_1 in this frame,

$$k_\perp^2 = Q^2 \left(\frac{\eta}{x} - 1\right) z(1 - z). \tag{3.48}$$

Note also that the case $\eta = x$ corresponds to a massless final state. Kinematically this can only happen if either p_1 or p_2 are infinitely soft (i.e., have zero energy), or if they are exactly collinear.

We therefore have

$$\frac{d\sigma^2(e+q)}{dx\, dQ^2} = \frac{1}{4N_c}\,\frac{1}{2\hat{s}}\,\frac{Q^2}{16\pi^2 s x^2} \int \frac{dz\, d\phi}{16\pi^2}\,\frac{8C_F N_c e^4 e_q^2 g_s^2}{k\cdot k'\, p_1\cdot p_2\, \eta p\cdot p_2} \left((p_1\cdot k)^2 + (\eta p\cdot k)^2 + (p_1\cdot k')^2 + (\eta p\cdot k')^2\right). \tag{3.49}$$

Rewriting in terms of our kinematic variables and averaging over ϕ, we have

$$\begin{aligned}
&\left\langle \frac{(p_1\cdot k)^2 + (\eta p\cdot k)^2 + (p_1\cdot k')^2 + (\eta p\cdot k')^2}{k\cdot k'\, p_1\cdot p_2\, \eta p\cdot p_2} \right\rangle_\phi \\
&= \frac{1}{y^2 Q^2} \left((1 + (1-y)^2) \left[\frac{1 + x_p^2}{1 - x_p}\,\frac{1 + z^2}{1 - z} + 3 - z - x_p + 11 x_p z \right] - y^2 \left[8 z x_p \right] \right),
\end{aligned} \tag{3.50}$$

where $x_p = x/\eta$. Two things are already clear: at this order we will have a non-zero longitudinal structure function, $F_L(x, Q^2)$; and the z integration, which runs from 0 to 1, will give a divergent contribution to F_2. This should worry us, since we are calculating a physical cross section, but let us continue for a while and see what happens.

Putting everything together we have

$$\frac{d\sigma^2(e+q)}{dx\,dQ^2} = \frac{C_F \alpha^2 e_q^2 \alpha_S}{2\eta x^2 y^2 s^2} \int_0^1 dz \left((1+(1-y)^2) \left[\frac{1+x_p^2}{1-x_p} \frac{1+z^2}{1-z} + 3 - z - x_p + 11 x_p z \right] - y^2 \left[8 z x_p \right] \right), \tag{3.51}$$

and hence

$$F_2(x,Q^2) = \sum_q \int_x^1 dx_p \, e_q^2 \frac{x}{x_p} f_q\left(\frac{x}{x_p}\right) \frac{C_F \alpha_S}{2\pi} \int_0^1 dz \left(\frac{1+x_p^2}{1-x_p} \frac{1+z^2}{1-z} + 3 - z - x_p + 11 x_p z \right). \tag{3.52}$$

The divergence at $z \to 1$ corresponds to kinematic configurations in which the outgoing gluon becomes exactly collinear with the incoming quark. Therefore in the Feynman diagram in which the gluon is attached to the incoming quark, the internal quark line becomes on-shell, causing the divergence. Note also that the coefficient of the divergence itself diverges at the point $x_p = 1$, at which the gluon is infinitely soft.

In order to study the divergence, let us first regulate it by calculating the contribution from emission with $k_\perp^2 > \mu^2$ (and assume $\mu^2 \ll Q^2$ for simplicity). Since $k_\perp^2$ is proportional to $(1-z)$ this will give us finite integrals. At any time, the full result can be obtained by setting $\mu \to 0$. We therefore obtain

$$F_2(x,Q^2) = \sum_q \int_x^1 dx_p \, e_q^2 \frac{x}{x_p} f_q\left(\frac{x}{x_p}\right) \frac{\alpha_S}{2\pi} \left(\hat{P}(x_p) \log \frac{Q^2}{\mu^2} + R(x_p) \right), \tag{3.53}$$

where the function $R(x_p)$ is finite. In the following we will not keep track of this function, although it would be essential for quantitative analysis. The function $P(x_p)$ we introduced in (3.53) is called the splitting function (or more strictly speaking the unregularized splitting function),

$$\hat{P}(x) = C_F \frac{1+x^2}{1-x}. \tag{3.54}$$

It actually describes the probability distribution of quarks produced in a splitting process, $q \to qg$ in which the produced quark has a fraction x of the original quark's momentum. (We will quantify this statement slightly more shortly.)

Obviously by regulating the divergence we have not removed it: physical cross sections are still supposed to be obtained by setting $\mu \to 0$, in which case F_2 is logarithmically divergent. However, before discussing what happens to this divergence, let us consider the virtual one-loop correction to $eq \to eq$. Since this diagram contains two quark-gluon couplings, when squared it would give an $\mathcal{O}(\alpha_S^2)$ correction. However, since it has the same final state as the lowest order diagram, we must consider the interference between the two, and this interference is $\mathcal{O}(\alpha_S)$, so we must include it.

We could obtain the result for the one-loop diagram by crossing from $e^+e^- \to q\bar{q}$. However, to illustrate the physics, it is sufficient to recall a few of its features. Firstly, since the external particles are the same as in the lowest-order process, the kinematics must be the same. In particular, it can only contribute at the point $\eta = x$. Secondly, as in the e^+e^- case, the interference of the one-loop and tree-level diagrams is divergent and negative. In fact the kinematic regions in which the one-loop integrand diverges are exactly the same as those of the $eq \to eqg$ contribution we have just considered: when the gluon is soft, or is collinear with either of the quarks.

It turns out that the divergence is exactly right to cancel the one we obtained above at $x_p \to 1$. In fact one finds that after including the one-loop contribution, one gets exactly the same formula as (3.53) except that the unregularized splitting function $\hat{P}(x_p)$ is replaced by the regularized one, $P(x_p)$,

$$P(x) = \hat{P}(x) + P_{\text{virtual}}(x). \tag{3.55}$$

Since the one-loop contribution has the same kinematics as the lowest-order process, $P_{\text{virtual}}(x)$ must be proportional to $\delta(1-x)$. $P(x)$ is therefore a distribution.

To define it, we will need to use a mathematical trick called the plus-distribution. Given some function $f(x)$, which is well-defined for all $0 \le x < 1$, we define a distribution $f(x)_+$ on the region $0 \le x \le 1$, as

$$f(x)_+ = f(x) - \delta(1-x) \int_0^1 dx' f(x'). \tag{3.56}$$

The plus-distribution is most useful when the function $f(x)$ is divergent at $x \to 1$. This means that for any other function $g(x)$, which is smooth at $x = 1$, we have the property

$$\int_0^1 dx\, f(x)_+ \, g(x) = \int_0^1 dx\, f(x) \left(g(x) - g(1) \right). \tag{3.57}$$

Provided that $g(x)$ goes to $g(1)$ sufficiently quickly, this integral is finite.

After including the virtual contribution, the splitting function is given by

$$P(x) = C_F \left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x) \right]. \tag{3.58}$$

This is actually the first correction to a function that describes the momentum distribution of quarks within quarks,

$$\mathcal{P}(x) \equiv \delta(1-x) + \frac{\alpha_{\mathrm{S}}}{2\pi} \log \frac{Q^2}{Q_0^2} P(x) + \mathcal{O}(\alpha_{\mathrm{S}}^2 \log^2), \tag{3.59}$$

where the distribution is defined to be a pure quark at scale Q_0 and probed at scale Q.

Inserting the full splitting function into (3.53), we find that the divergence at $x_p \to 1$ cancels between the real and virtual terms, but the divergence for $x_p < 1$ due to the region $z \to 1$ still remains.

3.2.2 Factorization of divergences

To understand why the results are still divergent even after including the virtual terms, and what ultimately happens to the divergences, we consider their physical origin. Like the e^+e^- annihilation case, we have singularities from regions in which the real gluon is collinear with either the incoming or outgoing quark, or is soft, and also from the virtual graph, as illustrated in Fig. 3.4. All these contributions were present in e^+e^- annihilation, but there we found that the divergences all cancelled to give a finite contribution. Why is the present situation different? In fact we find that here the magnitudes of the divergences are such as to cancel, but that the divergences arise in different regions of the x_p integral, so are prevented from cancelling.

In the e^+e^- case, we argued that the singular regions of real emission were indistinguishable from the lowest order process, since an infinitely soft gluon could not produce any hadrons and the jets produced by two collinear partons were indistinguishable from a single jet with their combined momentum. This statement is true here for the soft and final-state collinear contributions, but not the initial-state contribution. The final state of this contribution is indeed indistinguishable from the lowest order process (it has an additional jet collinear with the outgoing proton remnant, but this too gives a jet and the superposition of the two is indistinguishable from the proton remnant in the lowest order process). However, because we have used the parton model, we must convolute the partonic cross sections over an arbitrary (measured from experiment) probability distribution function, processes with different incoming momenta are effectively distinguishable. In all the singular regions, the final state of the process is massless, and this fact fixes the incoming momentum (to the value $Q/2x$ in the Breit frame), but in the initial-state singular process it is the internal line whose momentum gets fixed, as indicated in Fig. 3.4. Thus the incoming momentum in (c) is larger than in the other cases and its divergence, at $\eta > x$, cannot cancel the others, at $\eta = x$.

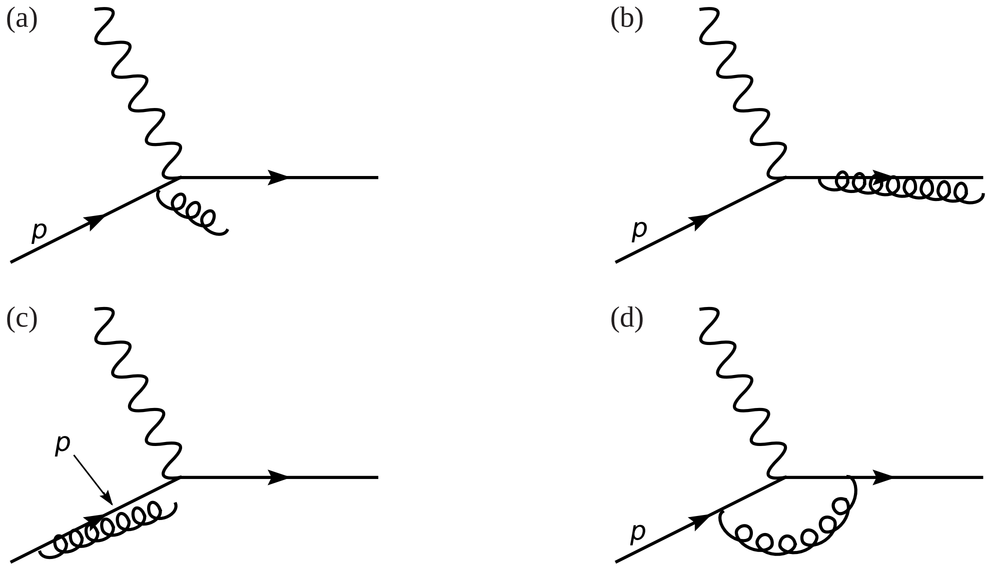


Fig. 3.4: Divergent contributions to DIS: (a) soft, (b) final-state collinear, (c) initial-state collinear, and (d) virtual. The label p shows which momentum in each contribution is fixed by the massless final-state condition.

As I mentioned earlier, these divergences come from the virtuality of the internal particle vanishing and hence the propagator diverging. Using the uncertainty principle, vanishingly small virtuality corresponds to arbitrarily long time-scales. This seems to be in direct contradiction with the assumption underlying the parton model, that the virtual photon takes an extremely rapid snapshot of the proton.

The problem is actually one of overcounting. We first introduced the pdfs, which are supposed to contain all information about the internal structure of the proton. Presumably this internal structure is the result of QCD interactions. We then tried to calculate the QCD corrections to the quark scattering cross sections, integrating over all final states, so all energy-scales. But these QCD corrections should somehow already be included in the internal dynamics of the proton.

To resolve this overcounting, we have to separate (or 'factorize') the different types of physics at different energy scales. Like in our discussion of renormalization, I will first try to give the physical picture in terms of a cutoff, before returning later to describe how factorization works in practice in dimensional regularization. We introduce the factorization scale μ, and call all physics at scales below μ part of the hadron wave function, and lump it into the parton distribution functions, and call all physics at scales above μ part of the partonic scattering cross section (or 'coefficient function').

Therefore we do in fact have a transverse momentum cutoff in the $eq \to eqg$ process and the form of (3.53) is correct.

Since physics at scales below μ is included in the pdfs and physics above is not, the pdfs themselves must become μ-dependent. We therefore have

$$F_2(x, Q^2) = \sum_q e_q^2 \int_x^1 dx_p \frac{x}{x_p} f_q\left(\frac{x}{x_p}, \mu^2\right) \left\{ \delta(1 - x_p) + \frac{\alpha_S}{2\pi}\left(P(x_p) \log \frac{Q^2}{\mu^2} + R(x_p)\right) + \mathcal{O}(\alpha_S^2) \right\}, \tag{3.60}$$

where the function $R(x_p)$ is not necessarily the same one as earlier, as the virtual contributions could have introduced some additional finite terms.

Note that the structure functions are now Q^2-dependent, violating Bjorken scaling. However, they also appear to be μ^2 dependent, which should worry us: μ was introduced in a completely ad hoc theoretical way: it simply separates physical processes into two parts that are dealt with in different ways, and the final result, which is the sum of the two parts, should not depend on where the separation was made. We return to discuss this point in more detail after calculating the μ^2-dependence of the pdfs.

It is important to emphasize that, although we have derived these formulae for the higher order corrections to DIS, the leading logarithmic behaviour is universal. In particular, for any quark-induced process with a hard scale Q, we expect a hadronic cross section of the form

$$\sigma_h(p_h) = \sum_q \int d\eta\, f_q\left(\eta, \mu^2\right) \left\{ \sigma_q(\eta p_h) + \frac{\alpha_S}{2\pi} \log \frac{Q^2}{\mu^2} \int dz\, P(z)\, \sigma_q(z\eta p_h) \right\}, \tag{3.61}$$

where $\sigma_q(p)$ is the partonic cross section for a quark of flavour q and momentum p.

3.2.3 DGLAP evolution equation

Although the pdfs are fundamentally non-perturbative and cannot be predicted from first principles at present, physics at scales close to μ^2 can be described perturbatively. We can therefore calculate the μ^2-dependence of the pdfs so that, given their value at some starting scale μ_0, for example from experimental measurements, we can calculate their values at all higher scales μ.

To do this, we use the fact just noted, that physical cross sections should not depend on μ^2. Therefore we should have

$$\mu^2 \frac{dF_2(x, Q^2)}{d\mu^2} = 0, \tag{3.62}$$

or at least, since we are working at $\mathcal{O}(\alpha_S)$,

$$\mu^2 \frac{dF_2(x, Q^2)}{d\mu^2} = \mathcal{O}(\alpha_S^2). \tag{3.63}$$

Applying this to (3.60), we obtain

$$\mu^2 \frac{d}{d\mu^2} f_q\left(x, \mu^2\right) = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dx_p}{x_p} f_q\left(\frac{x}{x_p}, \mu^2\right) P(x_p) + \mathcal{O}(\alpha_S^2). \tag{3.64}$$

Equation (3.64) is called the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (or DGLAP, or GLAP, or Altarelli–Parisi for short) evolution equation. Note that the rate of change of the pdf at some x value depends on its value at all higher xs.

To understand its physical content, it is useful to rewrite the splitting function,

$$P(x) = C_F \left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x) \right] = C_F \left(\frac{1+x^2}{1-x} \right)_+, \tag{3.65}$$

to give

$$\mu^2 \frac{d}{d\mu^2} f_q\left(x, \mu^2\right) = C_F \frac{\alpha_S}{2\pi} \int_x^1 \frac{dx_p}{x_p} f_q\left(\frac{x}{x_p}, \mu^2\right) \frac{1+x_p^2}{1-x_p} - C_F \frac{\alpha_S}{2\pi} f_q\left(x, \mu^2\right) \int_0^1 dx_p \frac{1+x_p^2}{1-x_p}. \tag{3.66}$$

The first term represents the fact that the pdf at a given x value is increased by quarks with higher x's reducing their momentum fractions by radiating gluons. The second term represents the fact that it is decreased by the quarks at that x reducing their momentum fractions by radiating gluons. Each contribution is divergent due to emission with $x_p \to 1$, i.e., infinitely soft gluon emission, involving an infinitely small change in x. However the two divergences exactly cancel because the number of quarks being lost to this x value by infinitely soft gluon emission is equal to the number being gained.

The DGLAP equation is most easily solved in moment space. For any function $f(x)$, we define

$$f_N = \int_0^1 dx\, x^{N-1}\, f(x), \tag{3.67}$$

the Mellin transform. Taking moments of both sides of (3.64), we obtain

$$\mu^2 \frac{d}{d\mu^2} f_{qN}\left(\mu^2\right) = \frac{\alpha_S}{2\pi} \int_0^1 dx\, x^{N-1} \int_x^1 \frac{dx_p}{x_p} f_q\left(\frac{x}{x_p}, \mu^2\right) P(x_p) + \mathcal{O}(\alpha_S^2) \quad (3.68)$$

$$= \frac{\alpha_S}{2\pi} P_N f_{qN}(\mu^2). \quad (3.69)$$

It is common to introduce the notation

$$\gamma_N(\alpha_S) = \frac{\alpha_S}{2\pi} P_N + \mathcal{O}(\alpha_S^2), \quad (3.70)$$

where γ_N is known as the anomalous dimension. If we assume that the coupling α_S is fixed, we can easily solve (3.69) with the boundary condition of given values for f_{qN} at some fixed scale μ_0,

$$f_{qN}(\mu^2) = f_{qN}(\mu_0^2) \left(\frac{\mu^2}{\mu_0^2}\right)^{\gamma_N(\alpha_S)}. \quad (3.71)$$

However, as we have seen, the renormalization of QCD means that the coupling constant becomes scale dependent, $\alpha_S(\mu^2)$, according to renormalization group equation

$$\mu^2 \frac{d}{d\mu^2} \alpha_S(\mu^2) = \beta(\alpha_S(\mu^2)) = -\frac{\beta_0}{2\pi} \alpha_S^2(\mu^2) + \mathcal{O}(\alpha_S^3). \quad (3.72)$$

Inserting the solution of the running coupling, Eq. (1.73), into (3.69), we obtain

$$f_{qN}(\mu^2) = f_{qN}(\mu_0^2) \left(\frac{\alpha_S(\mu_0)}{\alpha_S(\mu)}\right)^{\frac{P_N}{\beta_0}}. \quad (3.73)$$

Having the solution for f_q in moment N-space, we have to convert it back to x-space. This is done by the Inverse Mellin Transform, where f_{qN} is continued to the complex plane,

$$f_q(x) = \frac{1}{2\pi i} \int_C dN\, f_{qN}\, x^{-N}, \quad (3.74)$$

where the contour C runs parallel to the imaginary axis to the right of all poles. Because of the complexity of this process, the DGLAP equation is often solved simply by ‘brute force’ numerical solution of (3.64).

Beyond $\mathcal{O}(\alpha_S)$ the general structure of (3.69) and (3.72) remains unchanged: the anomalous dimension and β function simply become power series in α_S.

3.2.4 Scheme/scale dependence

Factorization, as introduced above, may seem pretty arbitrary. However it can be proved to all orders in perturbation theory. The most convenient way to do this is to use, instead of the transverse momentum cutoff we used above, dimensional regularization. When we calculate the NLO cross section in d dimensions, the divergence shows up as a pole, $1/\epsilon$. The coefficient multiplying this pole turns out to be the same splitting function we encountered earlier.

In d dimensions, we obtain for the structure function up to $\mathcal{O}(\alpha_S)$,

$$F_2(x, Q^2) = \sum_q e_q^2 \int_x^1 dx_p \frac{x}{x_p} \bar{f}_q\left(\frac{x}{x_p}\right) \left\{\delta(1-x_p) + \frac{\alpha_S}{2\pi}\left(\left(\frac{4\pi\mu^2}{Q^2}\right)^{\epsilon} \frac{-1}{\epsilon} P(x_p) + R(x_p)\right) + \mathcal{O}(\epsilon)\right\}, \quad (3.75)$$

where μ is the scale introduced to make the coupling constant dimensionless. Note that I have sneakily added a bar to f_q and that it is scale independent. $\bar{f}_q$ is known as the bare pdf. We now note that the distribution functions themselves are not physical observables, only their convolution with coefficient functions is. I can therefore define a modified set of distribution functions as follows:

$$x\,\bar{f}_q(x) \equiv \int_x^1 dx_p \frac{x}{x_p} f_q\left(\frac{x}{x_p}, \mu_F^2\right) \left\{\delta(1-x_p) - \frac{\alpha_{\mathrm{S}}}{2\pi}\left(\left(\frac{4\pi\mu^2}{\mu_F^2}\right)^{\epsilon} \frac{-1}{\epsilon} P(x_p) + K(x_p)\right)\right\}, \quad (3.76)$$

where μ_F is the (completely arbitrary again) factorization scale, and $K(x_p)$ is a completely arbitrary finite function to be discussed shortly. (To fit in with the standard notation, I should really multiply all occurrences of α_{S} by $1/\Gamma(1-\epsilon) = 1 - \gamma_E\epsilon + \mathcal{O}(\epsilon^2)$, but this will merely change the values of $R(x_p)$ and $K(x_p)$ which I do not specify anyway.) Combining (3.75) and (3.76), we end up with

$$F_2(x,Q^2) = \sum_q e_q^2 \int_x^1 dx_p \frac{x}{x_p} f_q\left(\frac{x}{x_p}, \mu_F^2\right) \left\{\delta(1-x_p) \right.$$
$$\left. + \frac{\alpha_{\mathrm{S}}}{2\pi}\left(P(x_p)\log\frac{Q^2}{\mu_F^2} + R(x_p) - K(x_p)\right) + \mathcal{O}(\alpha_{\mathrm{S}}^2)\right\}. \quad (3.77)$$

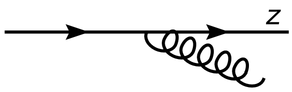
Note that this has the identical form to (3.60), except for the finite function. It is clear from (3.76) that $f_q(x,\mu_F^2)$ depends on the function $K(x_p)$. It therefore seems like we have no predictive power: the pdf and coefficient function each depend on the completely arbitrary function $K(x_p)$ and the completely arbitrary scale μ_F (note that all dependence on μ has again completely cancelled. As I said in the context of renormalization, many textbooks simply set it equal μ right from the start, but I consider this slightly confusing as they have quite different physical meaning. Having performed this manoeuvre, I henceforth drop the subscript $_F$). However, the factorization theorem proves, firstly that for any physical quantity, all dependence on $K(x_p)$ and μ will cancel and secondly that the scheme- and scale-dependent pdfs, $f_q(x,\mu^2)$ are universal (i.e., process-independent).

Two schemes are in common use, the $\overline{\mathrm{MS}}$ scheme in which $K(x_p)$ is zero, and the DIS scheme in which $K(x_p) = R(x_p)$, i.e. in which for $\mu = Q$ the parton model result is exact.

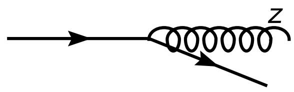
To understand the physical content of the scheme-dependence, it is worth while going back to the case with a cutoff. If, instead of a cut on transverse momentum we had used a cut on the virtuality of the internal quark line to separate the pdf from the coefficient function, we would have got exactly the same form as (3.60) except that $R(x_p)$ would have been a different function. In particular, it would differ by a $\log[(1-x_p)/x_p]$ term, together with some non-logarithmic terms. In fact, all logarithmic terms turn out to be the same with a p_t cutoff as in the $\overline{\mathrm{MS}}$ scheme, so for many purposes the two can be considered equivalent.

Although dependence on the scheme and scale must cancel in physical quantities, it is only guaranteed to do so after calculating to infinite orders of perturbation theory. At any finite order there can be some residual dependence. We must therefore have a procedure for choosing a value of μ. Essentially the identical discussion we had for the renormalization scale choice applies here. One can show that a structure like (3.60) continues to all orders of perturbation theory and that for every power of α_{S}, one gets a power of $\log Q^2/\mu^2$. Thus every order of perturbation theory contains terms like $\alpha_{\mathrm{S}}^n \log^m Q^2/\mu^2$, $m \le n$. It is clear that if μ is a long way from Q, the log can be large enough to compensate the smallness of α_{S} and the perturbative series will not converge quickly. One should therefore choose μ 'not too far' from Q.

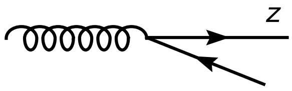
It is worth mentioning that one can set up DGLAP evolution equations for the Q^2-dependence of the structure functions, F_2 and F_L, themselves. These are then automatically scheme- and scale-independent even at finite orders of perturbation theory. This is sometimes known as the scheme-independent scheme.



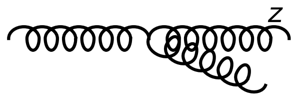
$$P_{qq}(x) = C_F \left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x) \right]$$



$$P_{qg}(x) = T_R \left[x^2 + (1-x)^2 \right]$$



$$P_{gq}(x) = C_F \left[\frac{1+(1-x)^2}{x} \right]$$



$$P_{gg}(x) = C_A \left[\frac{2x}{(1-x)_+} + 2\frac{1-x}{x} + 2x(1-x) \right] + \beta_0 \delta(1-x)$$

Fig. 3.5: The four DGLAP splitting functions of QCD

3.2.5 Initial-state gluons

As mentioned right at the start of this section, we also obtain $\mathcal{O}(\alpha_S)$ corrections from the process $eg \to eq\bar{q}$. Most of what we said above carries over in a straightforward way. Although there is no soft singularity or virtual term to cancel it, there is a collinear singularity. This corresponds to a two-step process in which a gluon splits to a q–$\bar{q}$ pair, one of which interacts with the photon. The singularity again corresponds to the virtuality of the internal quark line going to zero. This singularity can again be absorbed into a factorized universal pdf for the gluon. We end up with an additional contribution to the structure function of

$$F_2(x,Q^2) = \sum_q e_q^2 \int_x^1 dx_p \frac{x}{x_p} f_g\left(\frac{x}{x_p},\mu^2\right) \left\{ \frac{\alpha_S}{2\pi} \left(P_{qg}(x_p) \log\frac{Q^2}{\mu^2} + R_g(x_p) - K_{qg}(x_p) \right) + \mathcal{O}(\alpha_S^2) \right\}, \tag{3.78}$$

where the sum over q is over all 'light' flavours. We now have four different types of splitting function, illustrated in Fig. 3.5. The DGLAP equation now becomes a set of coupled equations:

$$\mu^2 \frac{d}{d\mu^2} f_a\left(x,\mu^2\right) = \sum_b \frac{\alpha_S}{2\pi} \int_x^1 \frac{dx_p}{x_p} f_b\left(\frac{x}{x_p},\mu^2\right) P_{ab}(x_p) + \mathcal{O}(\alpha_S^2). \tag{3.79}$$

In moment space, this can be conveniently written as a matrix equation (in general of $(2N_f{+}1)\times(2N_f{+}1)$ matrices, but for simplicity we show the case of only one flavour of quark):

$$\mu^2 \frac{d}{d\mu^2} \begin{pmatrix} f_{qN} \\ f_{\bar{q}N} \\ f_{gN} \end{pmatrix} = \begin{pmatrix} \gamma_{qqN}(\alpha_S(\mu)) & 0 & \gamma_{qgN}(\alpha_S(\mu)) \\ 0 & \gamma_{qqN}(\alpha_S(\mu)) & \gamma_{qgN}(\alpha_S(\mu)) \\ \gamma_{gqN}(\alpha_S(\mu)) & \gamma_{gqN}(\alpha_S(\mu)) & \gamma_{ggN}(\alpha_S(\mu)) \end{pmatrix} \begin{pmatrix} f_{qN} \\ f_{\bar{q}N} \\ f_{gN} \end{pmatrix}. \tag{3.80}$$

Exactly the same solution is obtained, but in matrix notation,

$$\begin{pmatrix} f_{qN}(\mu^2) \\ f_{\bar{q}N}(\mu^2) \\ f_{gN}(\mu^2) \end{pmatrix} = \exp \int_{\mu_0^2}^{\mu^2} \frac{d\mu'^2}{\mu'^2} \begin{pmatrix} \gamma_{qqN}(\alpha_S(\mu')) & 0 & \gamma_{qgN}(\alpha_S(\mu')) \\ 0 & \gamma_{qqN}(\alpha_S(\mu')) & \gamma_{qgN}(\alpha_S(\mu')) \\ \gamma_{gqN}(\alpha_S(\mu')) & \gamma_{gqN}(\alpha_S(\mu')) & \gamma_{ggN}(\alpha_S(\mu')) \end{pmatrix} \begin{pmatrix} f_{qN}(\mu_0^2) \\ f_{\bar{q}N}(\mu_0^2) \\ f_{gN}(\mu_0^2) \end{pmatrix}. \tag{3.81}$$

This is even more troublesome to do by the Inverse Mellin Transform, so the full set of DGLAP equations is almost always solved numerically.

Note that at higher orders of perturbation theory, even the zero entries in (3.80) become non-zero, as do contributions like $P_{qq'}(x)$.

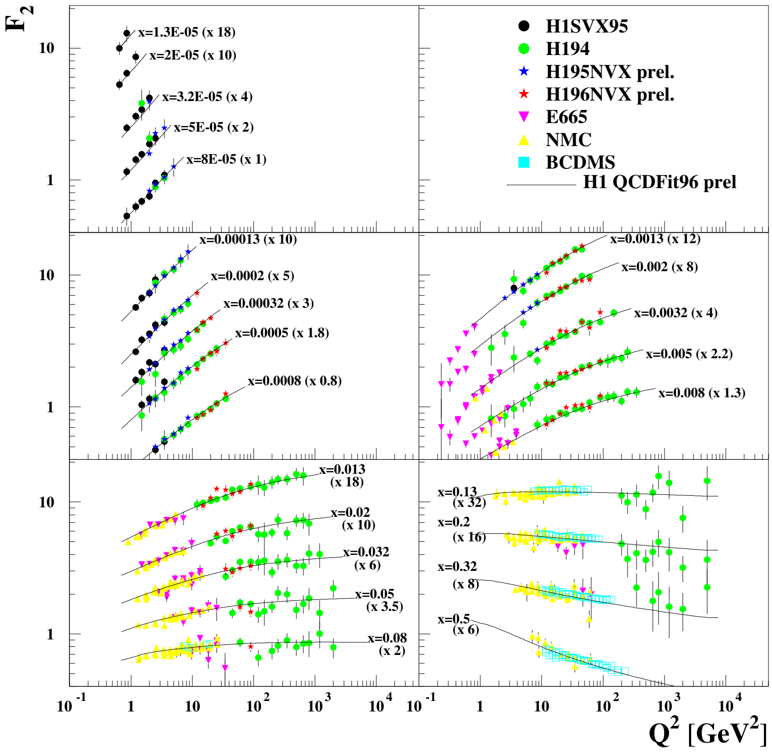


Fig. 3.6: Fit to the F_2 data over a wide range of Q^2 values, exhibiting violation of Bjorken scaling

3.2.6 Violation of Bjorken scaling

As we already noted, the factorization of initial-state singularities introduces a logarithmic Q^2 dependence into the structure functions and therefore a slow violation of Bjorken scaling. There is a close analogy with the renormalization of one-scale cross sections, where the energy-dependence was entirely due to the quantum corrections. Although the pdfs at some low scale are entirely non-perturbative and must be fit to data, the scale-dependence is entirely predicted by QCD and provides a stringent test over a wide range of energy scales. The result is impressive, see Fig. 3.6.

3.3 Summary

NLO calculations are hard! This is mainly because the real and virtual corrections are each divergent and must be regularized in some self-consistent way, for example with dimensional regularization. Unlike the ultraviolet divergences, which are isolated in well-localized pieces of the loop calculation and can effectively be removed by a redefinition of the Feynman rules, these divergences arise in different partonic contributions to physical observables. They must therefore be kept explicit until the very end of the calculation when all the partonic contributions are combined. Only then, provided our observable is infrared safe, will the real and virtual divergences cancel to yield a finite result.

Processes with incoming partons have extra divergences, arising from a miscancellation of the initial-state-collinear real and virtual contributions, which appear at different points in the integral over incoming momentum fraction. (It is worth mentioning that the same argument applies to the final-state distributions of identified hadrons, for example the momentum distribution of pions produced in e^+e^- annihilation.) These divergences have to be factorized into the non-perturbative, but universal, parton distribution functions at some factorization scale μ_F. This extra scale in the structure functions allows them to be Q^2-dependent. This Q^2-dependence is entirely driven by the μ_F^2-dependence of the parton distribution functions, which is predicted by the DGLAP evolution equations. Thus structure function data over a wide range of Q^2 provide a stringent test of perturbative QCD.

4 Summary

In this short course on QCD phenomenology, I have resisted the temptation to review the many important tests and studies of QCD that have been made over the years and have instead tried to concentrate on the key ideas that underpin them. These are:

– The gauge invariance of the theory, which allows us to write down the Lagrangian and which predicts some of the most important features of the theory: the universality of the coupling constant and the self-coupling of gluons, which ultimately leads to the negative β function and hence to asymptotic freedom at high energies and strong interactions at low energies.

– Renormalization and decoupling, which allow us to make predictive calculations at finite energy, without knowing the full structure of the theory to arbitrarily high energy and without the introduction of arbitrarily many input parameters. Renormalization is related to the quantum structure of the theory and introduces a dimensionful scale into even the scaleless Lagrangian of massless QCD, giving rise to energy-dependence of one-scale observables that would be energy-independent in the classical theory.

– Factorization and evolution, which allow us to use perturbation theory to calculate the interactions of hadrons, since all the non-perturbative physics gets factorized, into universal functions that can be measured in one process, like DIS, and then used to predict the cross sections for any other process. Again, this introduces a scale dependence into the parton model so that the structure functions of DIS, and other one-scale observables such as the Drell–Yan cross section, become scale dependent.

– Infrared safety, which ensure that the infrared singularities associated with soft and collinear emission cancel between real and virtual contributions, allowing the perturbative calculation of jet cross sections, without a detailed understanding of the mechanism by which partons become jets.

Together, these allow us to make sense of QCD, without having to solve the theory at all possible scales: unknown or uncalculable high- and low-energy effects can be renormalized, factorized and cancelled away. After all this, it is remarkable that most QCD phenomenology can be understood at least qualitatively from leading order perturbation theory with the one-loop renormalization group and DGLAP evolution equations. Higher order corrections, while essential for quantitative analysis, do not change this simple picture dramatically.

Acknowledgements

It is a pleasure to acknowledge the organizers, lecturers, tutors and students of the Latin American School of High Energy Physics at Recinto Quirama for making giving these lectures such an enjoyable experience.

References

[1] R.K. Ellis, W.J. Stirling, and B.R. Webber, *QCD and Collider Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Volume 8 (Cambridge University Press, 1996).

[2] M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, 1995).

[3] M. Gockeler, R. Horsley, A. C. Irving, D. Pleiter, P. E. L. Rakow, G. Schierholz, and H. Stuben, *A determination of the Lambda parameter from full lattice QCD*, Phys. Rev. D **73** (2006) 014513 [arXiv:hep-ph/0502212].

[4] M. Davier, S. Eidelman, A. Höcker, and Z. Zhang, *Confronting spectral functions from e+ e- annihilation and tau decays: Consequences for the muon magnetic moment*, Eur. Phys. J. C **27** (2003) 497 [arXiv:hep-ph/0208177].

[5] M. Davier, S. Descotes-Genon, A. Höcker, B. Malaescu, and Z. Zhang, *The determination of α_s from τ decays revisited*, Eur. Phys. J. C **56** (2008) 305 [arXiv:0803.0979 [hep-ph]].

[6] S. Bethke, *Experimental tests of asymptotic freedom*, Prog. Part. Nucl. Phys. **58** (2007) 351 [arXiv:hep-ex/0606035].