8

Large-mass expansion

A common situation in physics is that in investigating phenomena on a certain distance scale, one sees no hint of those phenomena that happen at much shorter distance scales. In a classical situation this observation seems evident. For example, one can treat fluid dynamics without any knowledge of the atomic physics that generates the actual properties of the fluids. However, in a quantum field theory this decoupling of short-distance phenomena from long-distance phenomena is not self-evident at all.

Consider an $e^+ - e^-$ annihilation experiment at a center-of-mass energy well below 10 GeV, the threshold for making hadrons containing the b-quark. There is, for practical (or experimental) purposes, no trace of the existence of this quark. However, the quark *is* present in Feynman graphs as a virtual particle, and can have an apparently significant effect on cross-sections. Our task in this chapter is therefore to prove what is known as the decoupling theorem. This states that a Feynman graph containing a propagator for a field whose mass is much greater than the external momenta of the graph is in fact suppressed by a power of the heavy mass. The physics at low energy is described by an effective low-energy theory that is obtained by deleting all heavy fields from the original theory.

The decoupling of heavy particles is not absolutely universal. One important and typical exception is that of weak interactions. Let us consider the interactions of hadrons at energies of a few GeV. The effective low-energy theory, in the sense just described, consists of strong and electromagnetic interactions alone, without weak interactions. So weak interactions should be ignorable at low energies. However, it is well known that there are in fact many observed processes, particularly decays, that are due entirely to weak interactions. The point is that, in the absence of weak interactions, these processes are exactly forbidden by symmetries, such as parity, charge-conjugation, and strangeness conservation. Weak-interaction amplitudes for the processes in question would be power-law corrections – suppressed by a power of energy divided by the mass of the W-boson – were it not that they are corrections to zero. The consequence of this particular situation is that, at low energies, weak interactions are

described by a *non*-renormalizable theory, viz., the four-fermion interaction. Efforts to find a renormalizable theory led to gauge theories, and a prediction of the W- and Z-bosons from phenomena at energies much lower than their masses: low-energy phenomena have indeed provided clues as to what might happen at much higher energy.

In this chapter, we will treat the cases where decoupling occurs. The theorem that tells us to expect decoupling to occur in many theories was formalized by Appelquist & Carazzone (1975) and Symanzik (1973). They work with a renormalizable theory in which some fields have masses very large compared with the others. They then consider Green's functions of the low-mass fields at momenta much less than the large masses. The theorem is that the Green's functions are the same as those in an effective low-energy theory obtained by deleting all of the heavy fields. Corrections are smaller by a power of momentum divided by a heavy mass. The sole effect of loops of heavy particles is that the couplings of the low-energy theory can have different values from those in the complete theory.

Since the renormalized couplings have no particular *a priori* value, the heavy particles are unobservable until close to threshold. The practical importance of the theorem is that one can understand low-energy physics without having a complete Lagrangian for all phenomena.

We will also show how the renormalization group can be applied in the computation of the relation between the couplings of the low-energy effective theory and those of the full theory.

There are many ramifications of the decoupling theorem, but we will not treat these. One of these is the detailed application of the decoupling theorem to gauge theories (see, for example, Kazama & Yao (1982)). Another is the large-mass expansion of Witten (1976) – where Green's functions of the heavy fields are computed; this expansion is used is deep-inelastic scattering.

We will also not treat the exceptions to the decoupling theorem. These can be treated by the same techniques as those used to prove the decoupling theorem itself. We have already mentioned weak interactions as one of the typical exceptions. Let us just note two other main classes of exception:

(1) In theories with spontaneous symmetry breaking, a mass is often made large by increasing a dimensionless coupling (Veltman (1977) and Toussaint (1978)). The decoupling theorem assumes that a mass is made large by increasing dimensional parameters.
(2) Some dimensionless couplings needed by power-counting violate renormalizability of the low-energy theory (see Collins, Wilczek & Zee (1978)).

In any event the effective low-energy theory is non-renormalizable.

It might be supposed that since General Relativity is non-renormalizable in perturbation theory, it contains some clues to phenomena at very high energies (see, for example, Hawking & Israel (1979)).

8.1 A model

We will restrict our attention to a very simple model. It is a ϕ^3 theory with two fields in six space-time dimensions:

$$\mathscr{L} = (\partial\phi_l)^2/2 + (\partial\phi_h)^2/2 - m^2\phi_l^2/2 - M^2\phi_h^2/2 - \mu^{3-d/2}[g_1\phi_l^3/6 + g_2\phi_l\phi_h^2/2] - \mu^{d/2-3}f\phi_l + \text{counterterms}. \quad (8.1.1)$$

Symmetry under $\phi_h \to -\phi_h$ has been imposed to cut down the number of possible couplings. Then (8.1.1) contains all couplings necessary for renormalizability. We assume that the renormalized mass, M, of the heavy field is made large while all other parameters are held finite. The factors of the unit of mass μ needed with dimensional regularization are explicitly indicated.

All our techniques can be readily extended to treat more complicated (realistic) theories.

As usual we have introduced a linear term in the Lagrangian to cancel the tadpole graphs. This is determined by the renormalization condition that $\langle 0|\phi_l|0\rangle = 0$.

The remaining counterterms can be put in the form

$$\begin{aligned}\mathscr{L}_{ct} = {} & (Z_l - 1)\partial\phi_l^2/2 + (Z_h - 1)\partial\phi_h^2/2 \\ & - [m^2(Z_m - 1) + M^2 Z_{mM}]\phi_l^2/2 - [M^2(Z_M - 1) + m^2 Z_{Mm}]\phi_h^2/2 \\ & - \mu^{3-d/2}[(g_{1B} - g_1)\phi_l^3/6 + (g_{2B} - g_2)\phi_l\phi_h^2/2] \\ & - \mu^{d/2-3}(f_B - f)\phi_l. \end{aligned} \quad (8.1.2)$$

As usual, we may choose the dimensionless renormalizations (viz., the Z's and the g_B's) to be independent of the dimensional parameters m^2, M^2, and f.

The decoupling theorem asserts that phenomena on energy scales much less than M are described by an effective low-energy theory whose Lagrangian has the form

$$\begin{aligned}\mathscr{L}_{eff} &= z\partial\phi_1^2/2 - m^{*2}z\phi_1^2/2 - \mu^{3-d/2}g^* z^{3/2}\phi_1^3/6 - \mu^{d/2-3}z^{1/2}f^*\phi_1 \\ &\quad + \text{counterterms} \\ &= \partial\phi^{*2}/2 - m^{*2}\phi^{*2}/2 - \mu^{3-d/2}g^*\phi^{*3}/6 - \mu^{d/2-3}f^*\phi^* \\ &\quad + \text{counterterms}. \end{aligned} \quad (8.1.3)$$

Here we have defined a scaled field $\phi^* = z^{1/2}\phi_l$. We will prove that g^*, m^* and the coefficient z can be chosen so that Green's functions of ϕ_l obtained from $\mathscr{L}_{\text{eff}}$ differ from those obtained from the full Lagrangian (8.1.1) by terms which are of the order of a power of external momenta divided by M.

It is usually convenient to work with the scaled field ϕ^* which has unit coefficient for its kinetic term in the basic Lagrangian. Then Green's functions in the full theory are related to Green's functions in the low-energy theory by

$$\begin{aligned}\tilde{G}_N(p_1,\ldots,p_N;g_l,g_h,m,M,\mu) &= \langle 0|T\tilde{\phi}_l(p_1)\ldots\tilde{\phi}_l(p_N)|0\rangle_{\text{full theory}} \\ &= z^{-N/2}\tilde{G}^*_N(p_1,\ldots,p_N;g^*,m^*,\mu)[1+O(1/M^a)] \\ &= z^{-N/2}\langle 0|T\tilde{\phi}^*(p_1)\ldots\tilde{\phi}^*(p_N)|0\rangle[1+O(1/M^a)], \quad (8.1.4)\end{aligned}$$

as $M\to\infty$ with $p_1,\ldots,p_N$ fixed. The fractional errors go to zero as a power of M times logarithms; the power is typically M^{-2}. We can therefore use $1/M^a$, with a slightly less than two, to bound the error. As is our convention, the tilde signs over the fields and Green's functions indicate a Fourier transform into momentum space.

8.2 Power-counting

In this section, we will establish the rules for finding the leading power of M in the value of a graph as $M\to\infty$. These form a simple generalization of Weinberg's theorem, and will involve us in understanding which regions of momentum space are important. We will mostly be interested in graphs for the Green's functions of the light field ϕ_l. Our aim will be to find those graphs that contain lines for the heavy field and that do not vanish as M goes to infinity.

8.2.1 *Tree graphs*

Because we choose to impose the symmetry $\phi_h\to-\phi_h$, the only tree graphs containing lines for the heavy field have heavy external lines. An example is Fig. 8.2.1. Since all momenta on the lines are fixed, and since the free ϕ_h-propagator is

$$i/(p^2-M^2)\sim -i/M^2,$$

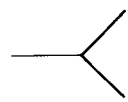


Fig. 8.2.1. A tree graph with a heavy line.

the behavior of any given tree graph as $M \to \infty$ is

$$M^{-2H}, \tag{8.2.1}$$

where H is the number of heavy lines. (We use the natural terminology of calling a line of a Feynman graph heavy or light according to whether its free propagator is for ϕ_h or ϕ_l respectively. Our graphical notation is that heavy lines are thicker than light lines.)

8.2.2 *Finite graphs with heavy loops*

Consider now a graph that has one or more loops but no ultra-violet divergences or subdivergences, and that has some heavy internal lines. The lowest-order example is Fig. 8.2.2(*a*), for the four-point function. Its external momenta (if small) may evidently be neglected on the lines of the loop, whose value is then

$$\Gamma_2 = \frac{g_2^4}{(2\pi)^6} \int \mathrm{d}^6 k \frac{1}{(k^2 - M^2)^4} = \frac{\mathrm{i} g_2^4}{384\pi^3 M^2}. \tag{8.2.2}$$

(We label the symbol Γ by the figure number.)

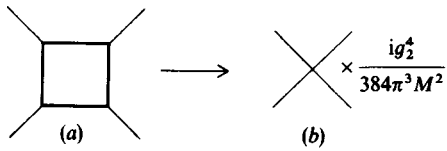


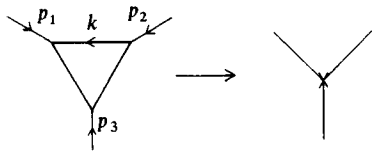
Fig. 8.2.2. Large-mass behavior of graph without an ultra-violet divergence.

The graph vanishes as $M \to \infty$. The precise power of M can be obtained by considering the possible regions of momentum space (after Wick rotation), as follows. Any region of k that is finite as $M \to \infty$ gives a contribution of order M^{-6}. Since the graph is UV finite, the only other possibility is $k = O(M)$. Simple power-counting gives M^{-2}, as found in (8.2.2). This power-counting is the same as for the UV degree of divergence.

The graph is negligible (by a power of M^2) compared to graphs with no heavy lines. If, nevertheless, we wanted its leading contribution, then it would be effectively the local four-point vertex symbolized in Fig. 8.2.2(*b*). The non-renormalizability of this coupling (when the space-time dimension is six) is tied to the negative power of M^2.

8.2.3 *Divergent one-loop graphs*

Consider the logarithmically divergent vertex graph in Fig. 8.2.3(*a*). After



(a) (b)

Fig. 8.2.3. Large-mass behavior of graph with an ultra-violet divergence.

minimal subtraction the loop gives

$$R(\Gamma_3) = \frac{ig_2^3}{64\pi^3}\left\{\tfrac{1}{2}\gamma + \int_0^1 dx \int_0^{1-x} dy \times \right.$$
$$\left. \times \ln\left[\frac{M^2 - (p_1^2 x + p_2^2 y)(1-x-y) - p_3^2 xy}{4\pi\mu^2}\right]\right\}$$
$$= \frac{ig_2^3}{128\pi^3}\left[\gamma + \ln\left(\frac{M^2}{4\pi\mu^2}\right) + O\left(\frac{p_i^2}{M^2}\right)\right]. \tag{8.2.3}$$

The same power-counting as for Fig. 8.2.2 confirms the power M^0 for the leading behavior as $M \to \infty$. There is also a logarithm. This occurs because two regions contribute to the leading behavior: the first is where the loop momentum k is of order M. The second region is the UV region where $k \to \infty$. After subtraction of the ultra-violet divergence a finite contribution remains.

Evidently the graph gives a contribution that increases with M. Fortunately the non-vanishing part of the loop is independent of the external momenta. So for large M, the loop is effectively a three-point vertex, as shown in Fig. 8.2.3(*b*). A proof which generalizes to higher order is to differentiate with respect to any external momentum. Since the differentiated graph is finite, it vanishes when $M \to \infty$, like a power of M.

Recall our statement of the decoupling theorem, that at low energies, we could calculate Green's functions from the effective low-energy Lagrangian (8.1.3). The result of our calculation of Fig. 8.2.3 is that the graph generates an extra piece in the ϕ_1^3 coupling of the low-energy theory. Let us therefore write

$$z^{3/2} g^* = g_1 - \frac{g_2^3}{128\pi^3}\left[\gamma + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right] + O(g^5). \tag{8.2.4}$$

We may drop the graph Fig. 8.2.3 and replace it by the order g^3 term on the right of (8.2.4). The loop has been replaced by a local vertex where all the lines come to a single point. This corresponds to the fact that the internal line is far off-shell and can only exist for a short time.

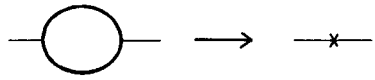


Fig. 8.2.4. Large-mass behavior of graph with a quadratic ultra-violet divergence.

The self-energy graph, Fig. 8.2.4, gives a leading term of order M^2. The value of the loop is

$$\begin{aligned} R(\Gamma_4) &= \frac{ig_2^2}{128\pi^3}\left\{(\gamma-1)(M^2-\tfrac{1}{6}p^2)\right. \\ &\quad \left. + \int_0^1 dx\,[M^2 - p^2x(1-x)]\ln\left[\frac{M^2-p^2x(1-x)}{4\pi\mu^2}\right]\right\} \\ &= \frac{ig_2^2}{128\pi^3}\left\{M^2\left[\gamma - 1 + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right] - \frac{p^2}{6}\left[\gamma + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right]\right. \\ &\quad \left. + O\left(\frac{p^2}{M^2}\right)\right\}. \end{aligned} \tag{8.2.5}$$

Again the loop momentum k can be either UV or of order M to contribute, so there will be at most a single logarithm of M^2/μ^2. Since we have to differentiate three times with respect to p_μ before obtaining a convergent graph, the non-vanishing terms, as $M \to \infty$, are quadratic in p. From the effective Lagrangian (8.1.3), we see that the graph may be replaced by a contribution to the basic self-energy vertex $i[(z-1)p^2 - (m^{*2}z - m^2)]$ in the low-energy theory, with

$$z = 1 - \frac{g_2^2}{768\pi^3}\left[\gamma + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right] + O(g^4), \tag{8.2.6}$$

$$zm^{*2} = m^2 - \frac{g_2^2M^2}{128\pi^3}\left[\gamma - 1 + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right] + O(g^4). \tag{8.2.7}$$

We can now compute g^* and m^*:

$$g^* = g_1 - \frac{g_2^2(g_2 - \frac{1}{4}g_1)}{128\pi^3}\left[\gamma + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right] + O(g^5), \tag{8.2.8}$$

$$\begin{aligned} m^{*2} = m^2 - \frac{g_2^2}{128\pi^3}&\left\{M^2\left[\gamma - 1 + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right]\right. \\ &\left. - \tfrac{1}{6}m^2\left[\gamma + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right]\right\} + O(g^4). \end{aligned} \tag{8.2.9}$$

Notice that there is a contribution of order M^2 to the self-energy and

hence to m^{*2}. In order to keep the physical mass of ϕ_1 finite and hence keep m^* finite as $M \to \infty$, we must let m^2 have a term proportional to $g_2^2 M^2$ (with higher-order corrections):

$$m^2 = \text{finite} + g_2^2 \frac{M^2}{128\pi^3}\left[\gamma - 1 + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right] + \text{higher order}. \quad (8.2.10)$$

On expanding m^{*2} in powers of coupling, we find

$$m^{*2} = \text{finite term in } m^2 + \tfrac{1}{6}\left(\frac{g_2^2 m^2}{128\pi^3}\right)\left[\gamma + \ln\left(\frac{M^2}{4\pi\mu^2}\right)\right] + \text{higher order}. \quad (8.2.11)$$

Since m is the mass parameter for the light field, it is generally considered unnatural to have to fine-tune it within a fractional accuracy of m^{*2}/M^2, as is required by (8.2.10), to obtain a finite value of m^* when $M \to \infty$. In the context of grand unified theories this is called the gauge hierarchy problem (Weinberg (1974, 1976), Gildener & Weinberg (1976)). It is hoped to solve it by finding a phenomenologically sensible theory with no need for fine-tuning.

8.2.4 *More than one loop*

We may have one of the divergent one-loop graphs occurring inside a larger superficially convergent graph. A typical example is Fig. 8.2.5. When

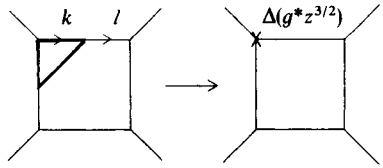


Fig. 8.2.5. Large-mass behavior of two-loop graph with an ultra-violet divergence.

$M \to \infty$ with the external momenta fixed, the only region of the loop momenta that gives a non-zero contribution is where the outer-loop momentum l is finite and the inner-loop momentum k is of order M or larger. So the heavy loop can be replaced by its effective low-energy vertex computed at (8.2.3). This procedure does not change the overall degree of divergence.

The situation at higher order or with overall-divergent graphs is more subtle as we will now see. The graph of Fig. 8.2.6 is typical. Now, it contains a subgraph, consisting of the heavy loop, which we have already considered.

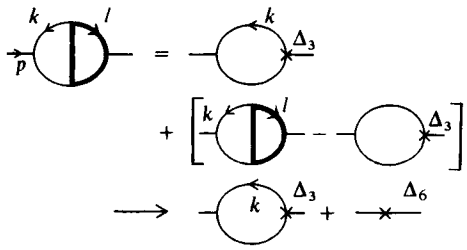


Fig. 8.2.6. Large-mass behavior of another two-loop graph with an ultra-violet divergence.

Therefore, the low-energy theory contains a graph where the heavy loop is replaced by a vertex using (8.2.4) for $g^* z^{3/2} - g$. This graph exactly reproduces the region where k is finite and l is large for Fig. 8.2.6. We add and subtract this graph from the original graph as indicated in the figure. The subtracted term (in square brackets) has a vanishing contribution from finite k (as $M \to \infty$). So we replace it by an effective vertex Δ_6. The same arguments as we used for one-loop self-energy, Fig. 8.2.4, show that it has three terms, proportional to p^2, m^2, and M^2, with coefficients polynomial in $\ln(M^2/\mu^2)$.

In this and in other graphs there are UV divergences for the whole graph and for subgraphs. Implicitly, the counterterm graphs are to be included. Provided we use mass-independent renormalization we are guaranteed that the counterterm graphs satisfy the same power-counting as the original graphs. In particular they are polynomial in the light masses. Thus the counterterm graphs do not change the power-counting and differentiation arguments that are crucial to our work.

8.3 General ideas

Structurally, the arguments in the last section appear similar to those we used in Chapter 7 to show that renormalization-prescription dependence can be compensated by finite counterterms. In fact, as we will see in the next section, Section 8.4, a proof of the decoupling theorem can be constructed exactly by changing the renormalization prescription. We will show that a renormalization prescription can be chosen to have a number of convenient properties, the most important of which is that the low-energy theory is constructed simply by deleting all heavy fields without changing the couplings and masses of the light fields. This property is called manifest decoupling, and we will explain it with the aid of an example in subsection 8.3.1.

Our approach follows the method given by Appelquist & Carazzone (1975) and Witten (1976). This approach generates the effective theory as a series of subtractions. The simplest non-trivial case is given in Fig. 8.2.6.

There is another approach due to Weinberg (1980) in which the decoupling is considered by first integrating over the heavy fields in the functional integral. (See also Ovrut & Schnitzer (1980).) This method is less convenient for treating graphs like Fig. 8.2.6, so we do not use it.

8.3.1 *Renormalization prescriptions with manifest decoupling*

Suppose we used BPH(Z) renormalization instead of minimal subtraction. Then the renormalization condition is that the terms up to $p^{\delta(\Gamma)}$ are zero in the Taylor expansion of a graph Γ about zero external momentum. Here $\delta(\Gamma)$ is the degree of divergence. For a graph with a single loop, consisting of a heavy line, these terms are precisely those that are non-vanishing as $M \to \infty$. Examples are given by the graphs of Figs. 8.2.3 and 8.2.4.

In fact, for a general graph, the effective low-energy theory in this renormalization presciption is obtained merely by deleting all graphs containing heavy lines, together with all their counterterm graphs. The values of the couplings and masses are not changed. Therefore the BPH(Z) prescription has the property we called 'manifest decoupling'. It might appear sensible always to use a renormalization prescription that has this property. However, for many purposes it is useful to use other renormalization prescriptions, e.g. minimal subtraction and its relatives. Particular cases are theories containing massless fields, especially non-abelian gauge theories, and theories with spontaneous symmetry breaking. In any case, it is good to have a direct method of proof of decoupling that can work with any prescription. Furthermore, a prescription like minimal subtraction is more convenient if one also wishes to compute high-energy behavior (Section 7.4) with the aid of the renormalization group. In fact, the method we will use will start from a mass-independent renormalization prescription defined for both the full theory and for the effective low-energy theory. Then the renormalization of the low-energy theory is extended to a renormalization prescription of the full theory in such a way as to satisfy manifest decoupling. This method was first stated by Collins, Wilczek & Zee (1978).

One renormalization prescription that gives manifest decoupling at low energies and that allows the use of renormalization-group methods at high energies is due originally to Gell-Mann & Low (1954). In this scheme, one makes subtractions at some arbitrarily chosen value of momentum. This

scheme was applied to the large-mass problem by Georgi & Politzer (1976). The disadvantage of this scheme, compared with the scheme that we will actually use, is that renormalization-group coefficients are explicitly functions of M/μ, and of m/μ:

$$\beta_1 = \beta_1(g_1, g_2; M/\mu, m/\mu).$$

This makes calculations complicated. Furthermore, this scheme obscures some symmetries.

8.3.2 *Dominant regions*

Before actually constructing a proof of the decoupling theorem, let us give a precise statement of the regions that give unsuppressed contributions (i.e. not suppressed by a power of M^2). We consider each graph in the full theory together with the set of subtraction graphs needed to cancel its divergences and subdivergences. We do not consider the subtraction graphs separately.

First of all, any graph with no heavy lines at all contributes without suppression.

A graph with one or more heavy lines cannot give a contribution unless at least one loop momentum is of order M. The contribution to a graph when M is large can be considered as the sum of contributions from various possible regions of momentum space. The regions can be specified by the sizes of the loop momenta. For our purposes, it is enough to classify a momentum as either finite or large. 'Large' we define to mean 'of order M or bigger'. We can do power-counting for each region in the obvious way. For the loops carrying large momenta, counting powers of M is the same as for the ultra-violet degree of divergence. This gives a factor M^δ, where δ is the ultra-violet degree of the lines carrying large momenta. A heavy line carrying finite momentum counts as M^{-2}. A light line carrying finite momentum counts as M^0.

The leading power of M for a graph is obtained as the maximum of the powers for the possible regions. Let us define $\delta_M(\Gamma)$ to be this highest power. In general there will be logarithmic enhancements. But Weinberg's theorem guarantees that the power $\delta_M(\Gamma)$ is correctly given by considering only the regions we have listed. The graphs treated in Section 8.2 provide examples of this procedure. (The subscript 'M' is to distinguish $\delta_M(\Gamma)$ from the ultra-violet degree of divergence.)

A region contributing to a leading power that is M^0 or bigger is symbolized by contracting to a point both the heavy lines and the lines carrying large momentum. The points represent vertices in the effective low-energy theory; we have already used this notation in Section 8.2, in the

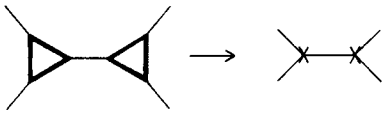


Fig. 8.3.1. Graph with two contracted subgraphs.

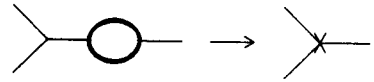


Fig. 8.3.2. One-particle reducible subgraphs may have to be contracted.

figures. In general (see Fig. 8.3.1) the contractions will result in several vertices. We include in our definition the restriction that a subgraph is only contracted to a point if it contains at least one heavy line. A contracted subgraph is 1PI in the light lines; for if it can be split into two parts by cutting a light line then that line is not carrying a large loop momentum. However, the contracted graph may be 1PR in the heavy lines. For example, in a theory where the symmetry $\phi_h \to -\phi_h$ is not valid, a graph like Fig. 8.3.2 gives a leading power M^0; the self-energy gives a power M^2 which cancels the $1/M^2$ in the propagator.

A subgraph that is contracted to a single vertex gives the same power of M as its UV power-counting. So

$$\dim(\text{subgraph}) = \text{power of } M + \dim(\text{couplings}).$$

Hence in a renormalizable theory (where couplings have non-negative dimension) the only contracted graphs that have a non-vanishing value as $M \to \infty$ correspond to vertices whose couplings have non-negative dimension. These vertices give the difference between $g^* z^{3/2}$ and g, etc. Thus the couplings in the effective low-energy theory satisfy the dimensional criterion for renormalizability. In a scalar theory, this implies actual renormalizability, provided all the couplings are used that have non-negative dimension and that obey the symmetries of the full theory.

8.4 Proof of decoupling

8.4.1 *Renormalization prescription R* with manifest decoupling*

Let us work with the theory defined by (8.1.1). We choose to renormalize it according to a mass-independent prescription, which we will denote by a symbol R. For definiteness we choose this to be minimal subtraction. By deleting all heavy fields from (8.1.1) and by changing the values of the couplings we obtain the form of expected low-energy theory (8.1.3). We choose to renormalize the low-energy theory by a mass-independent prescription R^*, which we also take to be minimal subtraction.

Our proof will consist of extending R^* to a renormalization of the full

theory. The extension will satisfy manifest decoupling. Since different renormalization prescriptions differ only by a reparametrization, the statement (8.1.4) of the decoupling theorem will hold. The structure of R^* in the full theory will give a 'mass-independent' form for g^*, m^{*2} and z:

$$\left.\begin{aligned} g^* &= g^*(g_1, g_2, M/\mu), \\ z &= z(g_1, g_2, M/\mu), \\ m^{*2} &= m^2 z_m(g_1, g_2, M/\mu) + M^2 z_{mM}(g_1, g_2, M/\mu). \end{aligned}\right\} \tag{8.4.1}$$

Mass independence means independence of the light mass. As before, to save notational complication we choose to renormalize the linear coupling $f\phi_1$ by the prescription that $\langle 0|\phi_1|0\rangle = 0$. We then ignore both the linear coupling and the tadpole graphs.

The reason we use mass-independent renormalization prescriptions for all the couplings other than the term linear in ϕ is that we can thereby make very clear the decoupling of phenomena at small mass scales from large mass scales. In addition, the renormalization-group equations for (8.4.1) are much simpler to work with than they would otherwise be.

It is convenient to define two concepts:

(1) A heavy graph is one containing at least one heavy line (i.e. a line for the heavy field ϕ_h).
(2) A light graph is one that contains no heavy lines.

For each basic graph Γ in the full theory we have a series of counterterm graphs that are used to cancel its divergences. If Γ is a heavy graph, then we also consider its counterterm graphs to be heavy graphs, even though they may contain no explicit heavy lines.

We have chosen a renormalization prescription R^* for the low-energy theory. This defines the renormalized value of any graph in the low-energy theory, and therefore of any light graph in the full theory. We now wish to extend this prescription to heavy graphs, in such a way that it satisfies manifest decoupling. That is, the renormalized value $R^*(\Gamma)$ of a heavy graph goes to zero as $M \to \infty$. The basic idea is to subtract such graphs at zero momentum. That this is a sensible procedure is easily seen by examining a few of the graphs from Section 8.2.

For example, we saw that Fig. 8.2.3 diverges logarithmically when $M \to \infty$, if we use minimal subtraction. But with zero-momentum subtraction we have

$$\begin{aligned} R^*(\Gamma_3) &= \frac{ig_2^3}{64\pi^3}\int_0^1 dx \int_0^{1-x} dy \ln\left[1 - \left(\frac{p_1^2 x}{M^2} + \frac{p_2^2 y}{M^2}\right)(1-x-y) - \frac{p_3^2 xy}{M^2}\right] \\ &= O(p^2/M^2) \quad \text{as } M \to \infty. \end{aligned} \tag{8.4.2}$$

Clearly, the difference between the two renormalizations is just the difference given by (8.2.4) for $g^* z^{3/2} - g$. At high energy we would use minimal subtraction – so M can be neglected compared with momenta – but at low energy we would use the R^* prescription – so that we can simplify calculations by dropping heavy graphs.

Fig. 8.4.1.

Consider next Fig. 8.4.1 for the self-energy of the heavy field. This graph contains both light and heavy lines. It would behave like $M^2 \ln(M)$ for large M, if we used minimal subtraction. Instead, let us define the subtraction by

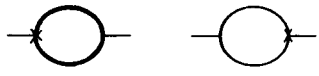
$$\begin{aligned} R^*(\Gamma_{4.1}) = {} & \frac{ig_2^3}{64\pi^3} \frac{\Gamma(2-d/2)}{(4\pi\mu^2)^{d/2-3}} \times \\ & \times \int_0^1 dx \{ [M^2 x + m^2(1-x) - p^2 x(1-x)]^{d/2-2} - (M^2 x)^{d/2-2} \\ & \qquad - (d/2-2)(M^2 x)^{d/2-3} [m^2(1-x) - p^2 x(1-x)] \} \\ & \xrightarrow[(d\to 6)]{} \frac{ig_2^3}{64\pi^3} \int_0^1 dx \left\{ [M^2 x + m^2(1-x) - p^2 x(1-x)] \times \right. \\ & \left. \times \ln\left[1 + \frac{m^2(1-x)}{M^2 x} - \frac{p^2(1-x)}{M^2} \right] - m^2(1-x) + p^2 x(1-x) \right\} \\ & = O(1/M^2) \quad \text{as } M \to \infty . \end{aligned} \tag{8.4.3}$$

Here, we observed that when $M \to \infty$ the dependence of the unrenormalized graph is linear in m^2 and p^2. So we expanded about $p = m = 0$ and subtracted the terms up to quadratic in m and p. This means that the counterterms are polynomial in m, i.e., 'mass-independence' holds good. Normally subtractions at zero mass and momentum have infra-red divergences, but the presence of a heavy line prevents this here.

As a final example let us examine the two-loop graph of Fig. 8.2.6. The unrenormalized graph cannot be expanded about $m = p = 0$ to give a counterterm, because there are two light lines. At $m, p \sim 0$ they give m- and p-dependence of the form

$$\int_{k \sim m \sim 0} d^6 k \frac{(\text{value of heavy loop at } p = k = 0)}{(k^2 - m^2)[(p+k)^2 - m^2]} \sim (p^2 + m^2) \ln(p^2 + m^2). \tag{8.4.4}$$

The right-hand side of this equation is schematic and symbolizes the



(a) (b)

Fig. 8.4.2. Counterterms for Fig. 8.2.6.

maximum powers and logarithms of m and of p that occur. However, to obtain the renormalized value of the graph we must first subtract subdivergences by the counterterms Fig. 8.4.2, constructed by the R^*-scheme. Now, the counterterm graph (*b*) has infra-red behavior exactly equal and opposite to that of (8.4.4), because the counterterm is minus the value of the heavy loop at $p = k = 0$. Thus the sum of the two graphs has the extra convergence we need. The overall counterterm is then linear in p^2 and m^2. There are no logarithms of m as $M \to \infty$.

8.4.2 *Definition of* R^*

To define the renormalization prescription R^* in general, we simply summarize and generalize what we have just done for particular graphs.

We define the renormalization prescription R^* in the full theory to be the same as our chosen prescription for the low-energy theory whenever it acts on a purely light graph. For a heavy graph Γ, we assume inductively that we have defined the quantity $\bar{R}^*(\Gamma)$ in the usual way to be the unrenormalized value of Γ plus counterterms in the R^*-scheme to cancel its subdivergences. If Γ has degree of divergence $\delta(\Gamma) > 0$, then its overall counterterm is defined by subtraction at $m = p = 0$. The renormalized value of Γ is $R^*(\Gamma) = \bar{R}^*(\Gamma) + C^*(\Gamma)$, as usual.

To define $C^*(\Gamma)$ precisely, we first expand $\bar{R}^*(\Gamma)$ in a Taylor series about the point where its external momenta and the light mass m are zero. Pick out the terms where momenta and m^2 occur with dimension up to $\delta(\Gamma)$, and let the counterterm $C^*(\Gamma)$ be the negative of these terms. Our examples tell us to expect that with such a counterterm:

(1) the leading $M \to \infty$ behavior is canceled,
(2) there are no IR singularities in the counterterm.

We must prove these statements in general. The proof will generalize from the simplest non-trivial case, Fig. 8.2.6. There, the UV divergent unrenormalized graph is not polynomial in m and p, but after subtraction of subdivergences by the R^*-scheme, it becomes polynomial. Then the R^*-prescription can legitimately generate the overall counterterm. Moreover, after subtraction of the subdivergences, the leading large-M behavior is also

polynomial in m and p with degree equal to the degree of divergence, so that it is cancelled by the overall counterterm.

Even with the subtractions for subgraphs, there are in general IR singularities in the Taylor expansion of a graph. For example, consider Fig. 8.4.1 and expand its integrand – see (8.4.3) – in powers of m^2 and p^2. All the terms beyond the second give divergences at $x = 0$; it is only the terms needed to cancel the UV divergence that are non-singular.

8.4.3 *IR finiteness of* $C^*(\Gamma)$

Suppose Γ is a heavy graph, 1PI in its light lines. Potential infra-red divergences in $C^*(\Gamma)$ arise when m and the external momenta are made small. They come from regions where some or all of the loop momenta are of order m. The simplest case is where all the internal momenta are of order m.

If Γ were a light graph, we would obtain a contribution of order $m^{\delta(\Gamma)}$, where $\delta(\Gamma)$ is the UV degree of divergence. So let us call $-\delta(\Gamma)$ the canonical IR degree of divergence of Γ. If $\delta(\Gamma) = 0$, this is a logarithmic divergence. If $\delta(\Gamma) > 0$, then the graph is finite as $m \to 0$. But to get the coefficients of the polynomial counterterms we differentiate up to $\delta(\Gamma)$ times with respect to m and the external momenta. The highest terms in the polynomial are therefore always logarithmically IR divergent, for a light graph.

However, Γ is actually a heavy graph. So at least one of its propagators counts as $1/M^2$ instead of $1/m^2$. Thus all the counterterms have an IR finite contribution from this region, where all its loop momenta are small.

This discussion is sufficient for all one-loop graphs. But multi-loop graphs have IR divergences coming from regions where only some loops have small momenta. For example, Fig. 8.2.6 has a divergence from the region where p and k are small, i.e., order m, and l is finite or large. This corresponds to IR degree -2, and is given by (8.4.4). As we saw, the IR divergence is canceled by the graph with a counterterm for the heavy loop.

The general case is that some light lines carry momenta of order m and

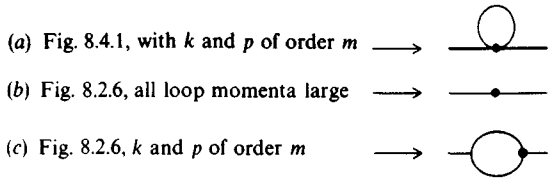


Fig. 8.4.3. Examples of reduced graphs.

the remainder of the lines either are heavy or carry large momentum. Each such region is symbolized by a reduced graph in which the subgraphs consisting of the lines with large momenta and of the heavy lines are contracted to points. Examples of reduced graphs are shown in Fig. 8.4.3. Note:

(1) Counterterm graphs can also have infra-red divergences. The counterterms are inside the vertices of the reduced graphs.
(2) All lines of reduced graphs are light, so at least one vertex of a reduced graph corresponds to a heavy subgraph.

We can write the infra-red degree of divergence for the region corresponding to a particular reduced graph γ as

$$\delta_{\mathrm{IR}}(\Gamma;\gamma) = -\delta(\Gamma) + \sum_{\substack{\text{reduced} \\ \text{vertices } V}} [\delta(V) + \text{IR degree of } R^*(V)]. \qquad (8.4.5)$$

The meaning of this equation can be seen from an example. Consider Fig. 8.2.6 when k and p are of order m. If the graph were purely light, we would have IR degree equal to -2, which is the negative of the UV degree. This would imply that the m^2 and p^2 terms in the expansion about $m = p = 0$ would be divergent. However, the single reduced vertex – as illustrated in Fig. 8.4.3(c) – has a counterterm. This counterterm ensures that the vertex's value is of order m^2/M^2 instead of m^0. The IR degree for the whole graph is thereby decreased by 2. The second term in (8.4.5), where the sum is over this single vertex, indicates this reduction. The degree for the region is then -4; we can therefore expand up to order m^2 and p^2 without an infra-red divergence. The terms of order m^4, p^4, etc., are infra-red divergent, but they are not needed for ultra-violet renormalization.

In the general case of (8.4.5), each reduced vertex V would contribute $-\delta(V)$ if it were light and all its internal lines had momenta of order m. But it actually contributes what we will now prove is a smaller amount. Remember that counterterm graphs also contribute, and we assume that counterterm vertices are included inside reduced vertices. The IR degree of $R^*(V)$ is its power as its external momenta are scaled like m. The possible cases for V are:

(1) If V is overall convergent and contains a heavy line, then its infra-red degree is greater than its ultra-violet degree. Fig. 8.4.3(a) has a vertex with UV degree -2 and IR degree zero.
(2) If V is overall divergent and contains a heavy line then ordinarily we would expect it to behave as m^0 when $m \to 0$ with fixed M. But we make

subtractions by the R^* scheme so that its behavior is actually $m^{\delta(V)+2}$. By induction we may assume its subtractions have no IR divergence.

Hence, in every region of momenta a heavy graph Γ always has at least one mechanism to reduce its IR degree below $-\delta(\Gamma)$ and none to increase it. Thus the overall counterterm $C^*(\Gamma)$ is IR finite. It is crucial to our inductive proof that we first subtract subdivergences by the R^* scheme.

8.4.4 *Manifest decoupling for R^**

A purely light graph is a graph in both the full theory and in the low-energy theory. It survives unaltered when we let $M \to \infty$. We will now prove that all the heavy graphs vanish when $M \to \infty$, given that we renormalize them by the R^* scheme.

To do this, decompose each heavy graph into its skeleton, i.e., a series of 1PI graphs connected by lines that are not part of any loop. Since a heavy line that is outside a loop vanishes as $M \to \infty$, all heavy graphs vanish as $M \to \infty$, if the 1PI graphs vanish.

The $M \to \infty$ limit of a 1PI graph can be related to an IR limit by scaling all masses and momenta:

$$M \to 1, \quad p \to p/M, \quad m \to m/M.$$

Then

$$\Gamma(p, m, M) = M^{d(\Gamma)}\Gamma(p/M, m/M, 1), \tag{8.4.6}$$

where $d(\Gamma)$ is the dimension of Γ. So Γ vanishes as $M \to \infty$ provided the infra-red behavior is less singular than $m^{-d(\Gamma)}$. But this is what we showed in the proof of IR finiteness of the counterterms. (Note that the dimension of a graph is greater than or equal to its UV degree of divergence.)

8.4.5 *Decoupling theorem*

We have constructed two renormalization prescriptions, labelled R and R^*, for the theory under consideration. The Green's functions in the schemes R and R^* are equal provided we make appropriate changes in the parameters:

$$\left.\begin{aligned} g_l &\to g^*, \\ g_h &\to g_h^*, \\ \text{coefficient of } \partial\phi_l^2/2 &\to z, \\ \text{coefficient of } \partial\phi_h^2/2 &\to z_h, \\ m^2 &\to m^{*2}, \\ M^2 &\to M^{*2}, \end{aligned}\right\} \tag{8.4.7}$$

This is just a particular case of a renormalization-group transformation, and is proved by Section 7.2. When $M \to \infty$ we may drop all heavy graphs in the R^* scheme (so also g_h^*, z_h, M^* drop out of consideration). This then gives (8.1.4), which is the decoupling theorem.

Mass-independence is true because we have arranged all counterterms to be polynomials in the light mass of the appropriate degree.

8.5 Renormalization-group analysis

When one computes a graph containing lines for fields with widely different masses, one finds, in general, that its value gets large as a power of the logarithm of the mass ratio. Such large coefficients are undesirable in a perturbation expansion, for they mean that the reliability of using a few low-order terms is worsened. This situation arises in both strong- and weak-interaction physics. We will now show how to combine the decoupling theorem and the renormalization group to do calculations without their being made unreliable by the large logarithms.

A convenient method is to use a mass-independent scheme (specifically minimal subtraction) for high-momentum calculations, where one often wishes to neglect all masses, and to use the R^* scheme, as defined in Section 8.4, at low momenta, where one wishes to neglect heavy graphs. An advantage of this method is a simplification of many of the calculations needed to match high-energy and low-energy calculations. One needs only the pole parts of graphs and the values at zero external momentum.

We will explain how to use this scheme in the toy theory (8.1.1). First let us write the RG equations for the Green's functions. For a Green's function of N_l light and N_h heavy fields, we have

$$\left(\mu\frac{\mathrm{d}}{\mathrm{d}\mu} + \tfrac{1}{2}N_l\gamma_l + \tfrac{1}{2}N_h\gamma_h\right)G_{N_l,N_h} = 0, \tag{8.5.1}$$

where

$$\mu\frac{\mathrm{d}}{\mathrm{d}\mu} = \mu\frac{\partial}{\partial\mu} + \beta_1(g_1,g_2)\frac{\partial}{\partial g_1} + \beta_2(g_1,g_2)\frac{\partial}{\partial g_2} - (M^2\gamma_M + m^2\gamma_{mM})\frac{\partial}{\partial M^2} - (m^2\gamma_m + M^2\gamma_{Mm})\frac{\partial}{\partial m^2}. \tag{8.5.2}$$

The RG coefficients are obtained from the renormalization counterterms as usual. Their lowest-order values are

$$\beta_1 = -\frac{1}{128\pi^3}(\tfrac{3}{2}g_1^3 + g_1g_2^2 + g_2^3) + \cdots,$$

$$\beta_2 = -\frac{1}{128\pi^3}(\tfrac{7}{4}g_2^3 + g_1 g_2^2 - \tfrac{1}{6}g_1^2 g_2) + \cdots,$$

$$\gamma_l = \frac{1}{192\pi^3}g_2^2 + \cdots,$$

$$\gamma_h = \frac{1}{384\pi^3}(g_1^2 + g_2^2)\ldots. \tag{8.5.3}$$

In the effective low-energy theory, the RG equation is

$$\left[\mu\frac{d^*}{d^*\mu} + \tfrac{1}{2}N\gamma^*\right]G_N^* = 0, \tag{8.5.4}$$

with

$$\left.\begin{aligned} \mu\frac{d^*}{d^*\mu} &= \mu\frac{\partial}{\partial\mu} + \beta^*\frac{\partial}{\partial g^*} - \gamma^* m^{*2}\frac{\partial}{\partial m^{*2}}, \\ \beta^* &= -\frac{3}{464\pi^3}g^{*3} + \cdots, \\ \gamma^* &= \frac{g^{*2}}{384\pi^3} + \cdots. \end{aligned}\right\} \tag{8.5.5}$$

To compare the low-energy theory and the full theory, we extended the renormalization scheme of the low-energy theory to a renormalization scheme R^* for the full theory. In this scheme the RG operator has the form

$$\mu\frac{d^*}{d^*\mu} = \mu\frac{\partial}{\partial\mu} + \beta^*\frac{\partial}{\partial g^*} + \beta_2^*\frac{\partial}{\partial g_2^*} - (M^{*2}\gamma_M^* + m^{*2}\gamma_{mM}^*)\frac{\partial}{\partial M^{*2}} - (m^{*2}\gamma_m^* + M^{*2}\gamma_{Mm}^*)\frac{\partial}{\partial m^{*2}}, \tag{8.5.6}$$

and the anomalous dimensions of the fields are γ^* and γ_h^*. In fact β^*, γ_m^*, and γ^* are identical to those in the low-energy theory (see (8.5.4) and (8.5.5)), while $\beta_2^* = \gamma_{mM}^* = \gamma_{Mm}^* = \gamma_M^* = \gamma_h^* = 0$. This is easily seen by examining the Green's functions which provide the normalization conditions for the renormalizations.

For example, consider the inverse of the heavy propagator when both p^2 and m^2 are much less than M^2:

$$1/G_{0,2} = -\mathrm{i}[p^2 - M^{*2} + O(p^4, m^2p^2, m^4)] \tag{8.5.7}$$

which satisfies

$$\mu\frac{d^*}{d^*\mu}G_{0,2}^{-1} = \gamma_h^* G_{0,2}^{-1}. \tag{8.5.8}$$

This is only consistent if $\gamma_h^* = \gamma_M^* = \gamma_{mM}^* = 0$.

8.5.1 *Sample calculation*

We wish to start with the full theory renormalized by minimal subtraction. In that version of the theory, we know the evolution of the couplings. Our aim is to compute Green's functions in the low-energy theory and the values of the mass and coupling. The low-energy effective couplings are

$$\left.\begin{aligned} g^{*2}(\mu) &= \frac{128\pi^3}{3\ln(\mu/\Lambda^*)} + \cdots, \\ g_2^* &= \text{fixed}, \\ m^{*2} &= \text{constant}\,[\ln(\mu/\Lambda^*)]^{-1/9}, \\ M^{*2} &= \text{fixed}. \end{aligned}\right\} \tag{8.5.9}$$

The effective couplings for the full theory with minimal subtraction are more complicated because they solve a coupled equation for two variables.

To make the transition between the schemes we compute the lowest-order divergent graphs. We equate the self-energy for ϕ_1 in the two schemes, with use of the Lagrangian (8.1.3) for the low-energy theory. This gives

Fig. 8.2.4 + pole counterterm

= Fig. 8.2.4 + zero-momentum counterterm + $i(z-1)p^2 - i(m^{*2}z - m^2)$.

We thus obtain z and m^{*2} as given by (8.2.6) and (8.2.7). To keep m^{*2} finite, and not of order M^2, we must replace m^2 by

$$m^2 + \frac{1}{128\pi^3} M^2 g_2^2 [\gamma - 1 + \ln(M^2/4\pi\mu^2)]. \tag{8.5.10}$$

Notice the presence of logarithms of M/μ. If they are large enough, they invalidate the use of perturbation theory to compute g^*, m^{*2}, and z. However, the equations we write are valid at any value of μ, so we may perform the calculations with μ of order M. After computing g^*, m^*, and z in terms of g_1, g_2, m, and M, we can evolve them to the value of μ that we wish to use for calculations in the low-energy theory.

A convenient point to do the matching is where $g_1 = g^*$, i.e., at $\mu^2 = \mu_0^2$, where $\mu_0^2 = M^2 e^{\gamma}/4\pi$. Then for a general value of μ we have

$$g^{*2}(\mu) = \frac{1}{1/g_1^2(\mu_0) + (3/128\pi^3)\ln(\mu/\mu_0) + \cdots}. \tag{8.5.11a}$$

A similiar equation holds for m^*:

$$m^{*2}(\mu) = m^2[(128\pi^3/3g_1^2) + \ln(\mu/\mu_0) + \cdots]^{-1/9}. \tag{8.5.11b}$$

The solution for z is more complicated since the renormalization-group equation for both renormalization prescriptions is needed.

8.5.2 *Accuracy*

We compute $g^*(\mu)$ in the low-energy theory by matching to the full theory at some μ_0 of order M and then evolving to an arbitrary renormalization mass μ from μ_0. Given the accuracy in $g^*(\mu)$ that we need for a particular calculation, we will find the order to which we must perform the matching and to which we must know β.

The RG equation for $g^*(\mu)$ gives

$$\ln(\mu/\mu_0) = \int_{g(\mu_0)}^{g(\mu)} \mathrm{d}g'/\beta(g'). \tag{8.5.12}$$

So if there are small errors $\Delta g(\mu_0)$ and $\Delta(1/\beta)$ in $g(\mu_0)$ and $1/\beta$ then the error in $g(\mu)$ is

$$\begin{aligned}\Delta g(\mu) &\sim \beta[g(\mu)]\left\{\frac{\Delta g(\mu_0)}{\beta[g(\mu_0)]} - \int \mathrm{d}g'\,\Delta\left[\frac{1}{\beta(g')}\right]\right\} \\ &= O[g(\mu)^3]\left\{\frac{\Delta g(\mu_0)}{O[g(\mu_0)^3]} - \int \mathrm{d}g'\,\Delta\left[\frac{1}{\beta(g')}\right]\right\}.\end{aligned}$$

Suppose we perform matching up to n_m-loop order; then the error in $g(\mu_0)$ is of order $g(\mu_0)^{2n_m+3}$. Suppose β is computed to n_β-loop order; then the error in $1/\beta(g)$ is of order $g^{2n_\beta-3}$. These translate to errors in $g(\mu)$ of order

$$g(\mu)^3 g(\mu_0)^{2n_m}$$

and

$$g(\mu)^3 \ln[g(\mu)/g(\mu_0)] \text{ if } n_\beta = 1,$$

or

$$O[g(\mu)^{1+2n_\beta}] + O[g(\mu)^3 g(\mu_0)^{2(n_\beta-1)}] \text{ if } n_\beta \geq 2.$$

For example, if we wish to perform reliable two-loop calculations, then we need Δg to be much smaller than g^3. This means that we need to do the matching correct to one loop and that the β-function is needed to two loops. This is the minimum accuracy needed to correspond to a fractional error on Λ (defined in Chapter 7) which is much less than unity.