Chapter 12
Effective Field Theories and Naturalness

Effective field theories are among the most powerful instruments in the toolbox of contemporary physics. Although the concept of effective field theory has been already discussed in Chap. 8, here we are going to provide a relatively elementary description of the relevant technology. Although rather unrealistic, the examples of effective field theories studied next serve the purpose of illustrating the relevant physics involved. The chapter will be closed with a discussion of the concept of naturalness, which plays a central role in modern particle physics. The reader is advised not to be scared by the technicalities of the Feynman diagram computations contained in the chapter. Most of the conclusions can be reached without caring too much about the precise value of the numerical prefactors.

12.1 Energy Scales in Quantum Field Theory

When introducing the renormalization group in Chap. 9 we did not go into a detailed evaluation of Feynman diagrams. In the present chapter we will provide some more details, particularly the computation of loop diagrams and how to extract divergences from them. The interpretation of the results coming from particle colliders like the Tevatron at Fermilab or the LHC at CERN requires extensive loop computations in the standard model. Simple cut-off methods or other regularization schemes are inadequate for this task.

The calculation of quantum corrections is necessary also when trying to derive effective field theories valid at low energies from more fundamental descriptions at high energies. This raises the important question of the separation of scales in quantum field theory. In the standard model, for example, we have widely separated energy scales. We have the mass of the $W^{\pm}$ boson $m_W \sim 10^2$ GeV characterizing the scale of the electroweak processes. At the same time there is another energy scale that can be constructed from the three fundamental constants of physics, Planck's constant $\hbar$, the speed of light c and Newton's constant G_N. This is the Planck mass defined by

L. Álvarez-Gaumé and M. Á. Vázquez-Mozo, *An Invitation to Quantum Field Theory*,
Lecture Notes in Physics 839, DOI: 10.1007/978-3-642-23728-7_12,

$$M_P = \sqrt{\frac{\hbar c}{G_N}} \sim 10^{19}\,\text{GeV}. \tag{12.1}$$

The theory might have additional "intermediate" high energy scales, such as the grand unification (GUT) energy at about 10^{15}–10^{16} GeV.

The important question to be addressed is how physical processes at energies below the lower scale, $E < m_W$, depend on higher scales such as the Planck or the GUT scale. The analysis of this question often invokes the *naturalness criterion* to be explained later. It should be stressed, however, that naturalness rather than a law of Nature should be seen as a good guiding principle to understand the sizes of the dimensionful parameters, specially masses, in the low-energy theory.

Later on we will argue that the fact that light fermions, such as the electron or the muon, have masses much smaller than the Planck scale should not be considered unnatural. The situation is different for scalar particles, such as the so-far hypothetical Higgs boson. In this case the fact that the Higgs mass is expected to be light with respect to other higher energy scales is quite unnatural. To see this let us take the point of view that the standard model is a valid physical description up to some energy scale Λ, that we can take to be the Planck, GUT or any other relevant scale. Identifying then Λ as the momentum cutoff in the quantum theory, one finds that the correction to the Higgs mass depends quadratically on it

$$\delta m^2 \sim \frac{g^2}{16\pi^2}\Lambda^2. \tag{12.2}$$

The conclusion is that the Higgs mass should be very sensitive to the details of the physics at higher energies since it strongly depends on the energy scale Λ at which the standard model should be replaced by a more complete description. This simple remark is the kernel of the famous *hierarchy problem*. Due to the strong dependence of the Higgs mass corrections on the scale of new physics, keeping m_{H} and Λ widely apart

$$m_{\text{H}}^2 = m_0^2 + \delta m^2 \ll \Lambda^2 \tag{12.3}$$

requires a fantastic fine tuning of m_0^2. This is what is meant by the statement that a light Higgs mass is unnatural.

The questions we just introduced will be formulated in more detail later on. We still do not have answers to many questions about naturalness. Their resolution will bring deep breakthroughs in our understanding of Nature at short and long distances.

12.2 Dimensional Regularization

To get started we present the method of dimensional regularization and renormalization that is universally used specially in theories with local gauge symmetries like the standard model. Introduced in the early seventies [1, 2], it is a remarkably clever way to regulate the integrals appearing in the perturbative calculation of scattering amplitudes.

In Chap. 8 we learned in a particular example how the calculation of Feynman diagrams with loops leads to divergent integrals over the momenta running in them, that should be regularized. In all cases the integrand has the structure of a rational function of the loop momenta. Dimensional regularization (DR) prescribes these integrals to be carried out in an arbitary dimension d instead of $d = 4$. The integration gives a finite result depending on d that can then be analytically continued to complex values of the dimension. The divergence of the original loop integral reappears as poles of different order in $d - 4$. This is a rather economical and efficient way of regularizing divergences.

Apart from its computational simplicity DR presents various advantages. The most important is that it automatically preserves the symmetries of the theory whenever they admit an extension to higher dimensions. This is the case of vector-like gauge theories like QCD where the gauge symmetry can be formulated in any dimension. This means that DR regularizes the theory without breaking gauge invariance.

Chiral (gauge) symmetries, on the other hand, are more problematic. The reason is that the notion of chirality is very "four-dimensional", and chiral gauge theories like the electroweak sector of the standard model require a very special treatment. The fact that chiral symmetries do not admit an extension to higher dimensions is related to the existence of the chiral anomaly studied in Chap. 9. The fact that DR frequently preserves the original symmetry is a big advantage with respect to other regularization procedures. Among its few drawbacks one should mention that so far all attempts to find a nonperturbative formulation of DR have been unsuccessful.

Since our presentation only requires the calculation of a few one-loop diagrams, we illustrate the use of DR with the basic integral

$$I_n(d, m^2) = \int \frac{d^d p}{(2\pi)^d} \frac{1}{(p^2 - m^2 + i\varepsilon)^n}. \tag{12.4}$$

For $d = 4$ this integral is divergent when $n = 0, 1, 2$. To evaluate it we begin with the integration over p^0. Due to the $i\varepsilon$ prescription, the integrand

$$\frac{1}{(p^2 - m^2 + i\varepsilon)^n} = \frac{1}{[(p^0)^2 - E_{\mathbf{p}}^2 + i\varepsilon]^n} \tag{12.5}$$

has poles located just above and below the real p^0 axis with real parts $\pm E_{\mathbf{p}}$. The integral over the real axis can be carried out applying Cauchy's theorem to the contour shown in Fig. 12.1. Since the contour encloses no poles and the integrand vanishes as $|p^0| \to \infty$, the integration over real p^0 can be expressed as the integral over the imaginary axis, namely

$$\int_{-\infty}^{\infty} \frac{dp^0}{2\pi} \frac{1}{[(p^0)^2 - E_{\mathbf{p}}^2 + i\varepsilon]^n} = i(-1)^n \int_{-\infty}^{\infty} \frac{dp_E^0}{2\pi} \frac{1}{[(p_E^0)^2 + E_{\mathbf{p}}^2]^n}. \tag{12.6}$$

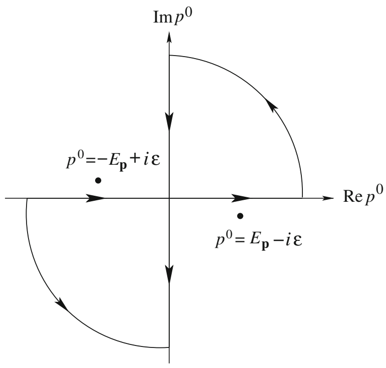


Fig. 12.1 Contour of integration used to evaluate the integral $I_n(d,m^2)$. The absence of poles in the first and third quadrants allows the integration contour along the real axis to be rotated to the imaginary axis (Wick rotation)

For $m \neq 0$ the poles are away from the imaginary axis and the $i\varepsilon$ prescription drops out of the second integral. Hence, Eq. (12.4) becomes

$$I_n(d,m^2) = i(-1)^n \int \frac{d^d p_E}{(2\pi)^d} \frac{1}{(p_E^2+m^2)^n}, \tag{12.7}$$

where p_E is an Euclidean d-dimensional momentum, $p_E^2 = (p_E^0)^2 + \mathbf{p}^2$.

To compute of the Euclidean integral (12.7) one exponentiates the integrand with the help of the identity

$$\frac{1}{a^n} = \frac{1}{\Gamma(n)} \int_0^\infty dt\; t^{n-1} e^{-az} \quad (a>0), \tag{12.8}$$

to write

$$\int \frac{d^d p_E}{(2\pi)^d} \frac{1}{(p_E^2+m^2)^n} = \frac{1}{\Gamma(n)} \int_0^\infty dt\; t^{n-1} e^{-tm^2} \int \frac{d^d p_E}{(2\pi)^d} e^{-tp_E^2}. \tag{12.9}$$

The Euclidean character of p_E ensures the convergence of the momentum integral. Using

$$\int \frac{d^d p_E}{(2\pi)^d} e^{-p_E^2} = \pi^{\frac{d}{2}}, \tag{12.10}$$

we arrive at the final result for the integral (12.7)

$$I_n(d,m^2) = \frac{i(-1)^n}{(4\pi)^{2+\frac{d-4}{2}}} \frac{\Gamma(n-2-\frac{d-4}{2})}{\Gamma(n)(m^2)^{n-2-\frac{d-4}{2}}}. \tag{12.11}$$

The dependence of $I_n(d,m^2)$ on the dimension can be analytically continued to complex values of d. The pole structure is derived from the properties of the Euler gamma function $\Gamma(z)$, which can be found in any book on special functions. Here we just notice that $\Gamma(z)$ has poles at nonpositive integer values of the argument, $z = 0, -1, -2, \ldots$ In the case at hand, for $n > 2$ the integral converges as $d \to 4$. When $n = 1, 2$, on the contrary, the expression contains either a factor of $\Gamma(1 - d/2)$ or $\Gamma(2 - d/4)$ in the numerator, both functions diverging in the limit $d \to 4$.

To find the behavior of $I_n(d, m^2)$ around $d = 4$ we use the Laurent expansion of the gamma function around its poles

$$\Gamma(-k + \varepsilon) = \frac{(-1)^k}{k!}\left[\frac{1}{\varepsilon} + \psi(k+1) + \mathscr{O}(\varepsilon)\right] \quad k \in \mathbb{N}, \tag{12.12}$$

where $\psi(z)$ is the dilogarigthm function, defined as the logarithmic derivative of the gamma function

$$\psi(z) = \frac{d}{dz}\log\Gamma(z), \quad \psi(k+1) = -\gamma + \sum_{n=1}^{k}\frac{1}{n}, \tag{12.13}$$

with $\gamma = -\psi(1) = 0.5772\ldots$ the Euler-Mascheroni constant. Applying this to our integral we find its behavior as $d \to 4$ to be

$$I_n(d, m^2) \stackrel{d\to 4}{\longrightarrow} -\frac{i(m^2)^{2-n}}{16\pi^2}\frac{2}{d-4} + \text{finite part}, \quad n = 1, 2. \tag{12.14}$$

Other integrals can be computed along similar lines. One of special interest is

$$I_n(d, m^2, q) = \int \frac{d^d p}{(2\pi)^d}\frac{1}{(p^2 + 2p\cdot q - m^2 + i\varepsilon)^n}. \tag{12.15}$$

Upon completing the square in the denominator and shifting the integration variable (a legitimate step once the integral is regularized), it reduces to an integral of the type (12.7)

$$I_n(d, m^2, q) = I_n(d, m^2 + q^2). \tag{12.16}$$

In the case of the integrals arising in the calculation of diagrams with higher loops we would find a collection of gamma functions that would produce poles at $d = 4$ of various orders

$$\frac{1}{(d-4)^L}, \frac{1}{(d-4)^{L-1}}, \ldots, \frac{1}{d-4}, \tag{12.17}$$

with L the number of loops in the diagram. It is quite remarkable in fact that the second and higher order poles are determined by the first order poles as we will show later on. This provides a very useful check on multi-loop computations.

The analytical continuation of the dimension in the integrals can be formulated in an axiomatic way. All we need is to define the operation of d-dimensional integration (with d complex) preserving the basic properties of multi-dimensional integrals, namely

$$\begin{aligned}\int d^d p\Big[f(p)+g(p)\Big] &= \int d^d p f(p) + \int d^d p g(p),\\ \int d^d p f(\lambda p) &= \lambda^{-d}\int d^d p f(p) \quad (\text{with } \lambda\in\mathbb{C}),\\ \int d^d p f(p+k) &= \int d^d p f(p).\end{aligned} \tag{12.18}$$

Since rational functions can be turned into exponentials by using identities like (12.8), the evaluation of any integral appearing in the computation of a Feynman diagram can be done using the previous properties together with the action of the d-dimensional integration on Gaussian functions, given by Eq. (12.10).

A simple perusal of the DR calculation carried out above shows that both $I_1(d, m^2)$ and $I_2(d, m^2)$ diverge exactly in the same way as $d \to 4$, namely with a simple pole. The conclusion is that DR is only sensitive to logarithmic divergences and all polynomial divergences are regularized to zero. As a matter of fact, in DR we can write

$$\int d^d p (p^2)^n = 0. \tag{12.19}$$

One way to see how this identity comes about is to consider $I_n(d, m^2)$ with $n \geq -1$ and take the limit $m \to 0$ in the region $\mathrm{Re}\, d \geq 2n$ and away from the poles of the gamma function in (12.11). The limit is then equal to zero and the result can be analytically continued to the whole complex d plane. DR is also useful in handling infrared divergences.

The fact that DR eliminates quadratic divergences might seem surprising in the light of the previous discussion of the hierarchy problem. Indeed, as DR regularizes the quadratic divergences to zero it seems that the whole hierarchy problem results from using a clumsy regulator, and that by using DR we could shield the Higgs mass from the scale of new physics. This is not the case, but for interesting reasons. In spite of DR the Higgs mass is still sensitive to high energy scales. If it is ever found with a low mass, we will also get relevant information on what shields its mass from the higher scales. Before explaining the interesting reasons, we need to develop more theory.

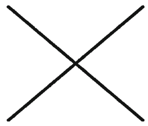
12.3 The ϕ^4 Theory: A Case Study

To get a better understanding of how a quantum field theory is regularized using DR we look into a very simple field theory: a massive real scalar field $\phi(x)$ with a quartic self-interaction

$$\mathscr{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{12.20}$$

As we learned in Chap. 6 the perturbative expansion is constructed using the Feynman rules. In this case we only have to specify one propagator and one vertex

$$\Longrightarrow \quad \frac{i}{p^2 - m^2 + i\varepsilon}$$



$$\Longrightarrow \quad -i\lambda$$

together with the delta function conservation $(2\pi)^4\delta^{(4)}(p_1 + p_2 + p_3 + p_4)$, where we use the convention that all momenta in the vertex are incoming. Since the scalar field is real, it does not carry charge and therefore the lines of the Feynman diagrams do not have orientation.

The quantization using DR requires defining the theory in d dimensions

$$S = \int d^d x \mathscr{L}(\phi, \partial_\mu\phi). \tag{12.21}$$

Since the dimensions of the fields and parameters in the action depend on d, it is useful to stop for a moment and carry out some dimensional analysis. In natural units $\hbar = c = 1$ the action is dimensionless and looking at the kinetic term we can fix the energy dimensions of the scalar field[1]

$$D_\phi = \frac{d-2}{2}. \tag{12.22}$$

The same analysis can be done for fermions and gauge fields with the respective result

$$D_\psi = \frac{d-1}{2}, \quad D_A = \frac{d-2}{2}. \tag{12.23}$$

The energy dimensions of the parameter of the scalar theory (12.20) are

$$D_m = 1, \quad D_\lambda = 4 - d. \tag{12.24}$$

In the case of scalar field theories with cubic self-interaction and/or Yukawa couplings to Dirac fermions, the dimension of the corresponding coupling constants are

$$\begin{aligned} \lambda'\phi^3 \quad &\Longrightarrow \quad D_{\lambda'} = 1 + \frac{4-d}{2} \\ g\phi\overline{\psi}\psi \quad &\Longrightarrow \quad D_g = \frac{4-d}{2} \end{aligned} \tag{12.25}$$

[1] Our choice of natural units allows us to specify the dimensions of all quantities in terms of powers of energy. Thus, for the coordinates we have $[x^\mu] = E^{-1}$, which we denote by $D_x = -1$.

In the particular case of the ϕ^4 example the dependence of the energy dimension of λ with d suggests replacing the coupling constant in the action (and therefore in the Feynman rules) by

$$\lambda \longrightarrow \mu^{4-d}\lambda, \tag{12.26}$$

where μ is an arbitrary energy scale. What we achieve with this is that λ is kept dimensionless for any value of d. In a theory with several couplings we would do the same with all of them using always the same scale μ.

We apply DR to the one loop renormalization of the ϕ^4 theory using the method of renormalized perturbation theory outlined in Sect. 8.3. The aim is to compute the renormalized Lagrangian

$$\mathscr{L}_{\text{ren}} = \frac{1}{2}\partial_\mu\phi_0\partial^\mu\phi_0 - \frac{m_0^2}{2}\phi_0^2 - \frac{\lambda_0}{4!}\phi_0^4, \tag{12.27}$$

depending on the bare parameters and fields. This can be written as $\mathscr{L}_{\text{ren}} = \mathscr{L} + \mathscr{L}_{\text{ct}}$ where

$$\mathscr{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\mu^{4-d}\phi^4 \tag{12.28}$$

depends only on renormalized couplings and fields and the counterterms have the structure

$$\mathscr{L}_{\text{ct}} = \frac{1}{2}A(d-4)\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2 B(d-4)\phi^2 - \frac{\lambda}{4!}\mu^{4-d}C(d-4)\phi^4. \tag{12.29}$$

The functions $A(d-4)$, $B(d-4)$ and $C(d-4)$ contain all the dependence of $\mathscr{L}_{\text{ren}}$ on the regulator d and are related to the bare quantities by

$$\begin{aligned} \phi_0(x) &\equiv \sqrt{Z_\phi(d-4)}\phi(x) = \sqrt{1+A(d-4)}\phi(x), \\ m_0^2(d-4) &= m^2\frac{1+B(d-4)}{1+A(d-4)}, \\ \lambda_0(d-4) &= \lambda\mu^{4-d}\frac{1+C(d-4)}{[1+A(d-4)]^2}. \end{aligned} \tag{12.30}$$

The time-ordered Green's functions of the renormalized fields computed using the Lagrangian $\mathscr{L}_{\text{ren}}$ are finite in the limit $d \to 4$. The field $\phi(x)$ interpolates between the vacuum and the one particle states and therefore the scattering amplitudes computed in terms of these Green's functions are finite as well. The renormalization conditions are then used to express the renormalized mass m and coupling constant λ in terms of measurable quantities.

We see now how this program is implemented. The first divergent Feynman diagram appears in the one-loop calculation of the two-point function

$$= \frac{1}{2}\lambda\mu^{4-d}\int\frac{d^dp}{(2\pi)^d}\frac{1}{p^2-m^2+i\varepsilon}$$

$$= \frac{1}{2}\lambda\mu^{4-d}I_1(d,m^2). \tag{12.31}$$

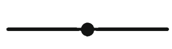
The factor of $\frac{1}{2}$ is a symmetry factor. We can take advantage of the calculations made in the previous section to isolate the divergent part of the diagram as $d \to 4$

$$= -i\frac{\lambda m^2}{16\pi^2}\frac{1}{d-4} + \text{finite part.} \tag{12.32}$$

To cancel this divergence we add a counterterm $-\frac{1}{2}\delta m^2\phi^2$ to the Lagrangian density where δm^2 is given by

$$\delta m^2 = -\frac{\lambda m^2}{16\pi^2}\frac{1}{d-4}. \tag{12.33}$$

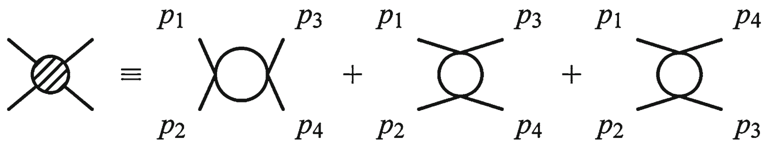
Adding this counterterm means to include in the Feynman rules a new vertex with two external legs



$$= -i\delta m^2(2\pi)^4\delta^{(4)}(p_1+p_2). \tag{12.34}$$

Its contribution to the two point function to order λ exactly cancels the divergent part of the one loop diagram (12.31). There is of course an ambiguity in the definition of the counterterm because in addition to the pole we could also have subtracted a finite part. For the time being, however, we choose not to do so.

The next divergent diagram in the ϕ^4 theory comes from the one-loop calculation of the four-point function. In fact there are three diagrams contributing at order λ^2



(12.35)

The last two diagrams differ in a permutation of the momenta p_3 and p_4. Since the corresponding legs are attached to different vertices the two diagrams are topologically nonequivalent. Applying the Feynman rules listed above, we find that the contribution of these three diagrams can be written as

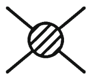
$$= \frac{\lambda^2}{2}\int\frac{d^dk}{(2\pi)^d}\frac{1}{k^2-m^2+i\varepsilon}\left[\frac{1}{(k+p_1+p_2)^2-m^2+i\varepsilon}\right.$$
$$\left.+\frac{1}{(k+p_1+p_3)^2-m^2+i\varepsilon}+\frac{1}{(k+p_1+p_4)^2-m^2+i\varepsilon}\right] \tag{12.36}$$

The $\frac{1}{2}$ in front of the integral is again a combinatorial factor associated with the symmetry of each of the three diagrams.

A look at the integrals (12.36) shows that for $d = 4$ they diverge logarithmically. To extract their divergent part we begin by exploiting the identity

$$\frac{1}{a_1 a_2} = \int_0^1 \frac{dx}{[x a_1 + (1-x) a_2]^2}, \tag{12.37}$$

where x is called a Feynman parameter. This reduces the contribution of the three diagrams to a combination of integrals of the type $I_2(d, m^2, q)$ computed above. Using the expansion of the integrals around $d = 4$ one arrives at the result

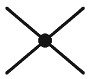


$$= \mu^{4-d} \frac{3i\lambda^2}{16\pi^2}\left(-\frac{2}{d-4} + \text{finite part}\right). \tag{12.38}$$

As with the two-point function, we can remove the divergence (12.38) by adding a counterterm $-\frac{\delta\lambda}{4!}\mu^{4-d}\phi^4$ to the Lagrangian density of the theory with

$$\delta\lambda = -\frac{3\lambda^2}{16\pi^2}\frac{1}{d-4}. \tag{12.39}$$

We can also incorporate this couterterm in the Feynman rules by adding the new vertex



$$= -i\delta\lambda(2\pi)^4\delta^{(4)}(p_1+p_2+p_3+p_4). \tag{12.40}$$

We have computed the only two one-loop 1PI diagrams that are divergent. It is not difficult to show that the one-loop contribution to the six-point function is finite. Thus, our calculation determines, at one loop order, the functions appearing in the counterterm Lagrangian (12.29)

$$\begin{aligned} A(d-4)_{1-\text{loop}} &= 0, \\ B(d-4)_{1-\text{loop}} &= -\frac{\lambda}{16\pi^2}\frac{1}{d-4}, \\ C(d-4)_{1-\text{loop}} &= -\frac{3\lambda}{16\pi^2}\frac{1}{d-4}. \end{aligned} \tag{12.41}$$

The bare parameters and the field renormalization can be computed from them using Eq. (12.29). In particular, we find that $Z_\phi(d-4)_{1-\text{loop}} = 1$ and the scalar field does not get renormalized at one loop.

Having reached this point, some remarks are in order. We notice that the construction of counterterms is intrinsically ambiguous because together with the divergent part we can also subtract a finite contribution. In our analysis we just removed the pole parts without imposing any renormalization condition at a particular value of the momenta. This is called minimal subtraction (MS). Most frequently, however, one uses a modified version of MS denoted by $\overline{\text{MS}}$ consisting in subtracting, together with the pole, also the term $\gamma - \log(4\pi)$. This results in a simplification of the expressions.

Both MS and $\overline{\text{MS}}$ are examples of *mass independent subtraction schemes* [3–5]. They owe their name to the fact that by subtracting just the pole at $d = 4$ (plus maybe other numerical constants) we get counterterms that are independent of any mass scale of the theory. As we will see very soon, this feature is very convenient for computational purposes and their use is crucial in the construction of effective field theories to be presented later.

In the study of the renormalization of QED carried out in Chap. 8 we learned that it is enough to consider the contribution of the 1PI irreducible diagrams, since all other diagrams can be written in terms of them. Here the situation is exactly the same, and we can focus our attention on the correlation functions obtained by summing all amputated 1PI Feynman diagrams with n external legs and incoming momenta $p_1, \ldots, p_n$. The correlation functions computed from the Lagrangian $\mathscr{L} + \mathscr{L}_{\text{ct}}$ are finite in the limit $d \to 4$, since the counterterms are constructed to cancel all divergences. These bare 1PI correlation functions depend on the bare parameters m_0 and λ_0 but are independent of the energy scale μ, which does not enter in (12.27).

The bare 1PI correlation functions can be written as [cf. (8.90) and the associated discussion]

$$\Gamma_n(p_i; m_0, \lambda_0, d-4)_0 = Z_\phi(d-4)^{-\frac{n}{2}}\Gamma_n(p_i; m, \lambda, \mu, d-4), \tag{12.42}$$

where in the right-hand side we have the correlation function computed with the renormalized fields. This depends only on the renormalized couplings m and λ, as well as on μ. Since the bare parameters do not depend on μ, the left-hand side is independent of this arbitrary scale, while the μ-dependence of the right-hand side is both explicit and implicit through the renormalized parameters. Taking the derivative with respect to μ we can write the analog of the renormalization equations discussed in Chap. 8

$$\begin{aligned}\Bigg[\mu\frac{\partial}{\partial\mu} + \beta\left(\lambda, \frac{m}{\mu}, d-4\right)\frac{\partial}{\partial\lambda} + \gamma_m\left(\lambda, \frac{m}{\mu}, d-4\right)m\frac{\partial}{\partial m}\\ -n\gamma\left(\lambda, \frac{m}{\mu}, d-4\right)\Bigg]\Gamma_n(p_i; m, \lambda, \mu, d-4) = 0.\end{aligned} \tag{12.43}$$

Since they only involve renormalized quantities these equations are regular in the limit $d \to 4$. The functions appearing on the left-hand side are defined by

$$\begin{aligned}\beta\left(\lambda,\frac{m}{\mu},d-4\right)&=\mu\frac{\partial\lambda}{\partial\mu},\\ \gamma_m\left(\lambda,\frac{m}{\mu},d-4\right)&=\frac{\mu}{m}\frac{\partial m}{\partial\mu},\\ \gamma\left(\lambda,\frac{m}{\mu},d-4\right)&=\frac{1}{2}\mu\frac{\partial}{\partial\mu}\log Z_\phi.\end{aligned} \tag{12.44}$$

They measure the change of λ, m and Z_ϕ with the scale μ.

The main reason to introduce the energy scale μ was to keep the dimensions of the expressions correct. It would be interesting to trade the evolution of the renormalized couplings with respect to μ by their evolution with respect to some physically meaningful energy scale, such as the one characterizing the process under study. To this end, we rescale all momenta in the correlation functions by a common factor $p_i \to sp_i$ and study how they change with s.

The 1PI function $\Gamma_n(sp_i;\lambda,m,\mu,d-4)$ has canonical dimension

$$D_n = 4 - n - \frac{d-4}{2}(n-2). \tag{12.45}$$

This formula can be heuristically justified by dimensional analysis of the contribution of a one-loop 1PI diagram with $n = 2k$ external legs. The canonical dimension D_n determines how the correlation function changes when we change the energy units in which the dimensionful parameters p_i, m and μ are measured. In mathematical terms this means that the correlation function is a homogeneous function of weight D_n, namely

$$\Gamma_n(\xi sp_i;\lambda,\xi m,\xi\mu,d-4) = \xi^{D_n}\Gamma_n(sp_i;\lambda,m,\mu,d-4), \tag{12.46}$$

for any $\xi > 0$. Euler's well known theorem for homogeneous functions implies that

$$\left(\mu\frac{\partial}{\partial\mu}+s\frac{\partial}{\partial s}+m\frac{\partial}{\partial m}-D_n\right)\Gamma_n(sp_i;\lambda,m,\mu,d-4)=0. \tag{12.47}$$

This equation is useful because it can be combined with Eq. (12.43) evaluated at sp_i to eliminate the derivative with respect to μ. Indeed, subtracting the two equations and taking the limit $d \to 4$ we arrive at the Callan–Symanzik equation

$$\begin{aligned}\Bigg\{-s\frac{\partial}{\partial s}+\beta\left(\lambda,\frac{m}{\mu}\right)\frac{\partial}{\partial\lambda}+\left[\gamma_m\left(\lambda,\frac{m}{\mu}\right)-1\right]m\frac{\partial}{\partial m}\\ +4-n\left[1+\gamma\left(\lambda,\frac{m}{\mu}\right)\right]\Bigg\}\Gamma_n(sp_i;m,\lambda,\mu)=0.\end{aligned} \tag{12.48}$$

It is at this point that the advantage of using a mass independent renormalization schemes becomes evident. Since the couterterms are mass independent, the functions appearing in the renormalization group equations only depend on the coupling constant λ

$$\left\{-s\frac{\partial}{\partial s}+\beta(\lambda)\frac{\partial}{\partial\lambda}+\Big[\gamma_m(\lambda)-1\Big]m\frac{\partial}{\partial m}\right.$$
$$\left.+4-n\Big[1+\gamma(\lambda)\Big]\right\}\Gamma_n(sp_i;m,\lambda,\mu)=0. \tag{12.49}$$

This fact makes it possible to formally integrate the equations in s to find how the 1PI correlation functions behaves under a simultaneous rescaling of *all* momenta. We introduce the functions $\overline{\lambda}(s)$ and $\overline{m}(s)$ solving the differential equations

$$s\frac{\partial}{\partial s}\overline{\lambda}(s)=\beta\Big(\overline{\lambda}(s)\Big),\qquad \frac{s}{\overline{m}(s)}\frac{\partial}{\partial s}\overline{m}(s)=\gamma_m\Big(\overline{\lambda}(s)\Big)-1 \tag{12.50}$$

with the initial conditions $\overline{\lambda}(1)=\lambda,\overline{m}(1)=m$. Since both $\beta(\lambda)$ and $\gamma_m(\lambda)$ can be computed order by order in renormalized perturbation theory, these equations completely determine the functions $\overline{\lambda}(s)$ and $\overline{m}(s)$.

Although not explicitly indicated, the function $\overline{\lambda}(s)$ also depends on the initial condition λ, whereas $\overline{m}(s)$ depends on both λ and m. This can be seen by rewriting (12.50) in integral form

$$\log s=\int_{\lambda}^{\overline{\lambda}(s)}\frac{dt}{\beta(t)},\qquad \overline{m}(s)=m\exp\left[\int_{\lambda}^{\overline{\lambda}(s)}dt\frac{\gamma_m(t)-1}{\beta(t)}\right]. \tag{12.51}$$

Differentiating with respect to λ and m we find that

$$\frac{\partial\overline{\lambda}}{\partial\lambda}=\frac{\beta(\overline{\lambda})}{\beta(\lambda)},\qquad \frac{\partial\overline{m}}{\partial m}=\frac{\overline{m}}{m},\qquad \frac{\partial\overline{m}}{\partial\lambda}=\overline{m}\left[\frac{\gamma_m(\overline{\lambda})-\gamma_m(\lambda)}{\beta(\lambda)}\right]. \tag{12.52}$$

Using these identities it is not difficult to show that the renormalization group equations (12.49) are formally solved by

$$\Gamma_n(sp_i;m,\lambda,\mu)=s^{4-n}\Gamma_n\Big(p_i,\overline{m}(s),\overline{\lambda}(s),\mu\Big)\exp\left[-n\int_1^s\frac{ds'}{s'}\gamma\Big(\overline{\lambda}(s')\Big)\right] \tag{12.53}$$

This solution gives the dependence of the 1PI Green's functions on the momentum rescaling factor s and can be used to determine their high-energy behavior by taking the limit $s\to\infty$. To get some intuition about the meaning of Eq. (12.53), we consider a massless theory sitting at a fixed point of the renormalization group flow where $\beta(\lambda^*)=0$. In this case (12.53) takes the simple form

$$\Gamma_n(sp_i;\lambda^*,\mu)=s^{4-n(1+\gamma^*)}\Gamma_n(p_i;\lambda^*,\mu). \tag{12.54}$$

This result would be the one expected from dimensional analysis in a theory where the scalar field $\phi(x)$ has energy dimensions $D_\phi=1+\gamma^*$ instead of the canonical value $D_\phi=1$. For this reason γ^* is called the anomalous dimension of the field. This name is also given by extension to the function $\gamma(\lambda, m/\mu)$ introduced in Eq. (12.44), although it should be noticed that, strictly speaking, it only admits such an interpretation at the fixed point.

12.4 The Renormalization Group Equations in Dimensional Regularization

So far we have studied some general properties of the renormalization group equations. To extract more precise information from them one needs to know the functions $\beta(\lambda)$, $\gamma(\lambda)$ and $\gamma_m(\lambda)$. This is going to be our next task, namely we are going to learn how to compute these functions using DR. The techniques we describe next were developed in [3, 4].

Our analysis is going to be as general as possible. For this reason we consider a theory characterized by a number of bare couplings denoted collectively by λ_{0k}, with $k = 1, \ldots, N$. They include coupling constants, masses and the wave function renormalization of the different fields. As prescribed by DR, we define the theory in dimension d and proceed to construct the counterterms required to cancel the divergences order by order in renormalized perturbation theory. A divergent L-loop diagram give rise to poles of the type shown in Eq. (12.17). Collecting the contributions to all orders in perturbation theory the bare couplings can be expressed as a Laurent series in $d-4$, namely

$$\lambda_{0k} = \mu^{D_k}\left[\lambda_k + \sum_{\nu=1}^{\infty}\frac{a_k^{(\nu)}(\lambda_\ell)}{(d-4)^\nu}\right], \tag{12.55}$$

where the coefficients $a_k^{(\nu)}(\lambda_\ell)$ depend on the renormalized couplings λ_ℓ. The energy dimensions D_k of the bare couplings are generically of the form

$$D_k = D_k^{(0)} + (d-4)D_k^{(1)}, \tag{12.56}$$

where $D_k^{(0)}$ is the dimension of λ_{0k} in four dimensions. The renormalized couplings have been rescaled by the appropriate powers of μ in order to make them dimensionless.

Each coefficient $a_k^{(\nu)}(\lambda_\ell)$ in the expansion (12.55) receives contributions from divergent L-loop diagrams with $L \geq \nu$. What makes these expressions still useful is that all coefficients $a_k^{(\nu)}(\lambda_\ell)$ with $\nu > 1$ can be expressed in terms of $a_k^{(1)}(\lambda_\ell)$, as we show next.

Let us differentiate Eq. (12.55) with respect to μ. Since the bare couplings are independent of this energy scale, the left hand side gives zero. The right-hand side, however, depends on μ both explicitly and implicitly through the renormalized couplings λ_k. We can take the derivative of the couplings to have the form

$$\mu \frac{\partial \lambda_k}{\partial \mu} = A_k(\lambda_\ell) + (d-4) B_k(\lambda_\ell) \tag{12.57}$$

and solve order by order in $d-4$. To begin with we get two terms proportional to $d-4$ that have to cancel. This determines the coefficient $B_k(\lambda_\ell)$ to be

$$B_k(\lambda_\ell) = -\lambda_k D_k^{(1)}. \tag{12.58}$$

Repeating this for the terms of order $(d-4)^0$ we find

$$\begin{aligned} A_k(\lambda_\ell) = {} & -\lambda_k D_k^{(0)} - D_k^{(1)} \left(1 - \lambda_k \frac{\partial}{\partial \lambda_k}\right) a_k^{(1)}(\lambda_\ell) \\ & + \sum_{j \neq k}^{N} D_j^{(1)} \lambda_j \frac{\partial}{\partial \lambda_j} a_k^{(1)}(\lambda_\ell). \end{aligned} \tag{12.59}$$

Finally, from the coefficients of the poles at $d=4$ the following recursion relation is obtained

$$\begin{aligned} 0 = {} & \left(D_k^{(0)} + A_k \frac{\partial}{\partial \lambda_k}\right) a_k^{(\nu)}(\lambda_\ell) + D_k^{(1)} \left(1 - \lambda_k \frac{\partial}{\partial \lambda_k}\right) a_k^{(\nu+1)}(\lambda_\ell) \\ & + \sum_{j \neq k}^{N} \left[A_j(\lambda_\ell) \frac{\partial}{\partial \lambda_j} a_k^{(\nu)}(\lambda_\ell) - D_j^{(1)} \lambda_j \frac{\partial}{\partial \lambda_j} a_k^{(\nu+1)}(\lambda_\ell) \right], \end{aligned} \tag{12.60}$$

where $A_j(\lambda_\ell)$ is given above in terms of $a_k^{(1)}(\lambda_\ell)$.

There are several interesting conclusions to be extracted from the relations just derived. First, taking the limit $d \to 4$ we find that the running of the coupling λ_k is fully determined by the coefficients of the single poles at $d=4$, namely

$$\begin{aligned} \mu \frac{\partial \lambda_k}{\partial \mu} = {} & -\lambda_k D_k^{(0)} - D_k^{(1)} \left(1 - \lambda_k \frac{\partial}{\partial \lambda_k}\right) a_k^{(1)}(\lambda_\ell) \\ & + \sum_{j \neq k}^{N} D_j^{(1)} \lambda_j \frac{\partial}{\partial \lambda_j} a_k^{(1)}(\lambda_\ell). \end{aligned} \tag{12.61}$$

At this point we have to recall that all couplings have been made dimensionless by rescaling them with powers of μ. This can be undone by a new rescaling

$$\lambda_k \longrightarrow \mu^{-D_k^0} \lambda_k \tag{12.62}$$

that has the effect of canceling the first term on the right-hand side of Eq. (12.61). This equation becomes even simpler when using a mass independent subtraction scheme. In this case the counterterms do not depend on the mass couplings of the theory and as a consequence they do not appear in the sum on the right-hand side of (12.61).

As we already explained, generically $a_k^{(1)}(\lambda_\ell)$ receives contributions to all loops. However, as promised, the recursion relations (12.60) give a way to compute $a_k^{(v)}(\lambda_\ell)$ with $v = 2, 3, \ldots$ in terms of $a_k^{(1)}(\lambda_\ell)$. This fact provides a very convenient method to check loop computations since, for instance, the coefficient of the two-loop $1/(d-4)^2$ pole is determined by the one-loop $1/(d-4)$ pole, and so on.

We go back now to the ϕ^4 theory we studied in the previous section. At one loop we only have to worry about two couplings m^2 and λ, since there is no wave function renormalization at this order. From Eqs. (12.30) and (12.41) we find the following expression for the bare couplings at one loop

$$\begin{aligned} m_0^2 &= m^2 \left(1 - \frac{\lambda}{16\pi^2}\frac{1}{d-4}\right), \\ \lambda_0 &= \mu^{4-d}\left(\lambda - \frac{3\lambda^2}{16\pi^2}\frac{1}{d-4}\right). \end{aligned} \tag{12.63}$$

We have used a mass-independent scheme, and the coefficients of the Laurent expansions of the bare couplings do not depend on the renormalized mass m. This simplifies the calculation of the renormalization group functions $\beta(\lambda)$ and $\gamma_{m^2}(\lambda)$

$$\begin{aligned} \beta(\lambda) &\equiv \mu\frac{\partial\lambda}{\partial\mu} = \frac{3\lambda^2}{16\pi^2} \\ \gamma_{m^2}(\lambda) &\equiv \frac{\mu}{m^2}\frac{\partial m^2}{\partial\mu} = \frac{\lambda}{16\pi^2} \end{aligned} \tag{12.64}$$

With this result we find that the beta function vanishes at zero coupling but it is positive for $\lambda > 0$. Applying what we learned from our analysis of Sect. 8.2 we conclude that $\lambda = 0$ is an infrared fixed point of the renormalization group flow. The one-loop beta function equation in (12.64) can be integrated to give

$$\lambda(\mu) = \frac{\lambda(\mu_0)}{1 - \frac{3}{16\pi^3}\lambda(\mu_0)\log\left(\frac{\mu}{\mu_0}\right)}, \tag{12.65}$$

where μ_0 is an arbitrary reference energy scale. This shows that the renormalized coupling grows with the energy μ. The coupling decreases as we go to lower energies, so perturbation theory becomes more and more reliable in this regime. This also indicates that the operator ϕ^4, which was originally marginal, becomes irrelevant once quantum corrections are included. Finally, we notice that the one-loop result (12.65) blows up at the nonperturbative energy scale

$$\mu = \mu_0 \exp\left[\frac{16\pi^3}{3\lambda(\mu_0)}\right]. \tag{12.66}$$

This Landau pole, similar to the one discussed for QED in Sect. 8.2, indicates that the theory becomes strongly coupled at high energies.

12.5 The Issue of Quadratic Divergences

We can return now to the problem of quadratic divergences introduced in Sect. 12.1. Instead of using DR we regularize the Euclidean integral (12.7) for $d = 4$ using a sharp cutoff $|p_E| < \Lambda$ and find the following leading behavior as $\Lambda \to \infty$

$$\int_{|p_E|<\Lambda} \frac{d^4 p_E}{(2\pi)^4} \frac{1}{(p_E^2 + m^2)^n} \sim \begin{cases} \frac{m^2}{8\pi^2}\left[\frac{\Lambda^2}{m^2} - \log\left(\frac{\Lambda^2}{m^2}\right)\right] & n = 1 \\ \frac{1}{8\pi^2}\left[\log\left(\frac{\Lambda^2}{m^2}\right) - \frac{1}{2}\right] & n = 2 \\ \frac{m^{4-2n}}{8\pi^2(n-1)(n-2)} & n > 2 \end{cases} \tag{12.67}$$

where we have dropped all terms that go to zero in this limit. These expressions contrasts with the DR result (12.11), where the integral diverges in the same way for $n = 1$ and $n = 2$, namely with a simple pole at $d = 4$.

The one-loop renormalization of the ϕ^4 field theory described in the previous section can now be implemented using the cutoff regularization. The cancelation of the divergent part of the diagram (12.31) gives the bare mass and coupling constant

$$\begin{aligned} m_0(\Lambda)^2 &= m^2 \left\{1 - \frac{\lambda}{16\pi^2}\left[\frac{\Lambda^2}{m^2} - \log\left(\frac{\Lambda^2}{m^2}\right)\right]\right\}, \\ \lambda(\Lambda) &= \lambda\left[1 - \frac{3\lambda}{16\pi^2}\log\left(\frac{\Lambda^2}{m^2}\right)\right]. \end{aligned} \tag{12.68}$$

We can invert the first equation to write the renormalized mass in terms of the bare parameters to first order in the bare coupling constant

$$m^2 = m_0(\Lambda)^2 + \frac{\lambda_0(\Lambda)}{16\pi^2}\left[\Lambda^2 - \log\frac{\Lambda^2}{m_0(\Lambda)^2}\right]. \tag{12.69}$$

Would the scalar theory be valid for arbitrary high energies, this would be the end of the story. The cutoff Λ would be an artifact of the quantization that should disappear at the end of the calculations. Physical results would only depend on the renormalized quantities m and λ. The situation is however different if we have reasons to believe that our theory is only valid up to certain energy scale at which new physics is expected to play a role. Then Eq. (12.69) has to be interpreted in Wilsonian terms (see Sect. 8.5) by regarding Λ as the energy above which our ϕ^4 theory is replaced by some unknown new dynamics. Just below this scale the leading part of the theory (not including irrelevant operators) is defined by the Lagrangian (12.27), the effect of the high energy degrees of freedom is codified in the cutoff dependence of the bare field and parameters $\phi_0(x, \Lambda)$, $m_0(\Lambda)$ and $\lambda_0(\Lambda)$. From this point of view m and λ are the parameters characterizing the theory at energies well below the cutoff scale, $E \ll \Lambda$.

The relation (12.69) between the low energy (renormalized) mass and the high energy bare parameters shows a strong dependence of the former on the cutoff Λ. Indeed, due to the term proportional to Λ^2 the value of the mass m will be determined by the cutoff scale unless the value of $m_0(\Lambda)$ is carefully chosen to cancel the contribution of the quadratic term up to many decimal places. The conclusion is that the preservation of the hierarchy $m \ll \Lambda$ requires an important fine tuning of the mass at the cutoff scale. This is the hierarchy problem.

Any theory with fundamental scalars is afflicted with this problem, including the standard model due to the presence of the Higgs field. The only exception are theories with Nambu-Goldstone bosons. They only include derivatives couplings preserving the invariance $\phi(x) \to \phi(x)+$ constant, thus forbidding any mass term in the action.[2]

We have seen in Sect. 12.2 that using DR there are no quadratic (or polynomial) divergences. The momentum integral of the one-loop self-energy diagram (12.31) has the same divergent behavior when $d \to 4$ as the milder logarithmically divergent integral appearing in the calculation of the four-point function. In fact, quadratically divergent integrals are not signaled in DR by higher order poles at $d = 4$, but by additional poles for $d < 4$. We can see this from Eq. (12.11). For the logarithmically divergent integral $I_2(d, m^2)$ the Gamma function in the numerator has a single pole for real positive d, namely at $d = 4$. In the case of $I_1(d, m^2)$, on the other hand, $\Gamma(\frac{2-d}{2})$ has, besides the pole at $d = 4$, another one at $d = 2$. Generically [6], in the integrals arising from L-loop diagrams these additional poles occur for fractional values of the dimension, $d = 4 - \frac{2}{L}$. This is how quadratic divergences are identified using DR.

The previous discussion might lead us to believe that the hierarchy problem is a regularization artifact that can be disposed of by a smart choice of the regulator. The whole thing, however, is more complicated. Integrating the second equation in (12.64) we find

$$m(\mu)^2 = m(\mu_0)^2 \exp\left[\int_{\lambda(\mu_0)}^{\lambda(\mu)} \frac{dx}{\beta(x)}\gamma_{m^2}(x)\right]. \tag{12.70}$$

This expression shows that the mass at the scale μ is proportional to the initial condition $m(\mu_0)$. This however does not clarify the issue, since in order to decide whether the ultraviolet sensitivity of the low energy parameters persists in DR we should find out how the initial condition $m(\mu_0)$ depends on the high energy scales.

We arrive at the conclusion that in order to understand the low energy role of quadratic divergences in DR we should look into the more general problem of understanding this regularization procedure in a Wilsonian setup. Therefore we turn to a brief description of effective field theories to address systematically the question of the separation of scales and how low-energy properties can be derived in a theory with

---

[2] The other known way of canceling quadratic divergences is to have supersymmetry (see Sect. 13.2), where the quadratically divergent corrections to the scalar masses are cancelled by the contribution of diagrams with fermion loops.

natural energy scales that are widely separated. As a bonus we will clarify the role played in Physics by nonrenormalizable field theories and arrive at the formulation of a criterion for naturalness [7] in quantum field theory.

12.6 Effective Field Theories: A Brief Introduction

One of the main reasons behind our progress in the understanding of physical processes is the fact that at a given length scale, and using the correct variables, our description of the physical phenomena is to a large extend independent of the physics at much shorter distances. This is a glaring fact in the history of Physics. For example, thermodynamics was formulated well before a microscopic description of the thermal processes in terms of statistical mechanics and atomic theory was available. Similarly, a moderately accurate calculation of the energy levels of the hydrogen atom is possible without concerning ourselves with the internal structure of the proton. Details such as its spin or charge radius have indeed an effect on the hydrogen spectral lines, but these are subleading corrections.

These examples, that can be multiplied at will, illustrate the basic fact that Physics largely deals with the formulation of effective theories describing physical phenomena within a certain range of energy scales with an acceptable accuracy. This is also the case in quantum field theory. At energies below the Fermi energy $1/\sqrt{G_F}$ weak interactions can be faithfully described using the Fermi theory. Only when higher energies become available experimentally, the effective low energy description has to be replaced by (or embedded in) a more general theory that takes into account the new degrees of freedom relevant for the exploration of new phenomena.

The basic ingredients in the building of effective field theories are the light degrees of freedom and the relevant symmetries of the problem. The latter provide the guiding principle to write a Lagrangian that would be the starting point for the calculation of observables. In Sect. 8.5 we learned that the infrared physics is dominated by relevant and marginal operators. A Lagrangian constructed using only these operators defines a renormalizable theory, such as QED, QCD or ϕ^4. Observables can be computed in terms of a limited number of parameters associated to the renormalized couplings of the relevant and marginal operators in the action.

The description in terms of relevant and marginal operators is very accurate in the deep infrared region. However, if we want to include the corrections due to new physics above an energy scale M we have to include irrelevant operators. These, generically, will appear in the action suppressed by the necessary powers of the scale M at which the new degrees of freedom become excited. This is precisely what happens in the case of Fermi's theory of weak processes, where β-decay is described by the four-fermion interaction of the Fermi theory discussed in Sect. 8.5 and M is set by the Fermi energy. Another example is provided by the description of the low energy properties of leptons and light quarks. At energies $E \ll m_b, m_t$ we do not have to include the b and t quarks as dynamical fields. They make themselves noticed, however, through irrelevant operators in the effective action for the light fields suppressed by powers of $1/m_b$ and $1/m_t$.

These examples should be enough to illustrate how as we go to higher energies (i.e., as we increase the power of our "magnifying glass") effective theories are replaced by a more complete description. In general this process repeats as we increase the energy. For example, the standard model is itself expected to be embedded at higher energies in a more general theory (maybe a GUT). Going to arbitrarily high energies would require a theory of everything that is not yet available.

The presence of irrelevant operators means that effective field theories are nonrenormalizable. This might seem to be a problem. The common lore states that nonrenormalizable theories do not have predictive power. The cancelation of the divergences in loop diagrams requires the introduction of an infinite number of different couterterms. Thus, in order to calculate observables one would need to specify an infinite number of parameters associated with the infinite number of irrelevant operators generated by quantum corrections.

This argument, however, is far too naive. In effective field theories we are interested in physical phenomena taking place in a range of energies much below the scale of new physics, $E \ll M$, and the contributions of the nonrenormalizable counterterms to a physical processes are weighted by powers of E/M. As it turns out, to a given degree of accuracy, there are only a few irrelevant operators that need be taken into account in our computations. Therefore, when looked at in the right way, nonrenormalizable theories are respectable and predictive.

It is important to realize that frequently it is experimental information that forces us to introduce higher-dimensional operators in a renormalizable theory. The classic example is the discovery of neutrino masses and mixing. The simplest way to accommodate this experimental fact in the standard model would be to include for each generation a sterile right-handed neutrino that is a singlet under the standard model gauge group $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)_Y$. This would generate Dirac mass terms for the neutrinos while preserving lepton number conservation. One can, however, take the widely accepted point of view that global symmetries such as lepton number conservation are mere accidental symmetries of the low energy theory that do not have to be preserved at high energies. In this case Majorana mass terms for the neutrinos are allowed.[3] The simplest way to generate these terms is by adding to the standard model Lagrangian the following dimension-five operator

$$\Delta\mathscr{L}_{\mathrm{SM}} = -\frac{1}{M}\sum_{i,j=1}^{3} g_{ij}\left(\overline{\mathbf{L}_i^C}\sigma^2\mathbf{H}\right)\left(\mathbf{H}^T\sigma^2\mathbf{L}_j\right) + \text{h.c.} \tag{12.71}$$

where g_{ij} are dimensionless coupling constants, $\mathbf{L}_i$ are the three lepton doublets introduced in Table 5.1, and $\mathbf{H}$ is the Higgs doublet (10.3). This term is gauge invariant,

3 As a matter of fact, once we decide that lepton number conservation is not a fundamental symmetry we can also introduce, in addition to the Dirac masses, Majorana mass terms for the right-handed neutrinos.

as can easily be shown using the definition of the charge conjugated spinors and the identity (10.6). Since this operator has dimension five it comes suppressed by M, the energy scale at which new physics is expected. Upon symmetry breaking, the Higgs doublet develops the vacuum expectation value (10.11) and the new term generates a Majorana mass for the three neutrinos [cf. equation (11.41)]

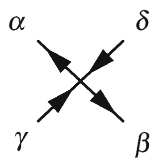
$$\Delta\mathscr{L}_{\text{SM}} = -\frac{\mu^2}{M}\sum_{i,j=1}^{3} g_{ij}\overline{\nu_i^C}\nu_j + \text{h.c.} \tag{12.72}$$

Using the experimental value of μ and the bounds for the neutrino masses it turns out that M can be as high as $M \sim 10^{15}\,\text{GeV} \gg m_W$. The discovery of neutrino masses may provide a hint to new physics at some high energy scale M. This would indicate that the standard model is an effective field theory, and therefore we should also include irrelevant operators to describe the effect of its ultraviolet completion.

After this long digression we study some basic features of effective field theories. Detailed introductions to the subject can be found in Ref. [8–11]; here we follow mainly the presentations of [12–14]. To illustrate our discussion we consider two unphysical toy models that however contain all the main features of more realistic effective field theories. The first is a non-renormalizable theory of a single Dirac spinor with a four-fermion interaction

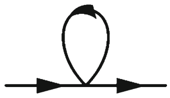
$$\mathscr{L} = \overline{\psi}(i\partial\!\!\!/ - m)\psi - \frac{a}{\Lambda^2}(\overline{\psi}\psi)^2 + \cdots \tag{12.73}$$

where a is a dimensionless coupling and the dots stand for higher-dimensional operators that we ignore. Using this Lagrangian we can study the effect of loop corrections induced by non-renormalizable interaction. These are the ones that, allegedly, would render the theory non-predictive. From our previous discussion we know that Λ sets the energy scale at which our nonrenormalizable Lagrangian should be completed with new degrees of freedom. We quantize the theory using this scale as a cutoff. The Feynman rules contain a single four-fermion vertex



$$= \frac{2ia}{\Lambda^2}\Big(\delta_{\alpha\delta}\delta_{\beta\gamma} - \delta_{\alpha\gamma}\delta_{\beta\delta}\Big), \tag{12.74}$$

and the only one-loop diagram contributing to the fermion self-energy is given by

$$-i\Sigma_{\alpha\beta}(\not{p},\Lambda) = \alpha$$  $$\beta$$

$$= -\frac{6am}{\Lambda^2}\delta_{\alpha\beta}\int^{\Lambda}\frac{d^4q}{(2\pi)^4}\frac{1}{q^2-m^2+i\varepsilon}. \tag{12.75}$$

As explained in Chap. 8 the physical mass of the fermion is defined by the zero of $\not{p} - m - \Sigma(\not{p}, \Lambda)$. Since the one-loop fermion self-energy is independent of the momentum, the mass correction is simply given by

$$\delta m = -\frac{6iam}{\Lambda^2}\int^{\Lambda}\frac{d^4q}{(2\pi)^4}\frac{1}{q^2-m^2+i\varepsilon} = -\frac{3am}{4\pi^2}\left[1+\frac{m^2}{\Lambda^2}\log\left(\frac{\Lambda^2}{m^2}+1\right)\right]$$
$$\sim -\frac{3am}{4\pi^2} \quad (\text{when}\, m \ll \Lambda). \tag{12.76}$$

To compute the integral we have performed a Wick rotation, as explained in Sect. 12.2, and integrated the Euclidean momentum in the range $|q_E| < \Lambda$. We have found that the leading correction to the mass is of order $(m/\Lambda)^0$, i.e., it is not suppressed by powers of Λ. This is not a peculiarity of the dimension-six operator chosen here. It also occurs for other higher-dimension operators.

This is a disastrous result. It leads to the conclusion that the fermion mass gets corrections from higher-dimensional operators that are not suppressed by the scale Λ and therefore are large even when we are considering energies much below the scale of new physics. What we are saying is that the value of low energy parameters is strongly influenced by what is going on at arbitrarily high energies. Thus, in order to compute the corrections to the mass of the fermion in our theory we would need to know the details of the dynamics at energies above the scale Λ.

The reason behind our failure in separating low energy physics from the details of the theory at high energies lies in the fact that we did not renormalize the theory. Instead of cutting off the integrals at the scale Λ we are going to regularize it using DR and a mass independent subtraction scheme. From the expressions derived in previous sections we can compute the fermion self-energy to be

$$\Sigma(\not{p}, \Lambda) = -\frac{6iam}{\Lambda^2}\int\frac{d^dq}{(2\pi)^d}\frac{1}{q^2-m^2+i\varepsilon}$$
$$= -\frac{3am}{8\pi^2}\left(\frac{m}{\Lambda}\right)^2\left[\frac{2}{d-4}+\gamma+\log\left(\frac{m^2}{4\pi\mu^2}\right)+\cdots\right]. \tag{12.77}$$

In the $\overline{\text{MS}}$ subtraction scheme we add a counterterm that cancels the pole in $d = 4$ together with the constants $\gamma - \log(4\pi)$. Then we find the following correction to the fermion mass

$$\delta m = -\frac{3am}{8\pi^2}\left(\frac{m}{\Lambda}\right)^2 \log\left(\frac{m^2}{\mu^2}\right). \tag{12.78}$$

This is a much nicer result. The mass correction is suppressed by powers of m/Λ, small in the regime where the effective field theory is applicable, $m \ll \Lambda$. In addition, the expression only depends logarithmically on μ. This energy scale is an artifact of the regularization and therefore should be absent of all physical quantities.

It is a general result that in a mass independent subtraction scheme, effective field theories produce a well defined expansion in powers of m/Λ or E/Λ, where E is the characteristic energy of the process under consideration. This means that to a given numerical accuracy only a few terms in the expansion should be considered. It is in this sense that effective field theories can be considered as predictive as renormalizable quantum field theories.

The reader might be puzzled at the comparison of the different results we have obtained for the mass renormalization using a cutoff, and DR plus a mass independent subtraction scheme. In fact, there is no contradiction between them. Physical predictions cannot depend on the way we choose to regularize and renormalize the theory. Cutting off the integrals at the scale Λ results in an infinite number of contributions to each order in $1/\Lambda$. Were we able to resum these terms we would obtain a result agreeing with the expression found using a mass independent scheme. The latter method provides a systematic way of organizing the $1/\Lambda$ contributions. As a consequence there is only a finite number of operators contributing to a given degree of accuracy.

Before closing our discussion of the Lagrangian (12.73) we mention the fact that the mass correction (12.78) is proportional to the mass m and therefore vanishes for $m = 0$. This apparently innocuous fact has a deep underlying explanation based on a symmetry enhancement of the theory at $m = 0$. Indeed, in the massless case both the kinetic term and the four-fermion interaction are invariant under the discrete chiral transformation

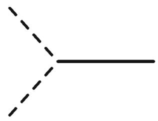
$$\psi \longrightarrow \gamma_5\psi, \quad \overline{\psi} \longrightarrow -\overline{\psi}\gamma_5. \tag{12.79}$$

This is however not a symmetry of the mass term, changing sign under it. Thus, the theory at $m = 0$ has an addition symmetry protecting the fermion from acquiring a mass through quantum corrections. This is why, in general, one can say that having a light fermion is *natural* in spite of the presence of a large energy scale Λ in the theory.

The second example we want to study is a renormalizable theory of two interacting real scalar fields with masses m and M, with $m \ll M$ and Lagrangian

$$\mathscr{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 + \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{M^2}{2}\Phi^2 - \frac{g}{2}\phi^2\Phi. \tag{12.80}$$

From the inspection of the interaction term we find that the Feynman rules contain the vertex



$$= -ig, \tag{12.81}$$

where the dashed and continuous lines represent respectively the light and heavy fields. The propagators are the usual ones for scalar fields with the appropriate values of the mass. To find the leading correction to the mass of the light field due to the heavy field, we consider the following diagram

$$\begin{aligned} &= -g^2\mu^{4-d}\int \frac{d^dq}{(2\pi)^d}\frac{i}{q^2-m^2+i\varepsilon}\frac{i}{(q+p)^2-M^2+i\varepsilon}. \\ &= g^2\mu^{4-d}\int_0^1 dx\, I_2\Big(d, xm^2+(1-x)M^2-x(1-x)p^2\Big). \end{aligned} \tag{12.82}$$

To write the second equality we have colleted together the two propagators using the trick (12.37), thus reducing the expression to a single integral of the type $I_2(d, m^2)$.

Let us take $m = 0$. To compute the leading correction to the mass it is enough to evaluate the previous diagram at zero momentum. An explicit calculation of the integral around $d = 4$ gives

$$\Bigg|_{p^2=m^2=0} = -\frac{ig^2}{16\pi^2}\left[\frac{2}{d-4}+\gamma+\log\left(\frac{M^2}{4\pi\mu^2}\right)+\cdots\right], \tag{12.83}$$

and in the $\overline{\text{MS}}$ subtraction scheme the mass corrections is found to be

$$\delta m^2 = \frac{g^2}{16\pi^2}\log\left(\frac{M^2}{\mu^2}\right), \tag{12.84}$$

which is nonzero even if we set $m = 0$. This simple calculation illustrates the important point that the mass of the scalar field is not protected against quantum corrections. The interaction with the heavy scalar produces a correction to the mass whose scale is set by the dimensionful coupling constant g, while depending only logarithmically on the mass of the heavy scalar. This means that having a scalar with mass well below $g/(4\pi)$ is *unnatural*.

A similar result would be obtained in the case of a light scalar field $\phi(x)$ coupled to a heavy fermion $\psi(x)$ of mass M through the Yukawa interaction

$$\mathscr{L}_{\text{int}} = g'\phi\overline{\psi}\psi. \tag{12.85}$$

Now, since g' is dimensionless, the scalar field mass m gets quantum corrections whose scale is set by the mass of the heavy fermion. This can be seen by computing the one-loop fermion correction to the scalar two-point function. Taking again $m = 0$ and using DR and the $\overline{\text{MS}}$ subtraction scheme the scalar acquires a mass of order

$$\delta m^2 \sim g'^2 M^2 \log\left(\frac{M^2}{\mu^2}\right). \tag{12.86}$$

Therefore there is no natural way to keep the mass of the scalar away from the scale of the heavy fermion.

The theories defined by the Lagrangians (12.73) and (12.80) show that while chiral symmetry makes light fermions *natural*, it is very difficult is to prevent scalar fields from acquiring masses of the order of the higher energy scales in the theory. This illustrates, in a toy model, the hierarchy problem: unlike fermions, low energy scalars are extremely sensitive to high energy scales.

These examples help also to clarify the issue of quadratic divergences in DR. We have seen that, in spite of the apparent absence of quadratic divergences in the Feynman integrals, the scalar masses in general are still quadratically sensitive to high energy scales. The hierarchy problem is therefore not an artifact of the regularization procedure.

12.7 Remarks on Naturalness

There are a number of hints, such as neutrinos masses and oscillations, indicating that the standard model is an effective theory. Since light scalars are unnatural in the presence of higher energy scales, it is necessary to explain what keeps the value of the Higgs mass below the scale 10^{15}–10^{19} GeV where new physics is expected, unless we are willing to accept an unnatural fine tuning of the Higgs mass. This can be rephrased as the fact that the ratio between the GUT or Planck mass scale (i.e., the scale of "new physics" M_{NP}) and the standard model energy scale $M_{\text{SM}} \sim 100$ GeV is a large number

$$\frac{M_{\text{NP}}}{M_{\text{SM}}} \sim 10^{13}-10^{16}. \tag{12.87}$$

The problem of accounting for large numbers has been around for a long time. Dirac [15] was worried about the emergence of large numbers in Physics, like the ratio between the strengths of the electromagnetic and gravitational interactions of protons and electrons. In a more modern context we can construct a dimensionless ration between the Fermi and Newton constants

$$\frac{G_F c^2}{G_N \hbar^2} \simeq 1.73 \times 10^{33}. \tag{12.88}$$

It is to Dirac's credit that he did not invoke any anthropic explanation. In his large number hypothesis he assumed that all these large dimensionless ratios should be related in a simple way to a single large number which he chose to be the age of the Universe. This led him to conclude that the fundamental constants of Nature vary with time.

We do not want to dwell any further on this subject. Excelent expositions of the notion of naturalness in high energy physics are available in the literature (see, for example, [16, 17]). In view, however, of the examples discussed in the previous section we find it necessary to state a *naturalness criterion* that probably most physicists would find acceptable:

> At any energy scale μ, a physical parameter or a set of physical parameters $\alpha_i(\mu)$ is allowed to be very small only if the replacement $\alpha_i(\mu) = 0$ would increase the symmetry of the system.

This criterion, originally formulated by Wilson [18] and further elaborated among others by 't Hooft [7] and Susskind [19], has been a guide for nearly four decades in the construction of theories beyond the standard model. The fact that the Higgs particle, if thought as an elementary scalar, has not yet been found adds a good deal of drama associated with naturalness.

This naturalness criterion may apply to particle physics, but in the broader context where gravity is included it is severely violated. In our discussion of effective field theories we have systematically forgotten the identity operators which, having zero dimension, should be dominant in the infrared. The reason why we could afford to ignore this operator so far is that we were not considering gravitational effects. The coupling of the identity operator receives contributions from the zero-point energy of all the quantum fields and, as long as gravity is left out of the game, can be simply ignored.

General relativity teaches us that all forms of energy gravitate, and this applies also to the zero-point energy of the quantum fields. Therefore once gravitational effects are considered there is no way to ignore the coupling of the identity operator to the gravitational field. This term is Einstein's famous cosmological constant Λ_c. Its contribution to the energy density of the Universe

$$\rho_\Lambda = \frac{\Lambda_c}{8\pi G_N} \tag{12.89}$$

can be measured from cosmological observations with the result

$$\rho_\Lambda \simeq (10^{-3}\,\text{eV})^4 = 10^{-48}\,\text{GeV}^4. \tag{12.90}$$

On the other hand, since ρ_Λ has dimensions of (energy)^4 and the only natural energy scale in gravity is the Planck mass $M_P \sim 10^{19}$ GeV, naturalness would require the scale of ρ_Λ to be set by M_P, that is

$$\rho_\Lambda \sim M_P^4 \sim 10^{76}\,\text{GeV}^4. \tag{12.91}$$

This means that there is a mismatch of more than 120 orders of magnitude between the natural and the measured value of the cosmological constant. To add to the puzzle, if we compare ρ_Λ with the cosmological critical density at the present time

$$\rho_c = \frac{3H_0^2}{8\pi G_N}, \tag{12.92}$$

we find the two values to be very close, $\rho_\Lambda \simeq 0.74\rho_c$.

Is it just serendipity that the measured value of the cosmological constant is so close to the Universe's critical density today? In view of this it is difficult to avoid the question of whether naturalness should only apply to particle physics in the absence of gravity. In fact, the problem of the apparent fine tuning of the cosmological constant is 60 or 70 orders of magnitude worse than the one for the Higgs mass. This brings in the question: why should gravity be excluded in naturalness arguments?

It is safe to say that currently nobody knows the answer to these questions. For the Higgs naturalness problems some scenarios have been suggested: supersymmetry and technicolor among others, that are likely to be tested soon. In the case of the cosmological constant, apart from anthropic arguments [20] or the string landscape [21, 22], there is very little to say. Large numbers are likely to continue haunting particle physicists and cosmologists for some time to come.

12.8 Coda: Heavy Particles and Decoupling

We have appraised mass independent subtraction schemes as the appropriate way to deal with the renormalization of effective field theories. They have, however, an important disadvantage: heavy particles do not decouple at energies below their masses, as expected from the Appelquist-Carazzone decoupling theorem [23].

A simple example showing this is provided by the calculation in QED of the contribution of a heavy fermion to the photon vacuum polarization (more details can be found in [13, 14]). The basic ingredient is the DR computation of the one loop polarization tensor

μ  ν

$$= \left(p_\mu p_\nu - p^2\eta_{\mu\nu}\right)\Pi(p^2; d-4) \tag{12.93}$$

where the polarization function $\Pi(p^2; d-4)$ is given by

$$\Pi(p^2; d-4) = -\frac{e^2}{12\pi^2}\left\{\frac{2}{d-4} + \gamma_E - \log(4\pi) + 6\int_0^1 dx\, x(1-x)\log\left[\frac{m_f^2 - x(1-x)p^2}{\mu^2}\right]\right\}. \tag{12.94}$$

Following the analysis of Sect. 8.3, the divergence in the previous expression is subtracted by adding to the QED Lagrangian a counterterm of the form $-\frac{1}{4}C(d-4)F_{\mu\nu}F^{\mu\nu}$. In the $\overline{\text{MS}}$ scheme we obtain

$$\Pi(p^2)_{\overline{\text{MS}}} = -\frac{e^2}{2\pi^2}\int_0^1 dx\, x(1-x)\log\left[\frac{m_f^2 - x(1-x)p^2}{\mu^2}\right]. \tag{12.95}$$

At leading order in the renormalized charge e the beta-function can be computed as

$$\beta(e) = \frac{e}{2}\mu\frac{\partial}{\partial\mu}\Pi(p^2)_{\overline{\text{MS}}} = \frac{e^3}{12\pi^2} \quad (\overline{\text{MS}} \quad \text{scheme}). \tag{12.96}$$

Alternatively, the same result can be obtained from the expression of the bare charge coupling contant

$$e_0 = e\mu^{\frac{4-d}{2}}\left(1 - \frac{e^2}{12\pi^2}\frac{1}{d-4}\right) + \text{finite part} \tag{12.97}$$

by applying the techniques introduced in Sect. 12.4. Since in a mass independent subtraction scheme the beta function is determined solely by the pole at $d = 4$, it does not depend on the value of the fermion mass. This is surprising because, on physical grounds, one would expect the fermion to decouple in the limit of large mass $m_f \to \infty$. In this limit the theory should have a vanishing beta function.

We repeat now the calculation of the beta function but using a mass dependent scheme. In particular we work in the so-called μ-scheme where the counterterm coefficient $C(d-4)$ is chosen in such a way that its contribution cancels the diagram (12.93) evaluated at the Euclidean momentum $p^2 = -\mu^2$. The renormalized polarization tensor is obtained by adding the contribution of the counterterm to the one-loop diagram, with the result

$$\Pi(p^2)_\mu = -\frac{e^2}{2\pi^2}\int_0^1 dx\; x(1-x)\log\left[\frac{m_f^2 - x(1-x)p^2}{m_f^2 + x(1-x)\mu^2}\right], \tag{12.98}$$

where the subscript μ indicates that we are dealing with the renormalized polarization function in the μ-scheme. The calculation of the beta function then gives

$$\begin{aligned}\beta(e) &= \frac{e}{2}\mu\frac{\partial}{\partial\mu}\Pi(p^2)_\mu \\ &= \frac{e^3}{2\pi^2}\int_0^1 dx\left[\frac{x^2(1-x)^2\mu^2}{m_f^2 + x(1-x)\mu^2}\right] \quad (\mu\text{-scheme}).\end{aligned} \tag{12.99}$$

Unlike the result (12.96) this beta function depends on the quotient between the fermion mass m_f and the subtraction point μ. In particular, when $m_f \ll \mu$ we find that $\beta(e)$ approaches the value (12.96), whereas in the opposite limit $m_f \gg \mu$ it tends to zero quadratically

$$\beta(e) \simeq \frac{e^3}{60\pi^2}\left(\frac{\mu}{m_f}\right)^2. \tag{12.100}$$

This is what we expect physically: a heavy fermion decouples from the low energy theory which asymptotically becomes a theory of free photons with vanishing beta function.

To understand why the $\overline{\text{MS}}$ subtraction scheme (or any mass-independent scheme for that matter) renders an incorrect result for the beta function of QED in the limit of large fermion mass we have to look at the renormalized polarization function (12.95). When the momentum goes below the fermion mass, $p^2 \ll m_f^2$, it approaches the value

$$\Pi(p^2)_{\overline{\text{MS}}} \simeq \frac{e^2}{72\pi^2}\log\left(\frac{m_f^2}{\mu^2}\right). \tag{12.101}$$

When $m_f \gg \mu$ this logarithm is large and, as a result, the perturbative expansion breaks down at low energies. This is the reason why the result obtained for the beta function is not reliable in this regime. The problem is absent in the mass dependent scheme used above, where the renormalized polarization tensor (12.98) vanishes in the limit of large fermion mass

$$\Pi(p^2)_{\mu} \simeq -\frac{e^2}{60\pi^2}\left(\frac{p^2+\mu^2}{m_f^2}\right) \longrightarrow 0. \tag{12.102}$$

The limit of heavy fermion mass is amenable to perturbation theory and the beta function can be reliably computed in this limit.

The way to deal with heavy particles in mass independent schemes is by integrating them out as we move down the energy ladder. At energies below the mass of a particle we have to use an effective field theory including only the light degrees of freedom at the corresponding scale, while the effects of the heavy fields are felt through higher-dimension operators. It is important to bear in mind that both the high and the low energy theories have the same light particle content, so they share the same infrared properties. They are however different in the ultraviolet, where the dynamics of the heavy particle distorts the high energy behavior of the effective field theory.

It is crucial that the description provided by the two theories be consistent at the threshold energy set by the mass of the particle that is being integrated out. This means, for example, that the scattering amplitudes of light particles cannot depend on whether we compute them using one theory or the other. The way to proceed is to match the Feynman graphs computed from the low energy effective field theory with the corresponding one-light-particle irreducible diagrams in the high energy theory. These are those Feynman graphs having only light particles on the external legs and that cannot be disconnected by cutting an internal light-particle line. These matching conditions implement the effects of heavy particles and high energy modes in the low energy effective field theory.

To summarize, the discussion carried out in this section shows that in the $\overline{\text{MS}}$ subtraction scheme, or any other mass-independent scheme, the decoupling of particles as we run from high to low energies has to be implemented by hand, integrating out the field that become heavy as we lower the energy. Thus, every time a particle