3
Basics of QCD

In this chapter, we review some basic properties of QCD that directly follow from its definition. This material is completely standard, and will form a foundation for the rest of the book. More details can be found in a standard textbook on quantum field theory, e.g., Peskin and Schroeder (1995), Srednicki (2007), Sterman (1993) and Weinberg (1995, 1996). For specific information on renormalization and the renormalization group see also my book on renormalization (Collins, 1984).

I first review how the theory is quantized and renormalized. Then I discuss the renormalization group (RG) and the calculation of the asymptotic freedom of QCD. A brief review of the flavor symmetries follows. Finally I show some of the complications that arise in perturbative calculations because some of the fields are much more massive than others.

3.1 Quantization

3.1.1 Definition; functional integral

A list of the fields of QCD and the formula for its gauge-invariant Lagrangian density (2.1) are sufficient to specify the theory, with the aid of general principles. Although there are some mathematical issues that have not been solved properly, it is standard to assume that the theory can be constructed (with some complications) through a functional integral. This gives Green functions, i.e., vacuum expectation values of time-ordered products of fields, as

$$\left\langle 0 \middle| T f[A, \psi, \bar{\psi}] \middle| 0 \right\rangle = \mathcal{N} \int \mathcal{D}A\, \mathcal{D}\psi\, \mathcal{D}\bar{\psi}\, e^{iS[A,\psi,\bar{\psi}]}\, f[A, \psi, \bar{\psi}]. \tag{3.1}$$

Here $f[A, \psi, \bar{\psi}]$ is a functional of the fields, e.g., a product $G^2(x)\, \bar{\psi}\psi(y)$. On the left-hand side, the fields are the quantum fields of QCD, time-ordered, while $|0\rangle$ is the true vacuum state. But on the right-hand side the fields are corresponding classical fields (Grassmann-valued in the case of the fermion fields ψ and $\bar{\psi}$). The normalization factor $\mathcal{N}$ is set so that $\langle 0|0\rangle = 1$.

From the Green functions can be reconstructed the state space and the operators. This includes an extraction of the particle content of the theory, from an examination of the positions of the poles in propagators and other Green functions. The S-matrix and scattering theory follow by the Lehmann-Symanzik-Zimmermann (LSZ) method. Note that the poles

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

of Green functions need not be the same as in free field theory, and so the particle content can be different from a free field theory of quarks and gluons.

3.1.2 Faddeev-Popov method; Feynman rules

The rules for Feynman perturbation theory are readily derived from the functional integral, with the Faddeev-Popov technique being used for gauge fixing. In this technique, a change of variables is used on sets of field configurations equivalent under gauge transformations. The implementation involves fermion scalar "ghost" fields, η_α and $\bar{\eta}_\alpha$. See most modern textbooks on QFT for details.

In the covariant gauges we will normally use, the gauge-invariant Lagrangian of (2.1) is replaced by

$$\mathcal{L} = \mathcal{L}_{\text{GI from (2.1)}} + \mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{GC}}, \tag{3.2}$$

where the gauge-fixing and "gauge-compensating" terms are

$$\mathcal{L}_{\text{GF}} = -\frac{1}{2\xi_0}(\partial \cdot A_{(0)\alpha})^2, \tag{3.3}$$

$$\mathcal{L}_{\text{GC}} = \partial_\mu \bar{\eta}_{0\alpha} \partial^\mu \eta_{0\alpha} + g_0 \partial^\mu \bar{\eta}_{0\gamma} f_{\alpha\beta\gamma} A^\beta_{(0)\mu} \eta_{0\alpha}, \tag{3.4}$$

in terms of bare quantities. This gives

$$\begin{aligned}\mathcal{L} = \bar{\psi}_0 (i \not{D} - m_0)\psi_0 - \frac{1}{4}(G^\alpha_{(0)\mu\nu})^2 - \frac{1}{2\xi_0}(\partial \cdot A^\alpha_{(0)})^2 \\ + \partial_\mu \bar{\eta}_{0\alpha} \partial^\mu \eta_{0\alpha} + g_0 \partial^\mu \bar{\eta}_{0\gamma} f_{\alpha\beta\gamma} A^\beta_{(0)\mu} \eta_{0\alpha}.\end{aligned} \tag{3.5}$$

Feynman rules for Green functions are derived in the usual way. In Sec. 3.2, we will formulate Feynman rules for renormalized Green functions with a counterterm method. Rules for elementary perturbation theory in terms of bare quantities can be obtained from those listed in Fig. 3.1 below by replacing each occurrence of $g\mu^\epsilon$ in that figure by the bare coupling g_0, and each renormalized quark mass m_f by the bare mass m_{0f}.

Note that without gauge fixing in the Lagrangian, Green functions of the elementary gauge-variant field operators are zero (Elitzur, 1975).

3.1.3 BRST symmetry

The full Lagrangian (3.2) is not gauge invariant, which considerably complicates the derivation and formulation of generalized Ward identities. The appropriate identities for non-abelian gauge theories were first found by Slavnov (1972) and Taylor (1971). The derivations were greatly simplified by Becchi, Rouet, and Stora (1975, 1976) and by Tyutin (1975), who discovered a new symmetry of the full Lagrangian.

This BRST symmetry is a supersymmetry, i.e., one that relates Bose and Fermi fields. It uses a parameter $\delta\lambda_0$ that takes its value in the fermionic part of some Grassmann algebra. For the gauge and matter fields, the BRST transformations are gauge transformations (2.4)

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

with $\omega_\alpha(x) = \eta_{0\alpha}(x)\delta\lambda_0$. Thus any gauge-invariant operator is also BRST invariant. The linear terms in the variation of the bare fields are

$$\delta_{\rm BRST}\psi_0 = -ig_0\eta_{0\alpha}\delta\lambda_0 t^\alpha\psi_0 = ig_0\eta_{0\alpha}t^\alpha\psi_0\delta\lambda_0, \tag{3.6a}$$

$$\delta_{\rm BRST}\bar\psi_0 = ig_0\bar\psi_0 t^\alpha\eta_{0\alpha}\delta\lambda_0, \tag{3.6b}$$

$$\delta_{\rm BRST}A^\alpha_{(0)\,\mu} = \left(\partial_\mu\eta_{0\alpha} + g_0 f_{\alpha\beta\gamma}\eta_{0\beta}A^\gamma_{(0)\,\mu}\right)\delta\lambda_0. \tag{3.6c}$$

The ghost and antighost fields transform as

$$\delta_{\rm BRST}\eta_{0\alpha} = -\tfrac{1}{2}g_0 f_{\alpha\beta\gamma}\eta_{0\beta}\eta_{0\gamma}\delta\lambda_0, \tag{3.6d}$$

$$\delta_{\rm BRST}\bar\eta_{0\alpha} = \frac{1}{\xi_0}\partial\cdot A_{0\alpha}\delta\lambda_0. \tag{3.6e}$$

It can readily be checked that the full Lagrangian is BRST invariant, up to a total derivative. With a slight exception, BRST transformations are also nilpotent. That is, applying successive BRST transformations with different anticommuting parameters $\delta\lambda_1$ and $\delta\lambda_2$ gives zero:

$$\left(\frac{\delta_{\rm BRST}}{\delta\lambda_0}\right)^2 \text{field} = 0. \tag{3.7}$$

The exception is that the second variation of $\bar\eta_0$ only vanishes after using the equation of motion for $A_{(0)}$; a third variation of the field is needed to get zero without use of the equations of motion.

A good formulation of the quantum theory associated with Faddeev-Popov quantization and BRST transformations is given by Nakanishi and Ojima (1990). In particular they give a full formulation of the conditions to be applied to physical quantum-mechanical states.

3.1.4 Relation to Euclidean lattice gauge theory

The functional integral contains an oscillating functional e^{iS}, and it can be defined by analytically continuing to Euclidean space-time, where the time coordinate becomes imaginary, $t = -i\tau$, and by then putting the theory on a lattice in a finite volume of space-time. The functional integral is then an ordinary finite-dimensional convergent integral (with suitable modifications for the fermion integrations). Numerical evaluation of these integrals by Monte-Carlo methods is the core of lattice gauge theory, a key technique for non-perturbative calculations in QCD (DeGrand and Detar, 2006).

The infinite-volume limit is an ordinary thermodynamic limit, but the continuum limit of zero lattice spacing is non-trivial, needing the use of renormalization: Sec. 3.2. However, there is not yet a completely rigorous proof that the limit exists.

The continuation back to real time is potentially problematic. Typical time dependence associated with high-energy states at large times, e^{-iEt}, corresponds to strongly suppressed exponentials $e^{-E\tau}$ in Euclidean time. Small errors in the Euclidean calculation, e.g., due

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

to the neglect of weak-interaction effects or purely numerical errors, do not automatically continue to small errors in the real-time formalism. Further research is clearly needed here. Euclidean lattice methods are not suitable for high-energy scattering problems.

For our purposes, it suffices to assume that some method exists to construct real-time functional integrals, as in (3.1).

3.2 Renormalization

Ultra-violet (UV) divergences appear in QCD (and in most other relativistic quantum field theories) when the continuum limit is taken. These were first found in perturbative calculations, but the divergences are a property of the exact theory, as is shown by a renormalization-group analysis, particularly using Wilsonian methods (Polchinski, 1984). The divergences are from large loop momenta, or, equivalently, from where interaction vertices approach each other in space-time. In renormalizable theories, like QCD, the divergences can be proved to be removed by a modification of the continuum limit, at least in perturbation theory.

1. The theory is first defined with a regulator[1] (or cutoff) of the UV divergences. Standard UV regulators are a non-zero lattice spacing or dimensional regularization.
2. All parameters of the theory consistent with its symmetries are made adjustable as functions of the cutoff. The parameters include the coefficients of terms like $i\bar{\psi}\partial_\mu\psi$.
3. When the limit of no UV cutoff is taken, the cutoff dependence of the parameters is chosen so as to remove the UV divergences and to obtain a non-trivial limiting theory.

Note that an entirely different status is to be given to the infra-red (IR) divergences that appear in perturbation theory for the S-matrix in theories such as QCD and QED that have massless fields. The S-matrix is derived given certain hypotheses about the large-time behavior of Green functions. But in a theory like QED with actual massless particles, these hypotheses are violated, while in QCD the particle content does not even correspond to the elementary fields. In either case, perturbative calculations must be adapted to the correct physics. But IR divergences do not affect the definition of the theory, only the interpretation of its solution, unlike the case of UV divergences.

The general ideas and methods of renormalization are explained in almost any modern QFT textbook, and a more specialized reference is Collins (1984), which is compatible with the presentation here.

3.2.1 Reformulating $\mathcal{L}$: bare parameters

To obtain finite Green functions, we use the freedom not only to change g_0 and m_0 in (3.5), but also to change the normalization of the fields, i.e., to do field strength renormalization.

[1] For mathematicians: In much of the mathematical literature, the word "regularization" has a different meaning, equivalent to physicists' "renormalization".

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

We therefore define the bare fields to be (square roots of) "wave-function renormalization" factors times renormalized fields: $A_{(0)\mu} = Z_3^{1/2} A_\mu$, $\psi_0 = Z_2^{1/2}\psi$, and $\eta_0 = \tilde{Z}^{1/2}\eta$. It is Green functions of the renormalized fields that are to be finite after removal of the UV cutoff. This gives the following formula for $\mathcal{L}$:

$$\begin{aligned}\mathcal{L} = {} & Z_2\bar{\psi}(i\partial\!\!\!/ - m_0)\psi - Z_2 Z_3^{1/2} g_0 \bar{\psi} t^\alpha A\!\!\!/^\alpha \psi \\ & - \frac{Z_3}{4}(\partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha)^2 + \frac{Z_3^{3/2} g_0}{2} f_{\alpha\beta\gamma}\left(\partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha\right) A_\mu^\beta A_\nu^\gamma \\ & - \frac{Z_3^2 g_0^2}{4}\left(f_{\alpha\beta\gamma} A_\mu^\beta A_\nu^\gamma\right)^2 - \frac{Z_3}{2\xi_0}(\partial \cdot A^\alpha)^2 \\ & + \tilde{Z}\partial_\mu \bar{\eta}_\alpha \partial^\mu \eta_\alpha + \tilde{Z} Z_3^{1/2} g_0 \partial^\mu \bar{\eta}_\gamma f_{\alpha\beta\gamma} A_\mu^\beta \eta_\alpha. \end{aligned} \tag{3.8}$$

Note that Z_2 could be a matrix relating bare and renormalized quark fields, diagonal in quark flavor, but color-independent.

Both of the formulae (3.5) and (3.8) define the same Lagrangian density; they differ only by a change of variables; the physical predictions are the same. Thus, provided that the correct LSZ prescription is used, the S-matrix and cross sections are unchanged under the field redefinitions.

The first form (3.5), with the bare fields, has unit coefficients for the terms $i\bar{\psi}_0\partial\!\!\!/\psi_0$, etc., which implies that the bare fields obey canonical (anti)commutation relations. This is a natural standard which then gives an invariant meaning to the normalization of the bare coupling and mass.

We have restricted the change of parameters to those that preserve gauge-invariance properties, admittedly with a renormalization of the definition of the gauge transformations. It is a theorem that this is sufficient to obtain finite Green functions. It can also be proved that ξ_0/Z_3 is finite, so that we can define $\xi_0 = \xi Z_3$ with ξ a finite renormalized gauge-fixing parameter; thus the gauge-fixing term in (3.8) has coefficient $1/\xi$. For proofs, see, for example, Collins (1984).

3.2.2 Renormalized BRST symmetry

The BRST transformations also need renormalization. This is done by a multiplicative renormalization of the parameter $\delta\lambda_0$:

$$\delta\lambda_0 = \delta\lambda Z_3^{1/2}\tilde{Z}^{1/2}. \tag{3.9}$$

In the resulting formulae (Collins, 1984, p. 297) for the renormalized BRST transformations of the renormalized fields, it is convenient to define

$$X = \tilde{Z} Z_3^{1/2} g_0/g_R, \tag{3.10}$$

where g_R is a finite parameter that is a version of the renormalized coupling to be introduced later. (The actual formula is $g_R = g\mu^\epsilon$; see (3.14).) The resulting renormalized BRST

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

transformations are finite:

$$\delta_{\mathrm{BRST,\,R}}\psi = -ig_R\eta_\alpha\delta\lambda t^\alpha\psi\ X, \tag{3.11a}$$

$$\delta_{\mathrm{BRST,\,R}}\bar{\psi} = ig_R\bar{\psi}t^\alpha\eta_\alpha\ X\,\delta\lambda, \tag{3.11b}$$

$$\delta_{\mathrm{BRST,\,R}}A^\alpha_\mu = \left(\partial_\mu\eta_\alpha\tilde{Z} + X\ g_R f_{\alpha\beta\gamma}\eta_\beta A^\gamma_\mu\right)\delta\lambda. \tag{3.11c}$$

The ghost and antighost fields transform as

$$\delta_{\mathrm{BRST,\,R}}\eta_\alpha = -\tfrac{1}{2}g_R X f_{\alpha\beta\gamma}\eta_\beta\eta_\gamma\delta\lambda, \tag{3.11d}$$

$$\delta_{\mathrm{BRST,\,R}}\bar{\eta}_\alpha = \frac{1}{\xi}\partial\cdot A_\alpha\delta\lambda. \tag{3.11e}$$

The finite operators on the right-hand sides of these equations are used in Slavnov-Taylor identities.

3.2.3 Counterterms, renormalized parameters, dimensional regularization

To implement renormalization in perturbation theory, we use a counterterm approach. The Lagrangian is split into three parts:

$$\mathcal{L} = \mathcal{L}_{\mathrm{free}} + \mathcal{L}_{\mathrm{b.i.}} + \mathcal{L}_{\mathrm{c.t.}}. \tag{3.12}$$

Free propagators correspond to the free Lagrangian $\mathcal{L}_{\mathrm{free}}$, which has the standard form in which appear derivative terms with unit coefficient, and mass terms with renormalized masses. The "basic interaction" Lagrangian $\mathcal{L}_{\mathrm{b.i.}}$ contains interaction terms, but with coefficients constructed using only finite renormalized couplings. Graphs constructed with only the basic interaction contain divergences in some of their one-particle-irreducible (1PI) subgraphs. The divergences are canceled by graphs in which divergent subgraphs are replaced by counterterm vertices derived from the counterterm Lagrangian $\mathcal{L}_{\mathrm{c.t.}}$. The rules for perturbation theory ensure that subdivergences in multiloop graphs are correctly canceled, order-by-order in an expansion in powers of the renormalized coupling.

Since the counterterms cancel the divergent contributions to loop graphs from UV momenta, it does not matter how UV divergences are regulated. After removal of the regulator, the same results are obtained for renormalized Green functions expressed in terms of renormalized parameters. The only requirement is a suitable adjustment of the finite parts of the counterterms when the method of UV regulation is changed.

For QCD perturbation theory, the most convenient UV regulator is often dimensional regularization, where the dimension n of space time is a continuous complex parameter, also written[2] as $n = 4 - 2\epsilon$. Although it is not known how to apply dimensional regularization to the exact theory, there are no problems in perturbation theory. A concrete mathematical

[2] *Warning*: Although this is the most common definition of ϵ, other definitions also appear in the literature, notably $\epsilon = n - 4$ and $\epsilon = 4 - n$.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

definition can be made (Wilson, 1973; Collins, 1984) by using an infinite dimensional space for momenta (and coordinates), and by using pathologies of infinite dimensional spaces to define integration so that it gives the scaling properties appropriate for a non-integer dimension.

3.2.4 Implementation in QCD

The free and basic interaction Lagrangians are defined to be

$$\mathcal{L}_{\text{free}} = \bar{\psi}(i\not{\partial} - m)\psi - \frac{1}{4}\left(\partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu\right)^2 - \frac{1}{2\xi}(\partial \cdot A_\alpha)^2 + \partial_\mu \bar{\eta}^\alpha \partial^\mu \eta^\alpha, \tag{3.13}$$

$$\begin{aligned}\mathcal{L}_{\text{b.i.}} = & -g\mu^\epsilon \bar{\psi} t^\alpha \not{A}^\alpha \psi + g\mu^\epsilon f_{\alpha\beta\gamma} A^{\beta\mu} A^{\gamma\nu} \partial_\mu A^\alpha_\nu - \frac{g^2\mu^{2\epsilon}}{4}\left(f_{\alpha\beta\gamma} A^\beta_\mu A^\gamma_\nu\right)^2 \\ & + g\mu^\epsilon f_{\alpha\beta\gamma} \partial^\mu \bar{\eta}^\gamma A^\beta_\mu \eta^\alpha.\end{aligned} \tag{3.14}$$

Here is introduced the well-known unit of mass μ for dimensional regularization, so that the renormalized coupling is $g\mu^\epsilon$, with g dimensionless for all ϵ. The Feynman rules that follow from these parts of $\mathcal{L}$ are listed in Fig. 3.1.

The counterterm Lagrangian is everything else in the full Lagrangian (3.8):

$$\mathcal{L}_{\text{c.t.}} = (Z_2 - 1)\bar{\psi} i\not{\partial}\psi - \left(g_0 Z_2 Z_3^{1/2} - g\mu^\epsilon\right)\bar{\psi}\not{A}^\alpha t^\alpha \psi + \ldots \tag{3.15}$$

In renormalized perturbation theory, the counterterm Lagrangian is treated as part of the interaction. We therefore have an extra set of vertices, the counterterm vertices, listed in Fig. 3.2. These are like those in the basic interaction, Fig. 3.1, but with modified coefficients, together with extra two-point vertices.

3.2.5 Mass-dependence and gauge-invariance relations for counterterms

In renormalization theory (e.g., Collins, 1984) the following is shown:

- The Ward identities that follow from gauge invariance imply that independent renormalization of the different interaction vertices is not needed; a single renormalization factor applied to g_0 is suitable. Thus gauge invariance is preserved.
- No counterterm proportional to the gauge-fixing term is needed. That is, $\xi_0 = Z_3\xi$, within the class of gauges we are using.
- With the exception of the mass parameters, the renormalization counterterms can be chosen to be independent of the quark masses.
- Renormalization of the bare coupling g_0 and the bare mass m_0 can be chosen to be independent of the gauge-fixing parameter ξ.
- The bare quark mass is linear in the renormalized mass: $m_{(0)f} = Z_m m_f + m_{00}$, with Z_m and m_{00} independent of mass. With dimensional regularization, we can set $m_{00} = 0$, so that $m_{(0)f} = Z_m m_f$.
- Z_2 and Z_m can be chosen to be independent of quark flavor. (But other choices of scheme can be useful in treating heavy quarks: Secs. 3.9 and 3.10.)

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

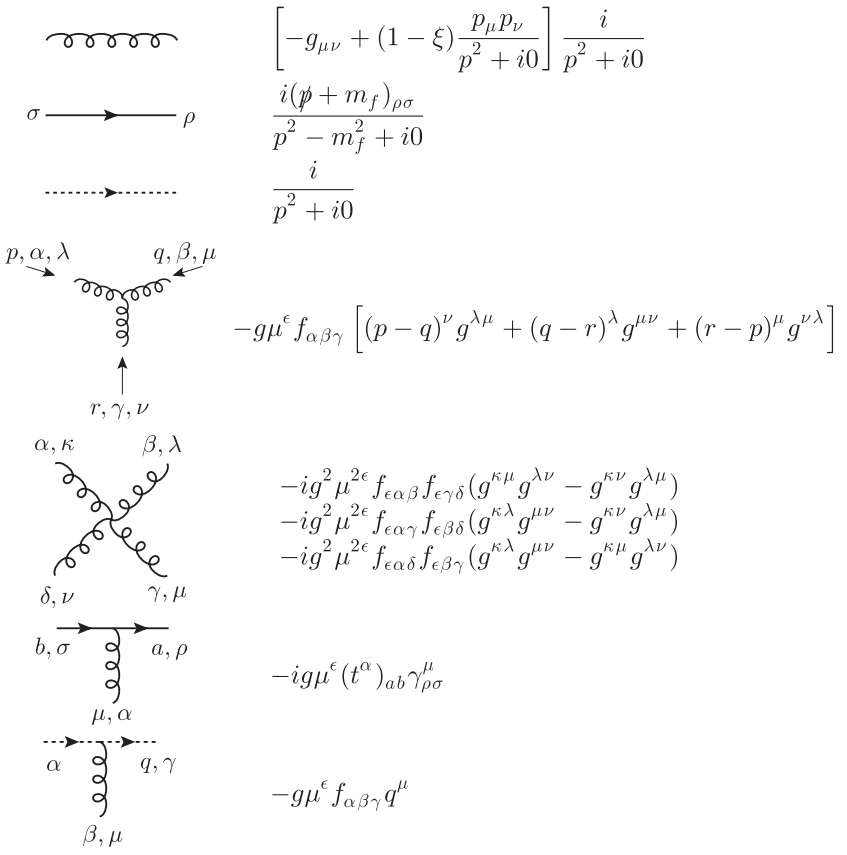


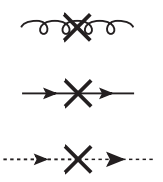
Fig. 3.1. Basic Feynman rules of QCD. The coupling has been replaced by $g\mu^\epsilon$, according to the standard convention for use in $4-2\epsilon$ dimensions. Propagators and vertices are diagonal in any indices (flavor or color) that are not explicitly indicated. For the renormalization counterterm vertices, see Fig. 3.2.

- Minimal subtraction (Sec. 3.2.6) is among the schemes to which the above statements on the lack of mass, flavor and gauge dependence apply.

3.2.6 Minimal subtraction

In a calculation order-by-order in the renormalized coupling, the requirement that a counterterm cancels its corresponding divergence determines the part of the counterterm that diverges as the UV regulator is removed, but not the finite part. A rule for determining the finite part is called a renormalization prescription. The most common in QCD calculations is minimal subtraction in its modified form, the $\overline{\rm MS}$ scheme due to Bardeen *et al*. (1978). When dealing with heavy quarks, it is convenient to apply a different scheme for graphs with heavy quark lines: Sec. 3.10.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press



$$i(Z_3-1)\left(-g_{\mu\nu}p^2+p_\mu p_\nu\right)$$

$$i\left[\not{p}(Z_2-1)-(Z_2Z_m-1)m_f\right]$$

$$i(\tilde{Z}-1)p^2$$

Vertex	Basic interaction	Counterterm
3-gluon	$g\mu^\epsilon$	$g_0 Z_3^{3/2} - g\mu^\epsilon$
4-gluon	$g^2\mu^{2\epsilon}$	$g_0^2 Z_3^2 - g^2\mu^{2\epsilon}$
Quark-gluon	$g\mu^\epsilon$	$g_0 Z_2 Z_3^{1/2} - g\mu^\epsilon$
Ghost-gluon	$g\mu^\epsilon$	$g_0 \tilde{Z} Z_3^{1/2} - g\mu^\epsilon$

Fig. 3.2. Counterterm vertices in QCD. The 2-point counterterms have diagonal dependence on all but Dirac indices for quarks and Lorentz indices for gluons. The other counterterm vertices simply correspond to vertices in Fig. 3.1 with the indicated modified coefficients for the coupling factors.

Definition

In the $\overline{\text{MS}}$ scheme, counterterms are pure poles at $\epsilon = 0$, except for unit-of-mass factors and a special factor S_ϵ for each loop:

$$g_0 = g\mu^\epsilon\left[1 + g^2 S_\epsilon \frac{B_{11}}{\epsilon} + g^4 S_\epsilon^2\left(\frac{B_{22}}{\epsilon^2} + \frac{B_{21}}{\epsilon}\right) + \dots\right], \tag{3.16}$$

$$Z_2 = 1 + g^2 S_\epsilon \frac{Z_{2,11}}{\epsilon} + g^4 S_\epsilon^2\left(\frac{Z_{2,22}}{\epsilon^2} + \frac{Z_{2,21}}{\epsilon}\right) + \dots, \tag{3.17}$$

etc. The rationale for the factor S_ϵ and its value are explained below. For normal UV divergences, the strength of the pole is at most $1/\epsilon^L$ in an L-loop counterterm. The only parameter on which the coefficients depend is the gauge-fixing parameter ξ, and this is absent for the bare coupling: the coefficients B_{ij} are pure numbers. In particular, the coefficients are independent of mass and of μ ('t Hooft, 1973; Collins, 1974).

The role played in renormalization by the unit of mass μ is quite central. It is commonly called the "renormalization mass" or "renormalization scale".

The $\overline{\text{MS}}$ scheme differs from the simplest minimal subtraction scheme by inserting a factor S_ϵ for each loop in the counterterms. This was motivated (Bardeen *et al.*, 1978) by the observation that in a one-loop calculation, there is an ϵ-dependent factor that naturally arises from an angular integration in $4-2\epsilon$ dimensions, and that would lead to certain universally occurring extra terms in renormalized Green functions. These are eliminated by choosing S_ϵ suitably. I define

$$S_\epsilon = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)} = 1 + \epsilon[\ln(4\pi) - \gamma_{\text{E}}] + O(\epsilon^2) \simeq 1 + 1.954\epsilon + O(\epsilon^2). \tag{3.18}$$

Here $\gamma_{\text{E}} = 0.5772\cdots$ is the Euler constant, and Γ is the gamma function.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

Although there are several ways in which the $\overline{\text{MS}}$ scheme has been defined in the literature, it can be proved (see problem 3.3) that all these definitions lead to identical renormalized Green functions at $\epsilon = 0$. For example, there are different formulae for S_ϵ, but only the order ϵ part of S_ϵ affects renormalized Green functions.[3] The equivalence of the definitions, to all orders of perturbation theory, applies to conventional Green functions, where the UV divergences give at most one pole per loop. But in Chs. 10 and 13, we will define quantities that have a double UV pole per loop. For these, it is the particular definition, (3.18), that gives the maximal simplification.

Advantages

Among the advantages of minimal subtraction is that it automatically preserves simple symmetries. For example, the counterterms for the 4-gluon interaction and for the 3-gluon interaction, etc., will automatically give counterterms with the correct gauge-invariance relations. Counterterms have their minimal mass dependence.

Mathematically, some care is needed in understanding the expansion about $g = \epsilon = 0$. Perturbative renormalization is done by first expanding in g and then analyzing the ϵ dependence. Real physics is defined with $\epsilon \to 0$ taken at fixed g. The direct perturbative calculation of the counterterms is really only valid in a region of g that shrinks to zero as $\epsilon \to 0$. This is enough to obtain the coefficients for renormalized perturbation theory, whose radius of applicability is not expected to shrink with ϵ.

As we will see in Sec. 3.5, renormalization-group methods can be used to calculate the true behavior of the bare parameters when the UV regulator (e.g., dimensional regularization, or a lattice) is removed with the renormalized couplings and mass fixed.

Renormalization group

A change of renormalization scheme, including a change of the unit of mass, can be compensated by a change in the numerical values of the renormalized parameters. All that changes is the parameterization of the set of renormalized theories by coupling(s) and masses. This is the subject of the renormalization group (RG) – Sec. 3.5 – which is a vital technique in perturbative QCD.

Minimal subtraction with other regulators

Although minimal subtraction is normally defined using dimensional regularization, the concept applies to any regularization method. With regularization by a lattice spacing a, one could define the counterterms in each order to be a polynomial in $\ln(a\mu)$ with no constant term. This would define a different scheme, related by a RG transformation to the standard $\overline{\text{MS}}$ scheme, which uses dimensional regularization.

3.3 Renormalization counterterms of QCD

Renormalization plays an essential role in perturbative QCD calculations. Not only does renormalization enable finite results to be obtained, but the counterterms themselves

[3] *Warning*: In comparing formulae for S_ϵ, note that some authors use a different convention for ϵ than this book.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

Fig. 3.3. (a) Quark self-energy graph. (b) Counterterm.

determine the renormalization-group coefficients that we will see are vital to predicting the scale dependence of measurable quantities. This is useful, since counterterms are much simpler to calculate than the finite parts of graphs.

This section reviews the renormalization of QCD at the one-loop level, giving a complete calculation for some parts and leaving the rest as an exercise. In Sec. 3.5, this will enable us to verify the key result of asymptotic freedom of QCD.

3.3.1 Wave-function renormalization

The wave-function and mass renormalization factors are obtained from propagator corrections, the "self-energy graphs". For the case of the quark, the one-loop graph and its counterterm are shown in Fig. 3.3. The graph's value is

$$\frac{g^2\mu^{2\epsilon}C_F}{(2\pi)^{4-2\epsilon}}\int \mathrm{d}^{4-2\epsilon}k\,\frac{\gamma^\mu(\not p-\not k+m)\gamma^\nu\left[-g_{\mu\nu}+(1-\xi)k_\mu k_\nu/(k^2+i0)\right]}{\left[(p-k)^2-m^2+i0\right]\left(k^2+i0\right)}, \tag{3.19}$$

where the C_F factor is from the color matrices $\sum_\alpha t_\alpha t_\alpha$, which gives $4/3$ in QCD. We combine the denominators using the Feynman parameter method (A.55), after which the momentum integral can be shifted so that the denominator loses its linear term in k. The use of standard Dirac algebra gives

$$\frac{g^2\mu^{2\epsilon}C_F}{(2\pi)^{4-2\epsilon}}\int_0^1 \mathrm{d}x\int \mathrm{d}^{4-2\epsilon}k\left\{\frac{(2-2\epsilon)\not p(1-x)-(4-2\epsilon)m+(1-\xi)(m-\not p x)}{\left[-k^2-p^2x(1-x)+m^2x-i0\right]^2}\right.$$
$$\left.-\frac{2(1-x)(1-\xi)(p^2\not p x^2+\not k\not p\not k)}{\left[-k^2-p^2x(1-x)+m^2x-i0\right]^3}+\text{terms odd in } k\right\}. \tag{3.20}$$

A Wick rotation gives a spherically symmetric integral in a Euclidean k variable in $4-2\epsilon$ dimensions, which can be performed analytically by using (A.34) and (A.50) and $\Gamma(1+\epsilon) = \epsilon\Gamma(\epsilon)$ to give

$$\frac{ig^2(4\pi\mu^2)^\epsilon C_F}{16\pi^2}\Gamma(\epsilon)\int_0^1 \mathrm{d}x\left[-p^2x(1-x)+m^2x-i0\right]^{-\epsilon}$$
$$\times\left\{(2-2\epsilon)\not p(1-x)-(4-2\epsilon)m+(1-\xi)\left[m-\not p x-\not p(1-x)(1-\epsilon)\right]\right.$$
$$\left.-\frac{\epsilon x^2(1-x)(1-\xi)p^2\not p}{\left[-p^2x(1-x)+m^2x-i0\right]}\right\}. \tag{3.21}$$

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

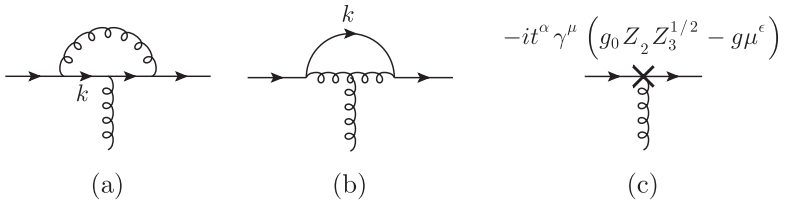


Fig. 3.4. (a) and (b) One-loop graphs for quark-gluon vertex. (c) Counterterm graph.

The pole at $\epsilon = 0$ is easy to extract, since $\Gamma(\epsilon) = 1/\epsilon + \text{finite}$, so that

$$\text{pole part of graph (a)} = \frac{ig^2 C_F}{16\pi^2\epsilon}\left[-3m + \xi(\not{p} - m)\right]. \tag{3.22}$$

We require the pole part plus the order g^2 part of the counterterm in Fig. 3.3(b) to be finite. This gives

$$Z_2 = 1 - \xi C_F \frac{\alpha_s S_\epsilon}{4\pi\epsilon} + O(\alpha_s^2), \tag{3.23}$$

$$Z_m = 1 - 3C_F \frac{\alpha_s S_\epsilon}{4\pi\epsilon} + O(\alpha_s^2). \tag{3.24}$$

Here, $\alpha_s = g^2/(4\pi)$, a commonly used definition analogous to the fine-structure constant in electromagnetism. The quantity S_ϵ is defined in (3.18), used to define the $\overline{\text{MS}}$ scheme.

Similar calculations for the gluon and ghost give

$$Z_3 = 1 - \frac{\alpha_s S_\epsilon}{4\pi\epsilon}\left[\left(\frac{\xi}{2} - \frac{13}{6}\right)C_A + \frac{4}{3}T_F n_f\right] + O(\alpha_s^2), \tag{3.25}$$

$$\tilde{Z} = 1 + \frac{\alpha_s S_\epsilon}{4\pi\epsilon} C_A \left(\frac{3}{4} - \frac{\xi}{4}\right) + O(\alpha_s^2). \tag{3.26}$$

In QCD, with its SU(3) gauge group and quarks in the triplet representation, the group theory coefficients used here are $C_A = 3$ and $T_F = 1/2$. See Sec. A.11 for more details. The quantity n_f is the number of quark flavors in QCD.

3.3.2 Quark-gluon vertex

To obtain g_0, we need to calculate one of the vertex functions. The simplest is the quark-gluon vertex, because it is only logarithmically divergent. The one-loop graphs and the counterterm are shown in Fig. 3.4. Now the UV divergence is independent of masses and external momenta. So we simplify the calculation by setting these variables to zero, and by ignoring any ϵ dependence that does not affect the pole. From the first graph we need

$$V_a = t_\alpha g\mu^\epsilon S_\epsilon \frac{g^2}{16\pi^4}(C_F - \tfrac{1}{2}C_A)\,\text{PP}\int_{\text{UV}} \frac{\mathrm{d}^{4-2\epsilon}k}{(k^2)^3}\left[-\gamma^\nu \not{k}\gamma^\mu \not{k}\gamma_\nu + \frac{1-\xi}{k^2}\not{k}\not{k}\gamma^\mu \not{k}\not{k}\right], \tag{3.27}$$

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

where "PP" means "pole part at $\epsilon = 0$". The subscript "UV" on the integration means that we restrict the integration to the UV region; we cut out a neighborhood of $k = 0$. The prefactors are those present in the lowest-order vertex. A Wick rotation and elementary spherically symmetric integrals over Euclidean k give the integral in terms of

$$\int_{\mathrm{UV}} \frac{\mathrm{d}^{4-2\epsilon}k}{(k^2)^2} = \frac{\pi^{2-\epsilon}}{\Gamma(2-\epsilon)} \int_{\mathrm{finite}}^{\infty} \frac{\mathrm{d}|k^2|}{|k^2|^{1+\epsilon}}, \tag{3.28}$$

so that

$$V_a = -it_\alpha \gamma^\mu g\mu^\epsilon S_\epsilon \frac{g^2}{16\pi^2\epsilon} \xi (C_F - C_A/2). \tag{3.29}$$

Similarly, graph (b) gives

$$\begin{aligned} V_b &= -t_\alpha g\mu^\epsilon S_\epsilon \frac{g^2}{16\pi^4} (\tfrac{1}{2} C_A) \\ &\quad \times \mathrm{PP} \int_{\mathrm{UV}} \mathrm{d}^{4-2\epsilon}k \, \frac{\gamma^{\kappa'} \not{k} \gamma^{\nu'}}{(k^2)^3} (-2k^\mu g^{\kappa\nu} + k^\nu g^{\kappa\mu} + k^\kappa g^{\nu\mu}) \\ &\quad \times \left(-g_{\nu'\nu} + (1-\xi)\frac{k_{\nu'}k_\nu}{k^2} \right) \left(-g_{\kappa'\kappa} + (1-\xi)\frac{k_{\kappa'}k_\kappa}{k^2} \right) \\ &= -it_\alpha \gamma^\mu g\mu^\epsilon S_\epsilon \frac{g^2}{16\pi^2\epsilon} C_A \left(\frac{3}{4} + \frac{3}{4}\xi \right). \end{aligned} \tag{3.30}$$

From these, we deduce the one-loop counterterm and, hence, from the previously calculated values of Z_2 and Z_3 we get the bare coupling:

$$g_0 = g\mu^\epsilon \left[1 - \frac{\alpha_s S_\epsilon}{4\pi\epsilon} \left(\frac{11}{6} C_A - \frac{2}{3} T_F n_f \right) + O(\alpha_s^2) \right]. \tag{3.31}$$

Note that the manipulations to obtain the coupling are performed with only the first two terms in a strict expansion in powers of g.

The results for the counterterms to higher order, up to four loops, can be deduced from the published values (Tarasov, Vladimirov, and Zharkov, 1980; Larin and Vermaseren, 1993; van Ritbergen, Vermaseren, and Larin, 1997; Czakon, 2005) of the RG coefficients, the primary ones being given in Sec. 3.7. See problem 3.2.

3.4 Meaning of unit of mass, renormalization scale

The unit of mass μ is a rather abstract concept seemingly tied to the use of dimensional regularization. It appears as a renormalization scale in renormalized quantities. We will see later (Sec. 3.5) that the value of the renormalization scale can be freely chosen, provided that the numerical value of the coupling and masses are adjusted in compensation. Perturbative calculations can be optimized in accuracy by a suitable choice of μ.

To understand how to choose μ, I now present a simple example that gives μ an intuitive meaning as approximating a cutoff in the physical dimension at a certain value of transverse momentum.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

Now, in many of our calculations for scattering, there will be preferred coordinates determined by the momenta of two of the particles. The Breit frame for DIS is a good example. Let us use these directions to fix a plane of t and z. Then for an integration over a momentum k, we first perform the k^0 and k^z integrals. After that we have a two (or $2-2\epsilon$) dimensional integral over a transverse momentum $\boldsymbol{k}_{\mathrm{T}}$, which is often rotationally symmetric. A generic one-loop integral, relative to a lowest-order calculation, is then

$$\begin{aligned} I_0 &= \frac{g^2\pi\mu^{2\epsilon}}{(4\pi)^{4-2\epsilon}} \int \mathrm{d}^{2-2\epsilon}\boldsymbol{k}_{\mathrm{T}}\, \frac{1}{\boldsymbol{k}_{\mathrm{T}}^2 + M^2} \\ &= \frac{g^2}{16\pi^2} \frac{\left(4\pi\mu^2\right)^{\epsilon}}{\Gamma(1-\epsilon)} \int_0^\infty \mathrm{d}k_{\mathrm{T}}^2\, \frac{(k_{\mathrm{T}}^2)^{-\epsilon}}{k_{\mathrm{T}}^2 + M^2}. \end{aligned} \tag{3.32}$$

The factor π in the first line is typical for a two-dimensional integral over the two longitudinal components of k. In an actual application, M would be a function of external longitudinal kinematic variables as well as of masses of particles and fields. For examples, see (9.4) and (10.137).

Using (A.50), we express the integral in terms of Γ functions, and then obtain the pole and finite part using (A.47):

$$I_0 = \frac{g^2}{16\pi^2}\Gamma(\epsilon)\left(\frac{4\pi\mu^2}{M^2}\right)^{\epsilon} = \frac{g^2 S_\epsilon}{16\pi^2}\left(\frac{1}{\epsilon} + \ln\frac{\mu^2}{M^2} + O(\epsilon)\right). \tag{3.33}$$

Subtraction of the $\overline{\mathrm{MS}}$ pole gives the renormalized value

$$I_R = \lim_{\epsilon\to 0}\left(I_0 - \frac{g^2 S_\epsilon}{16\pi^2\epsilon}\right) = \frac{g^2}{16\pi^2}\ln\frac{\mu^2}{M^2}. \tag{3.34}$$

Without the S_ϵ factor in the definition of the $\overline{\mathrm{MS}}$ counterterm, we would get an extra term containing $\ln(4\pi) - \gamma_{\mathrm{E}}$. The simple logarithmic dependence on the unit of mass μ is a general expectation, but for a more general integral the rest of the result will not be so simple and will not always have a simple analytic form.

To obtain an interpretation, we now rewrite the counterterm as a subtraction at the level of the integrand. Since the divergence is associated with the asymptotic large k_{T} behavior of the integrand, we consider an integral over this asymptotic behavior:

$$\frac{g^2\pi\mu^{2\epsilon}}{(4\pi)^{4-2\epsilon}} \int_{\boldsymbol{k}_{\mathrm{T}}^2 > C\mu^2} \frac{\mathrm{d}^{2-2\epsilon}\boldsymbol{k}_{\mathrm{T}}}{\boldsymbol{k}_{\mathrm{T}}^2} = \frac{g^2}{16\pi^2}\frac{\left(4\pi\mu^2\right)^{\epsilon}}{\Gamma(1-\epsilon)} \int_{C\mu^2}^\infty \frac{\mathrm{d}k_{\mathrm{T}}^2}{(k_{\mathrm{T}}^2)^{1+\epsilon}}. \tag{3.35}$$

The integral is of a power of k_{T}, so it is trivial to calculate. Since the extraction of the asymptotic behavior would otherwise expose an IR divergence, we put a lower limit proportional to μ^2 on the integration, with a coefficient C that is to be adjusted to obtain the correct finite part of the counterterm. The integral is

$$\frac{g^2 S_\epsilon}{16\pi^2\epsilon}\frac{e^{\gamma_{\mathrm{E}}\epsilon}}{\Gamma(1-\epsilon)} C^{-\epsilon}. \tag{3.36}$$

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

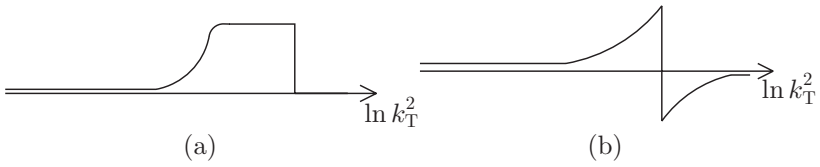


Fig. 3.5. (a) Integrand times k_T^2 of ((3.37) when $\mu \gg M$. (b) Same when μ is close to M.

Since the second factor is $1 + O(\epsilon^2)$, we can reproduce the renormalized graph by using (3.36) in place of the true $\overline{\text{MS}}$ counterterm, provided that we set $C = 1$. Then the renormalized graph is an integral in the physical dimension with a subtracted integrand:

$$I_R = \frac{g^2}{16\pi^3} \int \mathrm{d}^2\boldsymbol{k}_T \left(\frac{1}{\boldsymbol{k}_T^2 + M^2} - \frac{\theta(|\boldsymbol{k}_T| > \mu)}{\boldsymbol{k}_T^2} \right). \tag{3.37}$$

The integrand is plotted in Fig. 3.5. Because of the logarithmic behavior at large k_T, it is convenient to multiply the integrand by k_T^2 and to plot it against $\ln k_T^2$, to correspond to the integrand on the r.h.s. of an integral of the form

$$\int \mathrm{d}k_T^2 \, f(k_T) = \int \mathrm{d} \ln k_T^2 \, \left[k_T^2 f(k_T) \right]. \tag{3.38}$$

We now interpret (3.37), with a view to generalization.

- The natural expansion parameter for perturbation theory is $g^2/16\pi^2$, which arises as the product of the coupling, the factor $1/(2\pi)^4$ for a loop integral, and π^2 for an angular integration in four dimensions.
- This is multiplied by a group-theory factor and the number of graphs.
- In simple cases, renormalization can be performed by a subtraction of the asymptote of the integrand. The lower bound on k_T in the subtraction is commonly exactly μ.
- The coefficient S_ϵ defining the $\overline{\text{MS}}$ scheme is responsible for the cutoff being μ rather than a factor times μ. This gives a direct connection to the physical scale M in the integrand.
- In a more general graph, finite terms with modest, typically rational, numbers must be added. The need for this can be seen in the quark self-energy calculation, where the ϵ dependence of the numerator algebra enters.
- To get perturbative corrections of a natural size, μ should be close to the scale that is set by the transverse momentum dependence of the integrand, i.e., a scale characterizing the change from $1/k_T^2$ behavior at large k_T to constant behavior at small k_T.
- Although our example integral is exactly zero when $\mu = M$, this is not true in general; also M will generally be a function of external momentum. The best general statement is that for a single graph without a group-theory coefficient, the expected coefficient of $g^2/16\pi^2$ is a modest number of typical size unity if μ is close to a natural scale.
- For large values of μ, μ behaves like a cutoff on k_T in the unsubtracted integral.
- The rationale for these results suggests that they should approximately generalize to higher orders. In a well-behaved L-loop calculation, we can expect the result to be

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

roughly $(g^2/16\pi^2)^L$ times an effective number of graphs times a typical group-theory factor, provided again that μ is of the order of the physically relevant scale in transverse momentum.

- When we meet badly behaved situations, it is a good idea to search for explanations for large perturbative corrections in terms of the sizes of integrands in relevant kinematic regions.

3.5 Renormalization group

The general idea of renormalization prescribes only that counterterms cancel divergences; thus the finite parts of counterterms can be chosen freely. Within many schemes, like $\overline{\text{MS}}$, there is also a parameter μ that can be chosen freely. At first sight, the choices remove predictive power from the theory since any numerical value can be obtained from a one-loop integral with given external momenta. In reality, as explained more fully in textbook accounts of renormalization, this is not so. Instead we exploit the freedom in choosing μ to optimize the accuracy of finite-order perturbative calculations.

The complete theory is exactly invariant if when changing μ (or, more generally, the renormalization prescription) we also change the numerical values of the renormalized parameters of the theory. This is the renormalization-group (RG) invariance of the theory. An RG transformation amounts to a change in the partitioning of the full Lagrangian $\mathcal{L}$ into the three terms in (3.12). Thus it corresponds to a rearrangement of the perturbation expansion. The most important case for us is the transformation of the renormalized coupling and masses when the renormalization mass μ is changed.

3.5.1 RG evolution

When we perform RG transformations for changes of μ, keeping observable quantities fixed, each numerical value of μ corresponds to particular numerical values of the renormalized parameters $g(\mu)$, $m_f(\mu)$. When we change μ to another value μ', not only do the coupling and masses change, but also the normalization of the renormalized fields. So we write

$$\phi_i(x;\mu) = \zeta_i(\mu,\mu')\phi_i(x;\mu'). \tag{3.39}$$

Here i labels the different types of field (gluon, quark, etc.). A Green function therefore transforms as

$$G(\underline{p};\mu,g(\mu),\underline{m}(\mu)) = \prod_e \zeta_{i_e}(\mu,\mu')G(\underline{p};\mu',g(\mu'),\underline{m}(\mu')), \tag{3.40}$$

where $\underline{p}$ is the collection of external momenta of G, $\underline{m}$ is the set of renormalized masses, and the product is over the external lines, e, of G, with i_e labeling the corresponding types of field.

The S-matrix and hence cross sections are RG invariant. This is because an S-matrix element is obtained by applying to the corresponding off-shell Green function the following operations: (a) divide out a full external propagator; (b) multiply by the square root of the

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

residue of the particle pole; (c) put the external momenta on-shell. In this process there is a cancellation of the ζ factors for each external line. Exactly the same idea applies to Green functions of the composite external fields needed to obtain the S-matrix for composite particles.

We now determine equations for μ dependence of $g(\mu)$. The coefficients in the equations are obtained from the counterterms in the bare parameters, the starting point being RG invariance of the bare parameters, as is necessary to keep the physics unchanged. The normalizations of the bare parameters and the bare fields are fixed because terms like $i\bar{\psi}_0 \partial\!\!\!/ \psi_0$ have unit coefficients. Our discussion is tailored to the $\overline{\text{MS}}$ scheme, but the main principles and methods are general.

3.5.2 Coupling and mass

With a UV cutoff applied ($\epsilon \neq 0$), we hold the bare parameters g_0 and $m_{(0)f}$ fixed and vary μ. For g_0, we get

$$0 = \frac{\mathrm{d}}{\mathrm{d}\ln\mu} g_0(\mu, g(\mu), \epsilon) = \frac{\partial g_0}{\partial \ln\mu} + \frac{\mathrm{d}g}{\mathrm{d}\ln\mu}\frac{\partial g_0}{\partial g} = \epsilon g_0 + \frac{\mathrm{d}g}{\mathrm{d}\ln\mu}\frac{\partial g_0}{\partial g}. \tag{3.41}$$

We distinguish between a total derivative $\mathrm{d}/\mathrm{d}\mu$, with respect to *all* the μ dependence, and a partial derivative $\partial/\partial\mu$, for which the renormalized parameters $g(\mu)$, etc., are fixed. It is convenient to use a logarithmic derivative, given that renormalized graphs have a μ dependence that is polynomial in $\ln\mu$.

For the masses

$$0 = \frac{\mathrm{d}}{\mathrm{d}\ln\mu} m_0(m(\mu), g(\mu), \epsilon) = \frac{\mathrm{d}g}{\mathrm{d}\ln\mu}\frac{\partial m_0}{\partial g} + \frac{\mathrm{d}m}{\mathrm{d}\ln\mu}\frac{m_0}{m}, \tag{3.42}$$

where we used the lack of explicit μ dependence of m_0 in minimal subtraction.

It is convenient to use as the expansion parameter $\alpha_s/4\pi = g^2/16\pi^2$. Then from (3.41) we find

$$\frac{\mathrm{d}\alpha_s/4\pi}{\mathrm{d}\ln\mu} = \frac{g}{8\pi^2}\frac{\mathrm{d}g}{\mathrm{d}\ln\mu} = -\frac{\epsilon g}{8\pi^2}\frac{g_0}{\partial g_0/\partial g}. \qquad \text{(General } \epsilon\text{)} \tag{3.43}$$

The left-hand side is finite at $\epsilon = 0$, and therefore the right-hand side is finite also; all poles in ϵ must cancel. In the $\overline{\text{MS}}$ scheme each α_s in the counterterms is accompanied by a factor S_ϵ and all the terms in g_0 have negative powers of ϵ. We therefore find that the right-hand side has the form[4]

$$-2\epsilon\frac{\alpha_s}{4\pi} + S_\epsilon^{-1} 2\beta(\alpha_s S_\epsilon/(4\pi)), \tag{3.44}$$

where the only ϵ dependence is in the $-\epsilon\alpha_s$ term and in the explicit factors of S_ϵ.

At the physical space-time dimension, i.e., at $\epsilon = 0$, we use the perturbatively calculable β function to give an equation for the scale dependence of the coupling:

$$\frac{\mathrm{d}\alpha_s/4\pi}{\mathrm{d}\ln\mu} = 2\beta(\alpha_s/4\pi). \qquad (\epsilon = 0) \tag{3.45}$$

[4] The factor of 2 multiplying β is to correspond to the definition in Larin and Vermaseren (1993); this arises because these authors use derivatives with respect to $\ln\mu^2$ instead of $\ln\mu$.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

The one-loop value of the bare coupling, in (3.31), immediately gives

$$\begin{aligned}\beta(\alpha_s/4\pi) &= -\left(\frac{11}{3}C_A - \frac{4}{3}T_F n_f\right)\frac{\alpha_s^2}{16\pi^2} + O(\alpha_s^3) \quad \text{(for general group)} \\ &= -\left(11 - \frac{2}{3}n_f\right)\frac{\alpha_s^2}{16\pi^2} + O(\alpha_s^3). \qquad \text{(for SU(3))} \end{aligned} \tag{3.46}$$

Provided there are at most 16 quark flavors, which is true in currently known strong interactions, the coupling decreases with increasing scale, at least when it is small enough at the outset. The coupling does in fact go to zero as $\mu \to \infty$, as we will see, so that QCD is asymptotically free. The importance of this is clear from the previous chapter.

The results at higher order will be quoted in Sec. 3.7. Here we just note that β can be obtained from the *single* pole terms in g_0. With the conventions of (3.16), we get:

$$\beta(\alpha_s/4\pi) = \sum_{n=1}^{\infty} \frac{g^{2n+2}}{8\pi^2} n B_{n1}. \tag{3.47}$$

The finiteness conditions for $\mathrm{d}\alpha_s/\mathrm{d}\ln\mu$ enable the higher poles in g_0 to be computed in terms of the single poles.

The RG dependence of the mass is similarly obtained. A dimensionless function is obtained by using logarithmic derivatives:

$$\begin{aligned}\gamma_m(\alpha_s S_\epsilon/4\pi) &\stackrel{\text{def}}{=} \frac{\mathrm{d}\ln m}{\mathrm{d}\ln\mu} \\ &= \left(2\epsilon\alpha_s S_\epsilon/4\pi - 2\beta(\alpha_s S_\epsilon/4\pi)\right)\frac{\partial \ln Z_m}{\partial \alpha_s S_\epsilon/4\pi} \\ &= -6C_F\frac{\alpha_s}{4\pi}S_\epsilon + O(\alpha_s^2). \end{aligned} \tag{3.48}$$

Again, the divergences present in Z_m must cancel in this derivative in order that γ_m is finite. This time, it can be shown that the only ϵ dependence is in the S_ϵ multiplying α_s. This RG coefficient is usually less important in practice, since most pQCD calculations are performed with masses set to zero, or with a different scheme for heavy quarks.

The lack of mass dependence in the renormalization group coefficients β and γ_m follows from the mass-independence property of $\overline{\text{MS}}$ counterterms.

3.5.3 Anomalous dimensions and RG equations for Green functions

To unify the treatment of the RG transformation for renormalized fields, let us use the notation ϕ_i for the renormalized fields, with the label i denoting the type of field (gluon, flavor of quark, etc.). We define its anomalous dimension by

$$\gamma_i(\alpha_s S_\epsilon/4\pi, \xi)\phi_i = -\frac{\mathrm{d}\phi_i}{\mathrm{d}\ln\mu}. \tag{3.49}$$

Given that the corresponding bare field is $\phi_{(0)i} = Z_i^{1/2}\phi_i$, it follows that

$$\gamma_i(\alpha_s S_\epsilon/4\pi, \xi) = \frac{1}{2}\frac{\mathrm{d}\ln Z_i}{\mathrm{d}\ln\mu}. \tag{3.50}$$

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

A complication arises in gauge theories, from the gauge dependence of wave-function renormalization. Because of the relation $\xi_0 = \xi Z_3$, the gauge-fixing parameter obeys

$$\frac{\mathrm{d}\ln\xi}{\mathrm{d}\ln\mu} = -\frac{\mathrm{d}\ln Z_3}{\mathrm{d}\ln\mu} = -2\gamma_3. \tag{3.51}$$

Then the definition of γ_3 gives

$$\gamma_3 = \left(-\epsilon\frac{\alpha_s S_\epsilon}{4\pi} + \beta(\alpha_s S_\epsilon/4\pi)\right)\frac{\mathrm{d}\ln Z_3}{\mathrm{d}\alpha_s S_\epsilon/4\pi} - \gamma_3\frac{\partial\ln Z_3}{\partial\xi}. \tag{3.52}$$

Hence

$$\gamma_3 = \frac{\left(-\epsilon\frac{\alpha_s}{4\pi}S_\epsilon + \beta(\alpha_s S_\epsilon/4\pi)\right)\mathrm{d}\ln Z_3/\mathrm{d}(\alpha_s S_\epsilon/4\pi)}{1 + \partial\ln Z_3/\partial\xi}. \tag{3.53}$$

For the other anomalous dimensions, we have equations of the form

$$\gamma_2 = \left(-\epsilon\frac{\alpha_s S_\epsilon}{4\pi} + \beta(\alpha_s S_\epsilon/4\pi)\right)\frac{\mathrm{d}\ln Z_2}{\mathrm{d}\alpha_s S_\epsilon/4\pi} - \gamma_3\frac{\partial\ln Z_2}{\partial\xi}. \tag{3.54}$$

See Sec. 3.7 for the values of the anomalous dimensions.

From the above results follows the renormalization-group equation (RGE) for a renormalized Green function G. If G has n_2 external quark fields (and the same number of antiquarks) and n_3 external gluons, then

$$\frac{\mathrm{d}G}{\mathrm{d}\ln\mu} = -(2n_2\gamma_2 + n_3\gamma_3)\,G. \tag{3.55}$$

Exactly similar equations can be derived for other operator matrix elements, where the states can be other than the vacuum and the fields not simple products of the elementary fields of QCD at different space-time points. A simple example is the hadronic tensor $W^{\mu\nu}$ of DIS, (2.18). The electromagnetic current is a symmetry current of QCD and can be shown to have zero anomalous dimension. Hence $W^{\mu\nu}$ is RG invariant:

$$\frac{\mathrm{d}W^{\mu\nu}}{\mathrm{d}\ln\mu} = 0. \tag{3.56}$$

3.6 Solution of RG equations

3.6.1 General form of solution

The RG equations for the coupling, mass, and Green functions are readily solved to relate these quantities at different values of the $\overline{\mathrm{MS}}$, with the aid of integrals of β, γ_m and the anomalous dimensions:

$$\ln\frac{\mu}{\mu_0} = \int_{\alpha_s(\mu_0)/4\pi}^{\alpha_s(\mu)/4\pi}\frac{\mathrm{d}\alpha/4\pi}{2\beta(\alpha/4\pi)}, \tag{3.57}$$

$$\ln\frac{m(\mu)}{m(\mu_0)} = \int_{\mu_0}^{\mu}\frac{\mathrm{d}\mu'}{\mu'}\gamma_m(\alpha_s(\mu')/4\pi) = \int_{\alpha_s(\mu_0)/4\pi}^{\alpha_s(\mu)/4\pi}\mathrm{d}\alpha/4\pi\,\frac{\gamma_m(\alpha/4\pi)}{2\beta(\alpha/4\pi)}, \tag{3.58}$$

$$\ln\frac{G(\mu)}{G(\mu_0)} = -\int_{\mu_0}^{\mu}\frac{\mathrm{d}\mu'}{\mu'}\gamma_G(\alpha_s(\mu')/4\pi,\xi(\mu')) = -\int_{\alpha_s(\mu_0)/4\pi}^{\alpha_s(\mu)/4\pi}\mathrm{d}\alpha/4\pi\,\frac{\gamma_G(\alpha/4\pi,\xi(\mu'))}{2\beta(\alpha/4\pi)}. \tag{3.59}$$

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

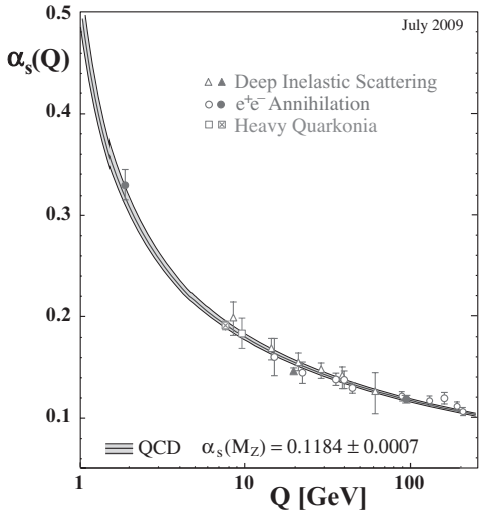


Fig. 3.6. QCD effective coupling. With kind permission from Springer Science+Business Media: Bethke (2009, Fig. 6). The lines represent the solution of the RGE for $\alpha_s(\mu)$ with the $\pm 1\sigma$ limits on the constant of integration. The scheme used is $\overline{\text{MS}}$ with a variable number of active quarks, as in Sec. 3.10. The data are, in increasing order of μ, from fits to the τ width, Υ decays, DIS, e^+e^- event shapes at 22 GeV at JADE, shapes at TRISTAN at 58 GeV, Z width, and e^+e^- event shapes at 91–208 GeV.

Here $\gamma_G = 2n_2\gamma_2 + n_3\gamma_3$ is the anomalous dimension of the Green function G, all of whose momentum and mass arguments we have suppressed.[5]

Since $\beta(\alpha/4\pi)$ is negative and $O(\alpha^2)$ at small coupling, (3.57) shows that $\alpha_s(\mu) \to 0$ as $\mu \to \infty$, i.e., that QCD is asymptotically free.

3.6.2 *Effective coupling; scale parameter* Λ

The μ dependence of the coupling underlies all other RG calculations in QCD, so a detailed analysis is useful. There is a one-parameter family of solutions of (3.45) for $\alpha_s(\mu)$, and the physical solution is specified, for example, by the value of coupling at a given scale (e.g., "$\alpha_s(M_Z) = 0.1184 \pm 0.0007$ in the $\overline{\text{MS}}$ scheme with five active flavors"). The physical solution is obtained by fitting the one parameter to data, with a result shown in Fig. 3.6.

One often-used procedure is the following, which is particularly useful for assessing the errors due to the limited accuracy with which RG functions are known. It was obtained (Buras *et al.*, 1977) basically by expanding $\alpha_s(\mu)$ in powers of $1/\ln\mu^2$ at large μ.

Let us write the expansion of β as

$$\beta(a_s) \stackrel{\text{def}}{=} \frac{\mathrm{d}a_s}{\mathrm{d}\ln\mu^2} = -\beta_0 a_s^2 - \beta_1 a_s^3 - \beta_2 a_s^4 - \beta_3 a_s^5 + O(a_s^6), \tag{3.60}$$

where $a_s = \alpha_s/4\pi = g^2/16\pi^2$. (The normalizations of all but β_0 differ from the less systematic conventions of the PDG; Amsler *et al.*, 2008.) In the solution (3.57) the integral

[5] Thus $G(\mu_0)$ means $G\big(\underline{p}; \mu_0, g(\mu_0), \underline{m}(\mu_0), \xi(\mu_0)\big)$.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

is of $1/\beta$, so we first separate out the singular parts of $1/\beta$, and represent the general solution of the RGE for $a_s(\mu)$ as

$$\ln \frac{\mu^2}{\Lambda^2} = \frac{1}{\beta_0 a_s} + \frac{\beta_1}{\beta_0^2} \ln(\beta_0 a_s) - f(a_s), \tag{3.61}$$

where

$$f(a_s) \stackrel{\text{def}}{=} \int_0^{a_s} \mathrm{d}a \left(-\frac{1}{\beta(a)} - \frac{1}{\beta_0 a^2} + \frac{\beta_1}{\beta_0^2 a} \right). \tag{3.62}$$

Here the constant of integration is represented by a parameter Λ, of the dimension of mass; it has an experimentally determined value[6] of around 200 MeV. The constant β_0 in the logarithm in the second term on the r.h.s. of (3.61) merely amounts to a standard convention for the definition of Λ whose rationale will become apparent below. When it is necessary to distinguish Λ from other similar parameters, we will add a subscript, as in Λ_{QCD}.

For small coupling, β is approximately $-\beta_0 a_s^2$, so that $a_s(\mu)$ behaves like $1/(\beta_0 \ln \mu^2)$ at large μ. To improve this estimate, we expand in powers of $1/\ln(\mu^2/\Lambda^2)$ (with some modifications as required). This gives

$$\frac{\alpha_s}{4\pi} = \frac{1}{\beta_0 \ln(\mu^2/\Lambda^2)} - \frac{\beta_1 \ln\ln(\mu^2/\Lambda^2)}{\beta_0^3 \ln^2(\mu^2/\Lambda^2)} + O\left(\frac{\ln^2 \ln(\mu^2/\Lambda^2)}{\ln^3(\mu^2/\Lambda^2)} \right). \tag{3.63}$$

Normally we would expect a term constant$/\ln^2(\mu^2/\Lambda^2)$, and the absence of this term is effectively the definition of Λ, and is exactly correlated with the use of $\ln(\beta_0 a_s)$ rather than $\ln a_s$ in (3.61). This convention is due to Buras *et al.* (1977). Then Λ can, in principle, be extracted from the large μ asymptote of $a_s(\mu)$:

$$\Lambda^2 = \lim_{\mu \to \infty} \mu^2 \exp\left[-\frac{1}{\beta_0 a_s} - \frac{\beta_1}{\beta_0^2} \ln(\beta_0 a_s) \right]. \tag{3.64}$$

Notice that this formula requires only the use of the known one- and two-loop terms in β, not any of the higher terms not all of which are known. Of course the higher terms will improve the accuracy of the measurement of Λ since $a_s(\mu)$ is only known at finite μ.

3.6.3 Dimensional transmutation

Suppose we were to approximate quark masses of QCD by zero. Since the masses of the light quarks are considerably smaller than the proton mass, this is in fact a useful approximation, for low-energy processes, if we keep only two (u, d) or three flavors (u, d, s), with the heavier quarks being removed according to the decoupling theorem. Then the mass of any particle, like the proton, would be a function of α_s and μ only. But by dimensional analysis it is μ times a function of α_s:

$$m_p = \mu F(\alpha_s) \qquad \text{in massless QCD.} \tag{3.65}$$

[6] Details depend on a treatment of heavy quark masses which we will present later (Sec. 3.10). The current best value with five active quarks is (Bethke, 2009) $\Lambda = (213 \pm 9)$ MeV.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

Since m_p is a physical mass, it is RG invariant, which fixes $\alpha_s(\mu)$ up to a multiplicative factor. It follows that m_p equals Λ times a pure number, K_p, which is a property of the solution of massless QCD: $m_p = \Lambda K_p$. The number K_p is non-perturbative and can be obtained from lattice QCD calculations.

Instead of specifying the theory by the numerical value of its dimensionless coupling g, we can instead specify a fixed mass parameter Λ. This is the property of dimensional transmutation (Coleman and Weinberg, 1973).

In fact, there is a certain sense in which even this parameter is illusory. Suppose we consider pure strong interactions with massless quarks. To completely define a measurement of the numerical value of Λ, we must specify a system of units, i.e., specify what a mass of numerical value unity means.[7] But with only the strong interaction under consideration, this can only mean some physical mass like the proton mass, which can be taken as a physical definition of a standard mass. So a measurement of Λ is really a measurement of the dimensionless ratio Λ/m_P, whose value is a unique prediction of the theory.

This is the sense in which *massless* QCD has no parameters. All real predictions of the theory are pure numbers. For example, a cross section as a function of center-of-mass energy $\sigma(E)$ is of the form $m_p^{-2} S(E/m_p)$, where S is a dimensionless function of a dimensionless variable. This function is in principle predicted with no parameters by massless QCD.

Since the masses of the three light quarks are known to give only a relatively small contribution to the nucleon mass, the above statements are approximately true in real QCD. The real intrinsic parameters of QCD are the quark masses, expressed in terms of a suitable chosen unit, e.g., Λ or m_p.

There is a contrast with QED, because of the different physics of its classical long-distance limit. For simplicity consider QED of a photon and electron field only. Then, again by dimensional transmutation, there is only one true parameter m_e/Λ_{QED}. As with its QCD analog, Λ_{QED} is in a region where the coupling is strong. In contrast to the QCD coupling, the QED coupling increases at large scales, and in fact Λ_{QED} is around the Planck scale. At low energies compared to m_e, the electron decouples, giving a free Maxwell field theory which we can solve completely and exactly. It therefore becomes much more sensible than in QCD to use an on-shell renormalization prescription, and to define the expansion parameter of the theory as the usual $\alpha \simeq 1/137$. Within pure single-lepton QED, we can take the unit of mass to be m_e.

Of course, weak coupling methods are very useful and accurate for normal phenomena in QED, including its bound states, in contrast to QCD, where perturbation theory has a more restricted range of applicability.

Although dimensional transmutation has reduced the number of genuine parameters in a quantum field theory by one compared with the apparent number, the parameter is regained when the theory is treated as a component of a more complete theory. For example, we can combine QED and QCD to get a complete theory underlying all nuclear, atomic and

[7] The last sentence was carefully worded to avoid confusion between the concept of unit of mass in dimensional regularization and the concept of the unit of mass in a system of units.

molecular phenomena. Then $m_e/\Lambda_{\rm QCD}$ is a parameter of the combined theory in addition to the intrinsic parameters of the separate theories.

3.6.4 Bare coupling

We used (3.41) to obtain the (finite) β function from the divergent perturbation expansion for the bare coupling. But we can also use it to obtain a formula for the bare coupling as a function of Λ and ϵ. From the ϵ-dependent β function given in (3.44) we get

$$\frac{\partial \ln a_0}{\partial a_s} = \frac{\epsilon}{\epsilon a_s - S_\epsilon^{-1}\beta(a_s S_\epsilon)}, \tag{3.66}$$

where again $a_s = g^2/(16\pi^2)$, while $a_0 = g_0^2/(16\pi^2)$ is the bare equivalent of $a_s\mu^{2\epsilon}$, with mass dimension 2ϵ. The solution is

$$\ln a_0 = \ln(a_s\mu^{2\epsilon}) + \int_0^{a_s S_\epsilon} {\rm d}a \left[\frac{\epsilon}{\epsilon a - \beta(a)} - \frac{1}{a}\right], \tag{3.67}$$

where the boundary condition is set by requiring $a_0/(a_s\mu^{2\epsilon}) \to 1$ as $a_s \to 0$ at fixed ϵ.

An important formula is obtained by expressing this in terms of Λ, and then taking the limit $\epsilon \to 0$ at fixed a_s. This gives

$$\frac{g_0^2}{16\pi^2} = \frac{1}{\beta_0}\epsilon^{1+\epsilon\beta_1/\beta_0^2}\left(\frac{\Lambda^2 e^{\beta_1/\beta_0^2+\gamma_{\rm E}}}{4\pi}\right)^{\epsilon}\left[1 + O(\epsilon^2)\right]. \tag{3.68}$$

When $\epsilon \to 0$, the $O(\epsilon^2)$ fractional correction can be dropped, since it is equivalent to a change in Λ by a fraction of order $O(\epsilon)$: since Λ determines the coupling in the renormalized theory, the correction does not affect renormalized Green functions at $\epsilon = 0$. From the β function, only the scheme-independent coefficients β_0 and β_1 are needed; the scheme choice is manifested in the numerical coefficient multiplying Λ^2. To provide a *full* specification of the renormalization of the theory, only the one- and two-loop renormalization counterterms in the coupling need to be known.

Similar results can be obtained when any regularization scheme is used. With a lattice regulator, we would have

$$\frac{g_0^2}{16\pi^2} = \frac{1}{-\beta_0 \ln(a^2\Lambda^2)} - \frac{\beta_1 \ln(-\ln(a^2\Lambda^2))}{\beta_0^3 \ln^2(a^2\Lambda^2)} - \frac{A}{\ln^2(a^2\Lambda^2)}, \tag{3.69}$$

where a is the lattice spacing, and the coefficient A can be computed (Hasenfratz and Hasenfratz, 1980; Dashen and Gross, 1981) from the perturbative expansion of the bare coupling computed to two-loop order, with the renormalized coupling being in the $\overline{\rm MS}$ scheme.

Other bare parameters and renormalization factors may be treated similarly.

3.7 Values of RG coefficients

The β function has been calculated in the $\overline{\rm MS}$ scheme up to three loops by Tarasov, Vladimirov, and Zharkov (1980) and by Larin and Vermaseren (1993), and to four loops by van Ritbergen, Vermaseren, and Larin (1997). The results have been confirmed by Czakon

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

(2005). The first three coefficients in β are rational numbers. With the notation of (3.60),

$$\beta_0 = \frac{11}{3}C_A - \frac{4}{3}T_F n_f, \tag{3.70a}$$

$$\beta_1 = \frac{34}{3}C_A^2 - 4C_F T_F n_f - \frac{20}{3}C_A T_F n_f, \tag{3.70b}$$

$$\beta_2 = \frac{2857}{54}C_A^3 + 2C_F^2 T_F n_f - \frac{205}{9}C_F C_A T_F n_f - \frac{1415}{27}C_A^2 T_F n_f + \frac{44}{9}C_F T_F^2 n_f^2 + \frac{158}{27}C_A T_F^2 n_f^2. \tag{3.70c}$$

The expression for the four-loop coefficient β_3 is more complicated and includes the irrational number ζ_3; the full expression is given in van Ritbergen, Vermaseren, and Larin (1997). The fact that even the three-loop coefficient is a rational number indicates a fundamental simplicity in the theory and in minimal subtraction that is certainly not apparent in straightforward calculations of Feynman diagrams. In the case of SU(3), i.e., for QCD, the coefficients are

$$\beta_0 = 11 - \frac{2}{3}n_f, \tag{3.71a}$$

$$\beta_1 = 102 - \frac{38}{3}n_f, \tag{3.71b}$$

$$\beta_2 = \frac{2857}{2} - \frac{5033}{18}n_f + \frac{325}{54}n_f^2, \tag{3.71c}$$

$$\beta_3 = \left(\frac{149\,753}{6} + 3564\zeta_3\right) - \left(\frac{1\,078\,361}{162} + \frac{6508}{27}\zeta_3\right)n_f + \left(\frac{50\,065}{162} + \frac{6472}{81}\zeta_3\right)n_f^2 + \frac{1093}{729}n_f^3 \approx 29\,243.0 - 6946.30 n_f + 405.089 n_f^2 + 1.499\,31 n_f^3. \tag{3.71d}$$

The anomalous dimensions have been computed by Larin and Vermaseren (1993) up to three loops, and by Czakon (2005) to four loops. The full results can be found in these papers.[8] Up to two-loop order, where the coefficients are rational, the values are

$$\gamma_2(\alpha_s/4\pi, \xi) = \frac{\alpha_s}{4\pi}C_F\xi + \left(\frac{\alpha_s}{4\pi}\right)^2\left(-\frac{3}{2}C_F^2 + \frac{25}{4}C_F C_A - C_F n_f + 2\xi C_F C_A + \xi^2\frac{1}{4}C_F C_A\right) + \ldots, \tag{3.72}$$

$$\gamma_3(\alpha_s/4\pi, \xi) = \frac{\alpha_s}{4\pi}\left[-\frac{13}{6}C_A + \frac{2}{3}n_f + \xi\frac{1}{2}C_A\right] + \left(\frac{\alpha_s}{4\pi}\right)^2\left[-\frac{59}{8}C_A^2 + 2C_F n_f + \frac{5}{2}C_A n_f + \xi\frac{11}{8}C_A^2 + \xi^2\frac{1}{4}C_A^2\right] + \ldots \tag{3.73}$$

In these equations, the value $T_F = 1/2$ was used.

[8] The definition of γ has different normalization conventions in different books and papers. The conventions of this book agree with those of Larin and Vermaseren (1993).

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

3.8 Symmetries and approximate symmetries of QCD

In this section, I summarize the standard set of exact and approximate symmetries of QCD. See Narison (2002, Chs. 53 and 54) for a recent account of many of their consequences, especially those that are not further referenced in this section.

3.8.1 Exact symmetries

The QCD Lagrangian is exactly invariant when any one of the quark fields is multiplied by a phase. By Noether's theorem this gives rise to conservation of the number of quarks (minus antiquarks) of each flavor: u-quark number, d-quark number, etc. The sum of all of these, the total quark number, is particularly important because it is not broken by flavor-changing weak interactions. Baryon number is simply one-third of total quark number, and its invariance was established long before QCD.

QCD is also invariant under each of the discrete symmetries of parity, charge conjugation, and time-reversal.

3.8.2 Note on "strong CP problem"

If QCD is specified simply as a renormalizable gauge theory, with an SU(3) gauge group and some set of quark fields in the triplet representation, then one extra term is permitted beyond those in the Lagrangian (2.1). In a standard normalization, the extra term has the form

$$\frac{\theta}{16\pi^2} G^{\alpha}_{\mu\nu} \tilde{G}^{\alpha\,\mu\nu}, \tag{3.74}$$

where $\tilde{G}^{\alpha}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} G^{\alpha\,\rho\sigma}$. The extra term breaks CP invariance, and there is a stringent observational bound on its coupling, $\theta \ll 10^{-9}$. It is considered problematic as to why θ is so small. This is the strong CP problem, which is reviewed along with possible solutions in Dine (2000).

3.8.3 Isospin and flavor SU(3)

If the up and down quarks were exactly equal in mass, QCD would be invariant under the isospin symmetry of SU(2) transformations on the u- and d-quark fields. This symmetry is quite accurate; we will apply it to the flavor dependence of parton densities and fragmentation functions in Secs. 6.9.7 and 12.4.8.

Rather less accurate is the flavor SU(3) symmetry that would be exact if the masses of the lightest three quarks, u, d, and s, were equal. SU(3) breaking is described by the quark mass terms, which correspond to the Q_3 and Q_8 terms of an operator transforming as an octet under flavor SU(3). Treated to first order in perturbation theory, these give a good description of the mass splittings within the well-known flavor-SU(3) octet and decuplet multiplets of hadrons.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

3.8.4 Symmetries at zero mass

The masses of the u and d quarks are quite small. When these masses are neglected, the QCD Lagrangian is further symmetric under separate SU(2) transformations on left- and right-handed quark fields defined by

$$\psi_L = \frac{1}{2}(1-\gamma_5)\psi, \qquad \psi_R = \frac{1}{2}(1+\gamma_5)\psi. \tag{3.75}$$

Then chiral $\mathrm{SU}(2)_L \otimes \mathrm{SU}(2)_R$ transformations have six parameters $\boldsymbol{\omega}_L$ and $\boldsymbol{\omega}_R$ for two commuting SU(2) groups, and the quark fields transform as

$$\begin{pmatrix} u_L \\ d_L \end{pmatrix} \mapsto e^{-i\boldsymbol{\omega}_L\cdot\boldsymbol{\sigma}/2} \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad \begin{pmatrix} u_R \\ d_R \end{pmatrix} \mapsto e^{-i\boldsymbol{\omega}_R\cdot\boldsymbol{\sigma}/2} \begin{pmatrix} u_R \\ d_R \end{pmatrix}. \tag{3.76}$$

The other fields (gluons, other quark flavors) are invariant. This symmetry is in fact spontaneously broken down to isospin SU(2). The low mass of the pions (about 140 MeV) relative to other hadrons is indicative of the expectation that they would be Goldstone bosons for spontaneously broken chiral symmetry in the limit of zero quark mass. Consequences can be successfully derived by the use of Ward identities together with the chiral transformation properties of the quark mass terms. These form much of the subject of current algebra.

3.8.5 Anomalies

When the u and d quarks are massless, the symmetry of their part of the Lagrangian appears also to include separate U(1) transformations on the left- and right-handed fields. (The quark-number symmetry corresponds to the same U(1) to both the left- and the right-handed fields.)

This symmetry is in fact anomalously broken. Thus, unlike the case of $\mathrm{SU}(2)_L \otimes \mathrm{SU}(2)_R$, there is no approximate Goldstone boson.

3.8.6 Chiral symmetry, hard scattering and factorization

When applying a factorization theorem like (1.1) there is a hard-scattering factor $\mathrm{d}\hat{\sigma}(\xi_a, \xi_b, i, j)$. This is normally computed with quark masses set to zero, and thus chiral symmetry applies to it.

Many consequences arise because at the quark-quark-gluon vertex, the coupling is only between quarks of the same helicity, and between quarks and antiquarks of the opposite helicities. That is, only the following transitions are possible:

$$q_L \leftrightarrow q_L + g, \qquad q_R \leftrightarrow q_R + g, \qquad q_L + \bar{q}_R \leftrightarrow g, \qquad q_R + \bar{q}_L \leftrightarrow g. \tag{3.77}$$

This produces many restrictions on the polarization dependence, as we will see in Secs. 11.6 and 13.16.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

3.9 Dealing with quark masses

Our basic technique for exploiting perturbation theory in QCD is to find quantities whose calculation has internal lines of Feynman graphs far off-shell, i.e., with some large virtuality Q^2. In these quantities we set the renormalization scale of order Q, so that the weakness of α_s at large scales allows the use of low-order perturbation theory, and we normally neglect quark masses.

However, there are quarks whose masses are not always negligible in these calculations, so that the general procedure needs modification to deal with heavy quarks. These are defined to be those quarks for which the coupling is small when the renormalization scale is of order the mass: $\alpha_s(m_q) \ll 1$. The known heavy quarks are c, b and t, with the remaining quarks and the gluon being called "light". The charm quark, of mass 1 to 1.5 GeV, is only marginally heavy, but, for robust observables, perturbation theory may be applicable at scales around the charm mass.

Clearly we need improved methods whenever Q, the physical of the process under consideration, is comparable to or smaller than the mass of one or more heavy quarks. First, we should not automatically neglect the mass. Second, the use of a mass-independent scheme, like $\overline{\rm MS}$, becomes unsuitable whenever the scale is *much* less than one of the quark masses.

The main issues are manifested in a calculation of the one-loop quark contribution to the gauge-field self-energy:

$$\begin{aligned}\Pi^{\mu\nu}_{\overline{\rm MS}} &= \sum_j \frac{-g^2\mu^{2\epsilon}\delta_{\alpha\beta}T_F}{(2\pi)^{4-2\epsilon}} \int \mathrm{d}^{4-2\epsilon}k \, \frac{\mathrm{Tr}\,(\not{k}+m_j)\gamma^\mu(\not{p}+\not{k}+m_q)\gamma^\nu}{\Big(k^2-m_j^2+i0\Big)\Big[(p+k)^2-m_j^2+i0\Big]} \\ &\quad + \text{counterterm}, \quad \text{with } \epsilon \to 0 \\ &= \sum_j \frac{-2i\alpha_s\delta_{\alpha\beta}T_F}{\pi}(-g^{\mu\nu}p^2+p^\mu p^\nu)\int_0^1 \mathrm{d}x\, x(1-x)\ln\frac{m_j^2-p^2x(1-x)}{\mu^2}.\end{aligned} \tag{3.78}$$

The following properties apply to this graph and more generally.

- If $|p^2|$ is large compared with m_j^2, then m_j can be neglected, with relative errors of order $m_j^2/|p^2|$.
- Furthermore, in the same situation, $|p^2| \gg m_j^2$, there is logarithmic dependence on p^2. The large logarithm can be removed by taking μ^2 of order $|p^2|$.
- If $|p^2|$ is much less than m_j^2, the integral approaches a constant, $\ln(m_j^2/\mu^2)$. In (3.78), this multiplies a factor quadratic in p, of the same momentum dependence as the UV counterterm.

The last item exemplifies the non-trivial part of the decoupling theorem for heavy particles (Appelquist and Carazzone, 1975). This theorem concerns a situation where we hold fixed the external scales of a Green function and make some internal mass much larger. Then the contributions of *convergent* graphs with the large internal mass are suppressed.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

The suppression fails whenever the heavy internal lines are in a divergent loop, but the unsuppressed contributions are equivalent to a contribution to renormalization counterterms. Thus the unsuppressed contributions can be eliminated by a choice of counterterm.

Suppose that we have the real-world situation that the quark masses are widely different. Then we can have a conflict in the choice of μ that eliminates large logarithms, whenever $|p|^2$ lies between two heavy quark masses, e.g., $m_t^2 \gg |p^2| \gg m_b^2$, which is common in practice. Different graphs for the same process involve different heavy quarks.

If we use $\overline{\text{MS}}$ renormalization, then, for the quarks that are heavy on a scale of p^2, we have logarithms $\ln(m_j^2/\mu^2)$, which can be removed by setting $\mu \sim m_j$. For the quarks that are light on a scale of p^2, we have logarithms $\ln(-p^2/\mu^2)$, which are removed by setting $\mu^2 \sim |p^2|$. When the quark masses and $|p^2|$ cover a wide range, we have incompatible conditions on μ.

The original way of using the decoupling theorem was to define a second theory in which all fields are omitted whose masses are much larger than the external scales. This is the low-energy effective theory (LEET) for a given set of heavy quarks. The renormalized parameters of the LEET have numerical values that, in general, differ from those of the full theory. These numerical values can be computed by comparing calculations of Green functions in the two theories and requiring that they give equivalent results.

A LEET removes from calculations quarks whose masses are much larger than the external scales. There can remain quarks with masses comparable to the external scales. For example, in a calculation at $Q \sim 5$ GeV, we would decouple the t quark, but none of the others, so that the LEET has five quark fields. But we could not neglect the mass of the b quark. Depending on the situation and required accuracy, we might be able to neglect the charm quark mass or might need to retain its mass. One normally neglects all three light quark masses in standard perturbative calculations.

For a full set of QCD calculations, we need to successively decouple the top, bottom and charm quarks. This gives us a series of effective theories with three, four and five quarks, with corresponding values of their $\overline{\text{MS}}$ couplings. Non-perturbative calculations at low scales are normally done in the 3-flavor effective theory; these include the well-known lattice Monte-Carlo simulations.

However, the method of LEETs has certain disadvantages, and in the next section I present a better method. The primary disadvantage of a LEET is that it is limited in the ultimate accuracy that it can achieve. For example, consider the 3-flavor effective theory. We could obtain it by sequential decoupling of the three heavy quarks. Now, the decoupling of the charm quark, to get the final 3-flavor LEET, assumes that it is much lighter than the previously decoupled bottom quark; so we have the leading term in an expansion in powers of m_c/m_b. But this ratio is only about $1/3$, so the errors could be quite large relative to a desirable accuracy. If instead we decouple both the charm and bottom quarks in one step, then the matching conditions would include logarithmic dependence on m_b/m_c, which would also reduce the accuracy.

A more general approach is to change the renormalization scheme to make decoupling more manifest. The simplest of such schemes is momentum-space subtraction, in which the

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

counterterms are chosen to set certain 1PI Green functions (and/or appropriate derivatives) to zero at a particular point in momentum space. For the quark self-energy, we could choose the renormalization point to be $p^2 = -\mu^2$, obtaining

$$\Pi^{\mu\nu}_{\rm MOM} = \sum_j \frac{-2i\alpha_s \delta_{\alpha\beta} T_F}{\pi} (-g^{\mu\nu} p^2 + p^\mu p^\nu) \times \int_0^1 {\rm d}x\, x(1-x) \ln \frac{m_j^2 - p^2 x(1-x)}{m_j^2 + \mu^2 x(1-x)}. \tag{3.79}$$

This scheme solves the difficulty of removing all large logarithms; these are eliminated by setting μ^2 of order $|p^2|$, independently of the size of m_j. Thus the scheme satisfies manifest decoupling, which means that we obtain the low-energy effective theory simply by deleting all graphs containing quarks much heavier than the external scale. The errors in doing this are a power of p^2 divided by the square of the mass of the lightest deleted quark.

But the scheme has two technical disadvantages. One is that gauge invariance is not automatically preserved. The defined momentum-space subtractions can only be applied to a limited set of 1PI Green functions, sufficient to determine an independent set of renormalization factors. The counterterms for the remaining 1PI divergent graphs are determined by gauge invariance, and will generally not have an obvious momentum-space definition. Indeed, a separate argument will be necessary to prove decoupling.

The second disadvantage is the practical one that the counterterms are mass dependent, so that the renormalization-group equations for the coupling and mass will be complicated and coupled. So the solution will be much more complicated and more difficult to overview. Moreover, the calculations of counterterms become algorithmically much more complicated: the exact values of off-shell Green functions are needed instead of just the pole part at $\epsilon = 0$. Calculation of on-shell Green functions is generally simpler than when they are off-shell, and calculations of the pole parts are even easier. This was nicely illustrated in our calculation of the quark-quark-gluon vertex graph in Sec. 3.3. This is an important issue, since high-order calculations are extremely expensive in time and effort, which rapidly increases with the order of the calculation. Moreover, for a given desired accuracy in a final phenomenological result, it is generally necessary to compute RG coefficients to one order higher than everything else, because the RG coefficients get integrated over a large range of scales, thereby increasing the effect of an error due to uncalculated higher-order corrections.

3.10 CWZ (ACOT) method for heavy quarks

A method that overcomes these complications was constructed by Collins, Wilczek, and Zee (1978) (CWZ). This method is actually a composite scheme, composed of a sequence of subschemes. The subschemes are parameterized by what is called the number of active quarks, $N_{\rm act}$. The active quarks are the $N_{\rm act}$ lightest, and the inactive are the remaining, heavier quarks. Since the gluon has zero mass, it is always treated as active. For a 1PI graph

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

containing only active quarks, normal $\overline{\text{MS}}$ counterterms are used. But *zero*-momentum subtractions are used for any 1PI graph that has at least one internal line for an inactive quark.

Normally, zero-momentum counterterms would have undesirable IR divergences in a theory with massless fields, like the gluons in QCD. But the presence of at least one massive line removes these divergences, to all orders of perturbation theory.

The CWZ scheme has the following advantages.

- Each subscheme automatically satisfies gauge invariance. That is, if the counterterms in the Lagrangian are determined by some minimal set of 1PI Green functions, then the remaining 1PI Green functions, with their counterterms determined by gauge invariance of $\mathcal{L}$ also obey the CWZ renormalization condition. No extra finite counterterms are needed.
- Manifest decoupling is satisfied in each scheme. In particular, the numerical value of the coupling in the LEET with N_{act} flavors and pure $\overline{\text{MS}}$ renormalization is the same as in the CWZ subscheme with N_{act} active quarks.
- The RG coefficients in each subscheme are mass independent and in fact exactly identical to those in the theory obtained by deleting the inactive quarks.
- This apparently violates the theorem that we have scheme independence of the one- and two-loop terms in β, and of the one-loop terms in the other RG coefficients. But the theorem only applies if the counterterms are mass independent, which is not the case here, when the number of active quarks changes.
- Normally, calculations of Green functions at zero external momentum are much easier than with a general external momentum.
- No IR divergences are induced by the use of zero-momentum subtractions.

Since there is a sequence of subschemes, relations must be derived between the renormalized parameters in the subschemes. This is quite straightforward, with some results listed below. Moreover, there are no large logarithms in relating the subscheme with N_1 active quarks to the scheme with $N_1 + 1$ active quarks, provided only that μ is of order the mass of the single quark that is making the transition between active and inactive. We will see examples later.

This scheme has become a standard, e.g., Bethke (2009). It extends quite simply to the treatment of parton densities, etc., in which case it is called the ACOT scheme, as expounded by Aivazis *et al*. (1994). It is the one I will use throughout this book, unless otherwise specified.

An important misapprehension needs to be eliminated from the beginning. This is that the $\overline{\text{MS}}$ scheme only applies to massless quarks. It is true that RG coefficients (and their generalizations) do not depend on the quark masses. For this and other reasons, it is often best to do many calculations with massless quarks. But there is no intrinsic reason for it to be restricted to massless quarks. The misapprehension is coupled with some severe conceptual misunderstandings concerning the factorization theorems of QCD, as we will see in later chapters.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

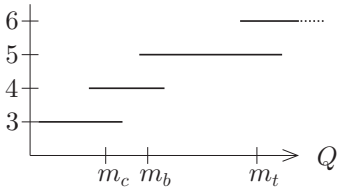


Fig. 3.7. Range of scales for which particular numbers of active flavors are appropriate.

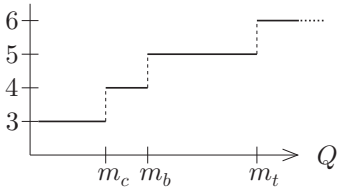


Fig. 3.8. Possible choice of switching points between CWZ subschemes.

3.11 Relating CWZ subschemes with different numbers of active quarks

For a particular CWZ subscheme with a given number, N, of active quarks, the vacuum polarization in (3.78) is replaced by

$$\Pi^{\mu\nu}_{\rm CWZ} = \frac{-2i\alpha_{s,N}\delta_{\alpha\beta}T_F}{\pi}(-g^{\mu\nu}p^2 + p^\mu p^\nu)\int_0^1 {\rm d}x\, x(1-x)$$
$$\times\left[\sum_{\text{active } j}\ln\frac{m_j^2 - p^2x(1-x)}{\mu^2} + \sum_{\text{inactive } j}\ln\frac{m_j^2 - p^2x(1-x)}{m_j^2}\right], \qquad (3.80)$$

where $\alpha_{s,N}(\mu)$ is the coupling appropriate to the subscheme. For a particular value of p^2, to eliminate large logarithms, we should (a) take μ^2 of order $|p^2|$, (b) make inactive all quarks with $m_j^2 \gg |p^2|$, and (c) make active all quarks with $m_j^2 \ll |p^2|$. Obviously, for quarks with $m_j^2 \sim |p^2|$ we have a choice of whether to make them active or inactive, as illustrated in Fig. 3.7. In the past, there was a tendency to make a definite switching point between subschemes: quark j was considered active if $\mu > m_j$, and inactive otherwise. But this is now seen as undesirable.

At one-loop, the relations between the subschemes are readily computed from the vacuum polarization graphs, as we will now see. Let us define $Z_{3,N}$ to be the value of Z_3 when the lightest N quarks are active, and similarly for $\tilde{Z}$ and the renormalized masses and coupling. Let $Z_{2,N,j}$ be the field strength renormalization for quark j.

3.11.1 Field-strength renormalization

At one-loop, the self-energies of the first N quarks and the ghost have no inactive quark lines, so $\overline{\rm MS}$ counterterms apply in both of the subschemes we are relating. Similar considerations

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

apply to the quarks which are inactive in both schemes.

$$\tilde{Z}_N = \tilde{Z}_{N+1} + O(\alpha_s^2), \tag{3.81}$$

$$Z_{2,N,j} = Z_{2,N+1,j} + O(\alpha_s^2), \qquad \text{if } j \le N \text{ or } j \ge N+2, \tag{3.82}$$

$$Z_{m,N,j} = Z_{m,N+1,j} + O(\alpha_s^2), \qquad \text{if } j \le N \text{ or } j \ge N+2. \tag{3.83}$$

Here, we use a notation in which the quark label j equals its sequence number in order of mass.

However, the counterterm for the gluon self-energy changes. From the earlier calculations we have

$$Z_{3,N} = Z_{3,\overline{\rm MS}} + \frac{\alpha_s S_\epsilon}{3\pi} T_F (n_f - N) \sum_{j>N} \left[\frac{\Gamma(\epsilon)}{e^{-\gamma_{\rm E}\epsilon}} \left(\frac{\mu^2}{m_j^2} \right)^\epsilon - \frac{1}{\epsilon} \right] + O(\alpha_s^2). \tag{3.84}$$

Bare quantities, including fields, are the same in *all* schemes. We therefore obtain the following relations between the fields and masses in the two subschemes:

$$A_N = A_{N+1} \left[1 + \frac{\alpha_s}{6\pi} T_F \ln \frac{\mu^2}{m_{N+1}^2} + O(\alpha_s^2) \right], \tag{3.85a}$$

$$\eta_N = \eta_{N+1}[1 + O(\alpha_s^2)], \tag{3.85b}$$

$$\psi_{j,N} = \psi_{j,N+1}[1 + O(\alpha_s^2)], \qquad \text{if } j \le N \text{ or } j \ge N+2, \tag{3.85c}$$

$$m_{j,N} = m_{j,N+1}[1 + O(\alpha_s^2)], \qquad \text{if } j \le N \text{ or } j \ge N+2, \tag{3.85d}$$

3.11.2 Coupling

Now consider the vertex for the ghost to a gluon. Its counterterm is pure $\overline{\rm MS}$ in both subschemes, and the counterterm is computed from g_0 and the Z factors as proportional to $g_0 \tilde{Z} Z_{3,N} - \mu^\epsilon g_N + O(g^5)$. The bare coupling is the same in both subschemes, so it follows that the renormalized coupling has the relation

$$\alpha_{s,N} = \alpha_{s,N+1} \left[1 - \frac{\alpha_s}{2\pi} T_F \ln \frac{\mu^2}{m_{N+1}^2} + O(\alpha_s^2) \right]. \tag{3.86}$$

Evidently, at the one-loop order, it is sufficient to compute the vacuum polarization.

Higher-order corrections to these relations have been made. For two-loop calculations, see Bernreuther and Wetzel (1982); Bernreuther (1983a, b). For three-loop calculations, see Chetyrkin, Kniehl, and Steinhauser (1997, 1998).

Exercises

3.1 Complete the calculation of the renormalization of QCD at one-loop order. The most economical method is probably to calculate the gluon, quark and ghost self-energies

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press

in addition to the quark-gluon vertex.[9] You will have thus verified for yourself the asymptotic freedom of QCD.

3.2 Given the values of the renormalization-group coefficients, reconstruct formulae for the $\overline{\text{MS}}$ renormalization factors for the coupling, and for the fields to at least two-loop order. You may find the results useful if you ever do serious perturbative QCD calculations.

(One method is to treat (3.41), etc., as differential equations determining renormalization factors from the RG coefficients. Solve these order-by-order in powers of the renormalized coupling. Then apply the boundary conditions that the Z factors and $g_0/g\mu^\epsilon$ go to unity at zero renormalized coupling.)

3.3 (**) There are competing definitions of the $\overline{\text{MS}}$ scheme. Show that these definitions all agree in the values of renormalized Green functions at $\epsilon = 0$, provided that S_ϵ in the different definitions agree to order ϵ.

3.4 Find the next term in the expansion (3.63) of the effective coupling. This will be $1/\ln^3(\mu^2/\Lambda^2)$ times a quadratic polynomial in $\ln\ln(\mu^2/\Lambda^2)$. To check your answer, see (9.5) of Amsler *et al*. (2008), but beware of different conventions for defining the β_j coefficients.

[9] The calculation of the three- and four-point gluon gluon functions is substantially more complicated, and should only be attempted if you have much time and wish to verify the general theorems on the renormalizability of non-abelian gauge theories. It is also possible to work with the ghost-gluon coupling, although this is a little more complicated, because it has a derivative coupling.

https://doi.org/10.1017/9781009401845.003 Published online by Cambridge University Press