3

Basic examples

In the later chapters we will develop methods to treat large-momentum behavior. The complete treatment becomes rather intricate at times, so this chapter is devoted to exposing in their simplest form the issues we will be discussing. We will do this by examining the self-energy graph in ϕ^3 theory. This will exhibit the basic phenomena which we will later be treating in detail.

We will see that (in four-dimensional space-time) the graph is renormalized by a mass counterterm. Then the concept of 'degree of divergence' will be introduced by varying d, the dimensionality of space-time. This device will enable us to see how simple power-counting methods determine what counterterms are needed. It will also introduce us to the method of dimensional regularization.

The renormalization group will be introduced by examining the behavior of the graph as its external momentum, p, is made large. By exploiting the arbitrariness in the renormalization procedure, we can reduce the size of higher-order contributions when p^μ is large.

3.1 One-loop self-energy in ϕ^3 theory

Consider the graph shown in Fig. 3.1.1 in the ϕ^3 theory of (2.16.1). We define its contribution to the self-energy to be i times the value of the graph with the external propagators removed:

$$\Sigma_1(p^2) = \frac{\mathrm{i}}{2}\frac{g^2}{(2\pi)^4}\int \mathrm{d}^4k \frac{1}{[k^2 - m^2 + \mathrm{i}\varepsilon][(p+k)^2 - m^2 + \mathrm{i}\varepsilon]}. \quad (3.1.1)$$

The overall factor $\frac{1}{2}$ is a symmetry factor.

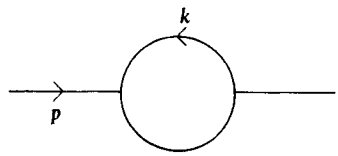

Fig. 3.1.1. One-loop self-energy graph in ϕ^3 theory.

When all components of k^μ get large, this integral diverges logarithmically. It is the simplest example of an ultra-violet divergence. As we will see, the divergence can be cancelled by a mass counterterm. But to explain the renormalization properly, we must discuss a number of other issues as well:

(1) The fact that if $|k^0| \simeq |\vec{k}|$ the divergence as k goes to infinity appears to be much worse.
(2) A precise way of formulating the statement that the divergence is cancelled by a mass counterterm.
(3) The arbitrariness inherent in the renormalization.
(4) The interpretation in coordinate space.

3.1.1 *Wick rotation*

The first of these problems is handled by recalling the Wick rotation into Euclidean space that was used to define the functional integral. This rotation determined the sign of the $i\varepsilon$ in the free propagator. The Wick rotation involved starting with imaginary time $t = -i\tau$, then performing the integral, and finally analytically continuing back to real time. In momentum space, this forces us to work with $k^0 = +i\omega$, the opposite sign appearing so that in the Fourier transformation $e^{ik\cdot x}$ is always a phase.

In the Euclidean formulation let us perform the k^0-integral first. The pole structure in the k^0-plane is shown in Fig. 3.1.2, when p^0 is imaginary. In this situation k^2 and $(p+k)^2$ are both negative, so that the integrand is positive definite. Observe that the factor i coming from the Wick rotation combines with the overall factor of i in (3.1.1) to make Σ_1 real. (We have $d^4k = id\omega d^3k$.)

Now rotate p^0 back to a real value. If $|p^0| < m$ then we have the situation shown in Fig. 3.1.3: there is no obstruction to rotating the k^0 contour to run along the real axis. It is only at this last step that there is a problem from the

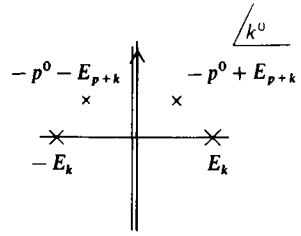

Fig. 3.1.2. The k^0-plane when p^0 is imaginary.

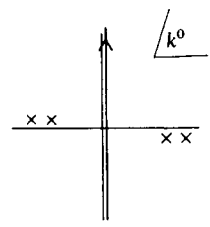

Fig. 3.1.3. The k^0-plane when p^0 is real, but $|p^0|$ is less than m.

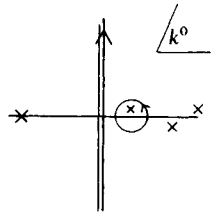

Fig. 3.1.4. The k^0-plane when p^0 is real, but $|p^0|$ is greater than m.

region of $|k^0| \simeq |\vec{k}|$. To avoid the problem we merely have to define the integral by rotating the k^0-contour to run along the imaginary axis.

Now continue p^0 to the region $|p^0| > m$. The case of negative p^0 is illustrated in Fig. 3.1.4. Again we Wick-rotate to imaginary k^0, but this time we pick up a pole term. Now the pole term occurs only when $(p^0)^2 > m^2 + (\vec{p} + \vec{k})^2 \equiv E_{p+k}^2$. Thus it contributes only in a finite region of $\vec{k}$; the UV divergence still comes from the integration over imaginary k^0.

The moral of all this is that the UV divergence is essentially Euclidean, i.e., we may regard k^0 as imaginary and $k^2 < 0$, $(p + k)^2 < 0$.

3.1.2 *Lattice*

We next need to quantify the divergence. The divergence comes from the asymptotic large-k behavior of the integrand which is $1/(k^2)^2$. Let us add and subtract a term with this behavior:

$$\Sigma_1 = \frac{ig^2}{32\pi^4}\left\{\int d^4k\left[\frac{1}{(k^2 - m^2 + i\varepsilon)[(p+k)^2 - m^2 + i\varepsilon]} - \frac{1}{(k^2 - \mu^2 + i\varepsilon)^2}\right]\right.$$
$$\left. + \int d^4k \frac{1}{(k^2 - \mu^2 + i\varepsilon)^2}\right\}. \tag{3.1.2}$$

The first integral is manifestly finite, for we have subtracted off the leading asymptotic behavior of the integrand. To avoid introducing an extra divergence at $k^2 = 0$ we have subtracted $1/(k^2 - \mu^2)^2$ rather than $1/(k^2)^2$. Since we add this term back on, the value of μ is irrelevant; Σ_1 is unchanged. The second term, while divergent, is independent of p. This is the fact that will enable us to cancel the divergence by a counterterm.

Of course we are manipulating divergent integrals, so that (3.1.2), as it stands, makes no sense. We will remedy this defect by using the fact that the theory is defined initially on a lattice. The propagators will then be different functions of momentum. However, the structure of (3.1.2) will be unchanged after imposing a cut-off, as we will now show.

To define the functional integral we not only had to Wick rotate time, but also had to put the theory on a space-time lattice, of spacing a. In the lattice theory, let the free propagator of a particle of mass M be $S_F(q; M, a)$. In the limit $qa \to 0$, this is just $i/(q^2 - M^2 + i\varepsilon)$. But it is zero if $qa > 1$, since high-momentum states do not exist on the lattice. (The reason is, of course, that when one makes a Fourier transformation on a discrete space, one only uses momentum modes with wave-lengths longer than a lattice spacing.) The self-energy on the lattice is finite, and (3.1.2) now reads:

$$\Sigma_1(p, m; a) = \frac{ig^2}{32\pi^4}\left\{\int d^4k[S_F(k; m, a)S_F(p + k; m, a) - S_F(k; \mu, a)^2] + \int d^4k S_F(k; \mu, a)^2\right\}$$
$$\equiv \Sigma_{1\,\mathrm{fin}}(p, m, \mu, a) + \Sigma_{1\,\mathrm{div}}(m, \mu, a). \tag{3.1.3}$$

All the integrals are now convergent, so (3.1.3) is a correct version of (3.1.2). As the lattice spacing goes to zero, the first integral approaches the first convergent integral in (3.1.2). The second integral diverges, but is independent of p. Thus (3.1.2) is not nonsensical, provided that the propagators are implicitly replaced by lattice propagators wherever necessary.

3.1.3 *Interpretation of divergence*

No matter how it is manipulated, the self-energy diverges in the continuum limit. The use of a lattice cut-off now enables us to quantify the divergence. From (3.1.3)

$$\begin{aligned}\Sigma_1 &= \frac{ig^2}{32\pi^4}\int d^4k S_F(k; \mu, a)^2 + \Sigma_{1\,\mathrm{fin}} \\ &= \frac{-g^2}{32\pi^4}\int_{\omega^2 + \vec{k}^2 < 1/a^2} d\omega d^3k \frac{1}{(\omega^2 + \vec{k}^2 + \mu^2)^2} + \text{finite} \\ &= \frac{-g^2}{16\pi^2}\int_0^{1/a} dk \frac{k^3}{(k^2 + \mu^2)^2} + \text{finite} \\ &= \frac{-g^2}{16\pi^2}\ln 1/a + \text{finite} \qquad \text{as } a \to 0. \end{aligned} \tag{3.1.4}$$

Thus we can interpret the divergence as follows: Let Σ be (i times) the sum of self-energy graphs. (As usual, the self-energy graphs are graphs for the propagator that have the external lines amputated and that cannot be split

Fig. 3.1.5. Summation of self-energy graphs into propagator.

into disconnected parts by cutting a single line.) Then the full propagator is

$$\tilde{G}_2(p^2) = \mathrm{i}/(p^2 - m^2 - \Sigma + \mathrm{i}\varepsilon). \tag{3.1.5}$$

This equation is illustrated in Fig. 3.1.5; the propagator is the sum of a geometric series involving the self-energy. The actual mass m_{ph} of the particle is determined by the pole position, $p^2 = m_{\mathrm{ph}}^2$. Evidently m_{ph}^2 is not m^2 but $m^2 + \Sigma(p^2 = m_{\mathrm{ph}}^2, m^2)$. In other words, the self-energy represents the dynamical contribution to the mass coming from the interactions. The divergence (3.1.4) is independent of p^2, so it is precisely a contribution to the mass. (We ignore higher orders for now.)

Traditionally, one observes that it is convenient to parametrize the theory, not by the mass parameter m that cannot be observed directly but by the physical mass m_{ph}. One writes the mass term in $\mathscr{L}$ as

$$-m_0^2\phi^2/2 = -m_{\mathrm{ph}}^2\phi^2/2 - \delta m^2\phi^2/2. \tag{3.1.6}$$

The first term is left in the free Lagrangian, so that the free propagator is $\mathrm{i}/(p^2 - m_{\mathrm{ph}}^2 + \mathrm{i}\varepsilon)$. But the second term – called the mass counterterm – is put into the interaction Lagrangian, and adjusted so that the full propagator has a pole at $p^2 = m_{\mathrm{ph}}^2$. The counterterm exactly cancels the dynamical contribution to the particle's mass. This is the basic idea of renormalization. It is physically irrelevant that δm^2 happens to diverge.

In perturbation theory, δm^2 is determined as a power series in g. To $O(g^2)$ we have, in addition to Fig. 3.1.1, the graph of Fig. 3.1.6 corresponding to the mass counterterm in (3.1.6). The self-energy to $O(g^2)$ with the new parametrization is the sum of Figs. 3.1.1 and 3.1.6. We call it the renormalized self-energy:

$$\begin{aligned}\Sigma_{1\mathrm{R}} &= \{\Sigma_1(p^2; m_{\mathrm{ph}}^2, a) + \delta m^2\}\big|_{a\to 0} \\ &= \Sigma_{1\mathrm{fin}}(p^2; m_{\mathrm{ph}}^2; \mu, 0) + [\delta m^2 + \Sigma_{1\mathrm{div}}].\end{aligned} \tag{3.1.7}$$

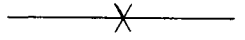

Fig. 3.1.6. Counterterm graph to the self-energy.

We adjust δm^2 first of all to cancel the divergence in $\Sigma_{1\mathrm{div}}$, so that the term in square brackets is finite as $a \to 0$. Then we adjust the finite part so that

$$\delta m^2 + \Sigma_{1\mathrm{div}} = -\Sigma_{1\mathrm{fin}}(m_{\mathrm{ph}}^2, m_{\mathrm{ph}}^2, \mu). \tag{3.1.8}$$

For the renormalized self-energy we find

$$\Sigma_{1R} = \frac{ig^2}{32\pi^4}\int d^4k \left\{ \frac{1}{(k^2 - m_{ph}^2 + i\varepsilon)[(p+k)^2 - m_{ph}^2 + i\varepsilon]} - \frac{1}{(k^2 - \mu^2 + i\varepsilon)^2} \right\}$$

$$- \text{value at } p^2 = m_{ph}^2. \tag{3.1.9}$$

Note that the μ-dependence cancels.

Equation (3.1.9) gives the value of the self-energy in the continuum theory correct to $O(g^2)$.

3.1.4 *Computation*

One way to calculate Σ_{1R} is to differentiate with respect to p^2. (It is of course a Lorentz scalar.) Integrating the result gives Σ_{1R}; the constant of integration is fixed by the renormalization condition that $\Sigma_{1R}(m_{ph}^2)$ is zero.

We have:

$$\frac{\partial \Sigma_{1R}}{\partial p^2} = \frac{p^\mu}{2p^2}\frac{\partial}{\partial p^\mu}\Sigma_{1R}$$

$$= \frac{-ig^2}{32\pi^4 p^2}\int d^4k \frac{p\cdot(p+k)}{(k^2 - m^2 + i\varepsilon)[(p+k)^2 - m^2 + i\varepsilon]^2}. \tag{3.1.10}$$

This is identical to what we would have obtained from the unrenormalized expression (3.1.1) without regard to the fact that it is divergent. We could have written it down directly without going through the long explanation that we used. But then there would have been no defense to the argument that we are manipulating meaningless quantities and that therefore quantum field theory makes no sense.

Since (3.1.10) is finite it can be easily calculated by using a Feynman parameter representation and then by shifting the k-integral:

$$\frac{\partial \Sigma_{1R}}{\partial p^2} = \frac{ig^2}{16\pi^4 p^2}\int_0^1 dx\, x \int d^4k \frac{p\cdot(p+k)}{[m_{ph}^2 - p^2 x - 2p\cdot kx - k^2 - i\varepsilon]^3}$$

$$= \frac{ig^2}{16\pi^4 p^2}\int_0^1 dx\, x \int d^4k \frac{p\cdot[k + p(1-x)]}{[m_{ph}^2 - p^2 x(1-x) - k^2 - i\varepsilon]^3}$$

$$= \frac{-g^2}{32\pi^2}\int_0^1 dx \frac{x(1-x)}{[m_{ph}^2 - p^2 x(1-x)]}$$

$$= \frac{g^2}{32\pi^2}\frac{\partial}{\partial p^2}\left\{ \int_0^1 dx \ln[m_{ph}^2 - p^2 x(1-x)] \right\}. \tag{3.1.11}$$

Using the condition $\Sigma_{1R} = 0$ at $p^2 = m_{ph}^2$ now gives

$$\Sigma_{1R} = \frac{g^2}{32\pi^2} \int_0^1 dx \ln \left[\frac{m_{ph}^2 - p^2 x(1-x)}{m_{ph}^2 (1 - x + x^2)} \right]. \tag{3.1.12}$$

This integral can be worked out analytically.

3.2 Higher order

The graph of Fig. 3.1.1 is not the only divergent graph in the theory. In Chapter 5 we will discuss the general theory of renormalization and we will see how to extend the removal of divergences to all orders. In this section we will only consider a class of graphs which have divergences generated because Fig. 3.1.1 occurs as a subgraph. Examples are given in Fig. 3.2.1.

One property should be clear. This is that the divergences come from subgraphs all of whose lines are part of a loop. A general way of characterizing these subgraphs is to define the concept of a one-particle-irreducible graph or subgraph. A one-particle-irreducible (1PI) graph is one which is connected and cannot be made disconnected by cutting a single line. A graph which is not 1PI is called one-particle-reducible (1PR). The graphs in Fig. 3.2.1 are all one-particle-reducible, since they all have one or more lines that when cut leave the graph in two disconnected pieces. The self-energy subgraph of Fig. 3.1.1 consisting of the two lines in the loop is 1PI. This identical subgraph occurs several times in the graphs of Fig. 3.2.1.

We introduced a mass counterterm into the interaction, so that the

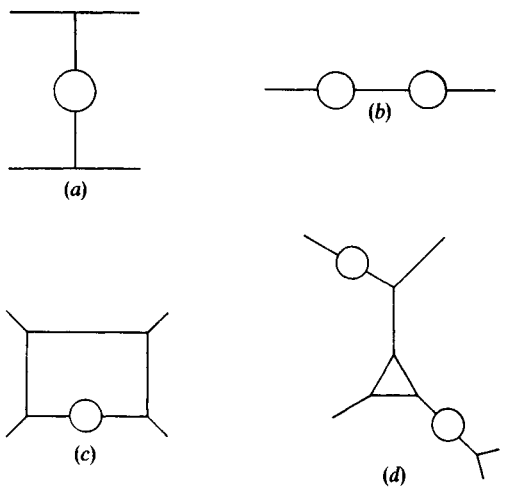

Fig. 3.2.1. Graphs containing the one-loop self-energy as a subgraph.

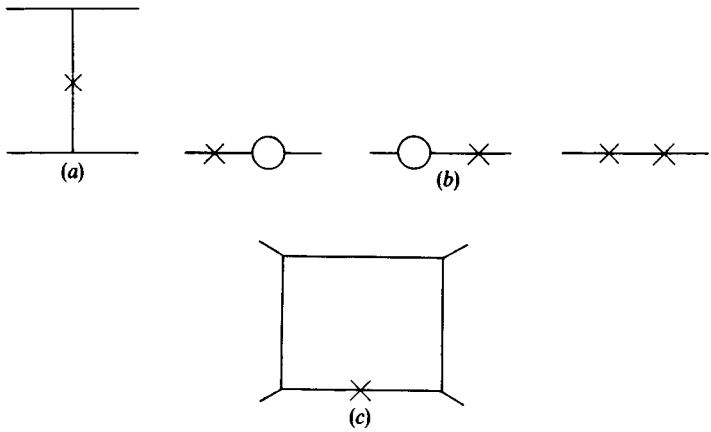

Fig. 3.2.2. Counterterm graphs to Fig. 3.2.1.

counterterm graph Fig. 3.1.6 cancels the divergence in Fig. 3.1.1. Clearly the counterterm vertex can be used as an interaction anywhere in a graph. In fact, all graphs containing it can be found as follows: (*a*) take a graph with the loop of Fig. 3.1.1 occurring as subgraph one or more times, but with no mass counterterm vertices; (*b*) replace one (or more) of the occurrences of the loop by the counterterm. The terms generated from Figs. 3.2.1 (*a*)–(*c*) are shown in Fig. 3.2.2.

Evidently, the sum of the original graph and its counterterm graph(s) is just the original graph with every occurrence of the loop replaced by its renormalized value $-i\Sigma_{1R}$. It is sensible to keep the counterterm associated with the loop, thereby considering the loop plus the counterterm as a single entity.

In the case of the graphs of Fig. 3.2.1 the result of this procedure is to make the graphs finite. The generalization to an arbitrary graph will be worked out in Chapter 5.

3.3 Degree of divergence

We saw how the Wick rotation ensured that the UV divergence of the one-loop self-energy is a purely Euclidean problem. The divergence is from the region $|k^\mu| \to \infty$, without any regard to direction. Thus, simple power-counting determined that there is a logarithmic divergence. The power-counting involves merely counting the powers of k in the integral for large k. The divergence is logarithmic as the lattice spacing, a, goes to zero.

Power-counting in this fashion works for a general graph to determine

what is called either the 'overall degree of divergence' or the 'superficial degree of divergence'. This we will discover in Chapter 5. There we will also see how the value of the degree of divergence determines the particular counterterm vertices needed for a theory. This will enable us to determine whether or not a theory is renormalizable by invoking arguments revolving around the dimension of the coupling.

In this section we will vary the dimensionality, d, of space-time in the calculation of the self-energy graph, Fig. 3.1.1. We can then explore the relation between the degree of divergence, as determined by power-counting, and the momentum dependence of the counterterm.

The integral for Fig. 3.1.1 is now

$$\Sigma_1(p^2, m^2, d) = \frac{ig^2}{2(2\pi)^d} \int d^d k \frac{1}{(k^2 - m^2 + i\varepsilon)[(p+k)^2 - m^2 + i\varepsilon]}. \quad (3.3.1)$$

The space-time has one time dimension and $d-1$ space dimensions. In the Feynman rules the factors $(2\pi)^4$ get replaced by $(2\pi)^d$, since they arise from the result

$$\int d^d x e^{ik\cdot x} = (2\pi)^d \delta^{(d)}(k). \quad (3.3.2)$$

The number of powers of k in the integral is now $d-4$; we call this the degree of divergence of the graph. If d is less than four, then the graph is convergent. But whenever d is greater than or equal to four, the graph diverges. Our discussion in the previous sections tells us we must try to renormalize it by adding a counterterm.

Now, differentiating once with respect to p^μ gives convergence if $d = 4$:

$$\frac{\partial \Sigma_1}{\partial p^\mu} = \frac{-ig^2}{(2\pi)^d} \int d^d k \frac{(p+k)_\mu}{(k^2 - m^2)[(p+k)^2 - m^2]^2}, \quad (3.3.3)$$

with the degree of divergence of the integral being reduced by one to $d-5$. If $d = 5$ the integral diverges logarithmically. However one might surmise that the divergence comes from the piece of the integrand proportional to k_μ and that symmetrizing by $k \to -k$ would kill the divergence. This is in fact true, but let us be more simple-minded.

Differentiating again with respect to p gives a result with degree of divergence $d-6$:

$$\frac{\partial^2 \Sigma_1}{\partial p^\mu \partial p^\nu} = \frac{ig^2}{(2\pi)^d} \int d^d k \frac{\{2(p+k)_\mu (p+k)_\nu - g_{\mu\nu}[(p+k)^2 - m^2]\}}{(k^2 - m^2)[(p+k)^2 - m^2]^3}. \quad (3.3.4)$$

To recover Σ_{1R} at $d = 5$ we integrate twice. There are more constants of integration that appear as an additive contribution of the form: $A + B_\mu p^\mu$.

However, we must require Lorentz invariance of Σ, so the B_μ term is eliminated, and we are left with a mass term as the only counterterm. This is the first and simplest example of the use of a symmetry argument to determine the form that we will allow for a counterterm.

3.3.1 ϕ^3 at $d = 6$

Let us now go to $d = 6$. Differentiating Σ_1 three times gives a finite result with degree of divergence $d - 7$. Integrating to obtain Σ_{1R} gives arbitrary integration constants of the form $A - Bp^2$, where again we have used Lorentz invariance. If we went to the lattice we would find divergences proportional to $1/a^2$, $m^2 \ln(a)$ and $p^2 \ln(a)$. The fact that these terms have dimension 2 corresponds to the fact that the integral for Σ_1 has degree of divergence 2. To make it finite we must not only use a mass counterterm but also a counterterm proportional to p^2; the total we will call $\delta m^2 - \delta Z p^2$. This is generated by a counterterm

$$-\delta m^2 \phi^2/2 + \delta Z(\partial\phi)^2/2 \tag{3.3.5}$$

in the Lagrangian.

Evidently the value of the degree of divergence is reflected as the maximum number of derivatives or powers of p in the counterterms. Equally, it is reflected in the integration constants that appear when we recover Σ_{1R} from the differentiated Σ_1. These two phenomena happen for a general graph, as we will see later. The method of proof will in fact be to differentiate each graph enough times with respect to its external momenta until it is finite.

The $(\partial\phi)^2$ counterterm in (3.3.5) is of course an example of the wave-function renormalization introduced in Section 2.3. We can interpret it physically by examining the propagator. The propagator for the bare field can be expressed in terms of the propagator of the renormalized field:

$$\begin{aligned}\tilde{G}_{2(0)} &\equiv \langle 0|T\tilde{\phi}_0(p)\phi_0(0)|0\rangle \\ &= Z\langle 0|T\tilde{\phi}(p)\phi(0)|0\rangle \\ &= iZ/(p^2 - m_{ph}^2 - \Sigma_{1(0)} - \delta m^2 + \delta Z p^2 + i\varepsilon) \\ &= iZ/(p^2 - m_{ph}^2 - \Sigma_{1R}).\end{aligned} \tag{3.3.6}$$

Note the distinction between bare and renormalized fields. The residue of the particle pole in the propagator of an interacting field is in general not unity:

$$\tilde{G}_{2(0)} = iR_{(0)}/(p^2 - m_{ph}^2 + i\varepsilon) + \text{finite}, \quad \text{as} \quad p^2 \to m_{ph}^2. \tag{3.3.7}$$

Examination of its spectral representation demonstrates that $0 \leq R(0) < 1$,

because the field ϕ_0 has canonical equal-time commutation relations. (The proof is given, for example, in Section 16.4 of Bjorken & Drell (1966).)

Now we can always change the definition of $\phi \equiv Z^{-1/2}\phi_0$ by multiplying Z by a finite factor. So it is possible to adjust Z so that the renormalized propagator has a pole of unit residue at $p^2 = m_{\text{ph}}^2$:

$$\tilde{G}_2 = \mathrm{i}/(p^2 - m_{\text{ph}}^2 + \mathrm{i}\varepsilon) + \text{finite}, \quad \text{as} \quad p^2 \to m_{\text{ph}}^2. \tag{3.3.8}$$

In this case we identify $R_{(0)}$ with Z. The renormalized self-energy satisfies

$$\Sigma_{1R} = \partial\Sigma_{1R}/\partial p^2 = 0 \quad \text{at} \quad p^2 = m_{\text{ph}}^2. \tag{3.3.9}$$

When integrating to obtain Σ_{1R} from the finite derivative of Σ_1, this condition enables the integration constants to be determined. It is called the mass-shell renormalization condition.

3.3.2 *Why may Z be zero and yet contain divergences?*

As a property of the exact theory we know that $0 \leq Z < 1$, if we adopt the renormalization condition (3.3.9); Z is definitely finite. However its perturbation expansion starts at $1 + g^2[C\ln(a) + \text{finite}] + \cdots$. (Here C is a constant.) This seems to be infinite as $a \to 0$, rather than finite. We resolve the contradiction by realizing that we should not expect higher-order terms to be small if the one-loop correction is large. For example, we could have the series

$$Z = \left[\frac{1}{1 - g^2[(C/D)\ln(a) + \text{const.}]}\right]^D, \tag{3.3.10}$$

where D is a positive number. Then $Z \to 0$ as $a \to 0$ even through the one-loop term goes to infinity. In any event we see that C must be positive (otherwise if we fix a and let $g \to 0$ there is a region with $Z > 1$).

It would appear impossible to derive a formula like (3.3.10) since it involves summing all orders of perturbation theory. Moreover it involves an analytic continuation from within its radius of convergence $|\ln(a)| < D/g^2C$ to $|\ln(a)| = \infty$. This seems to make no sense at all since perturbation series are in general asymptotic series rather than convergent series. However, we will see in Chapter 7, on the renormalization group, that we can find a systematic method of calculating Z and the other renormalizations in the limit $a \to 0$. It relies on the so-called asymptotic freedom of the theory. The behavior (3.3.10) will turn out to be essentially correct, in asymptotically free theories.

No matter what the truth is, it should be clear that the divergences in perturbation theory as $a \to 0$ need not be reflected as divergences in the exact theory, but only as singularities.

3.3.3 *Renormalizability and non-renormalizability*

Suppose we now go to a dimension $d > 6$. For example, let us set $d = 8$. Then the one-loop self-energy has a quartic divergence. The necessary counterterms contain up to four derivatives of the field:

$$-\delta m^2\phi^2/2 - \delta Z(\partial\phi)^2/2 + E(\Box\phi)^2/2. \tag{3.3.11}$$

The quartic term is not of the form of any term in the original Lagrangian, so it cannot be obtained by renormalization of the Lagrangian. When this situation occurs the theory is called non-renormalizable. Non-renormalizability is *a priori* a good reason for dropping the theory from consideration. There are possible ways to avoid this, but we will leave discussion of this until later.

There is a simple argument that helps in the determination of whether or not a theory is renormalizable. It links dimensional analysis and power counting. Consider a one-particle-irreducible graph Γ. Let its degree of divergence be $\delta(\Gamma)$, and let its mass dimension be $d(\Gamma)$. Now Γ is the product of a numerical factor, a set of couplings, and an integral. As we see from (3.3.1) the degree of divergence is the dimension of the integral. Therefore, if we let $\Delta(\Gamma)$ be the dimension of the couplings in Γ, we have

$$d(\Gamma) = \delta(\Gamma) + \Delta(\Gamma). \tag{3.3.12}$$

Now consider the counterterms to Γ. These form a polynomial of degree $\delta(\Gamma)$ in the external momenta. For each term C in the polynomial, let $\delta(C)$ be the number of derivatives and let $\Delta(C)$ be the dimension of the coefficient, so that its dimension is the same as Γ:

$$\delta(C) + \Delta(C) = d(\Gamma). \tag{3.3.13}$$

Now, the maximum number of derivatives in the counterterms is

$$\delta(\Gamma) = d(\Gamma) - \Delta(\Gamma).$$

If the couplings have negative dimension, then $\delta(\Gamma)$ can be made arbitrarily large by going to a graph of high enough order. In the absence of miraculous cancellations this tells us to expect non-renormalizability. If the couplings have zero or positive dimension, we have a finite number of counterterm vertices, since $-\Delta(\Gamma)$ is bounded above by zero and $d(\Gamma)$ decreases as the number of external lines increases.

If the couplings never have negative dimension, we observe that the coefficients of the counterterms satisfy

$$\begin{aligned}\Delta(C) &= d(\Gamma) - \delta(C)\\ &\geq d(\Gamma) - \delta(\Gamma)\\ &= \Delta(\Gamma),\end{aligned}$$

so that these coefficients also have non-negative dimension. This, in the simplest cases, is sufficient to ensure renormalizability.

For example, in the ϕ^3 theory ϕ has dimension $d/2 - 1$. (Recall that $\mathscr{L}$ has the dimension of an energy-density.) Then $\Delta(g) = 3 - d/2$. So if $d > 6$ the theory is non-renormalizable, as we saw by example. If $d \leq 6$, there are only a finite number of possible counterterms, since these are restricted to have coefficients of non-negative dimension.

3.4 Renormalization group

3.4.1 *Arbitrariness in a renormalized graph*

The infinities of a renormalizable theory amount to divergent dynamical contributions that renormalize the parameters in the Lagrangian. Traditionally one thinks of renormalization as the procedure of working with measured quantities instead of the corresponding bare quantities. The most obvious case is that of the mass m_{ph} of the particle corresponding to an elementary field. However to take the traditional view is much too restrictive.

This issue can be understood by looking at strong interactions. There we have a theory, QCD, in which free particles corresponding to the elementary fields do not appear to exist (*pace* LaRue, Phillips & Fairbank (1981)). So arises the hypothesis of quark confinement – not proved from QCD, so far – according to which quarks are never isolated particles. Even so, the theory has quark masses, which can be measured (up to considerable uncertainties for the light quarks). But one cannot identify these masses with the directly measurable masses of free quarks. One must only speak of mass parameters, measured, in this case, rather indirectly.

There is no problem in taking this point of view. For example, we write the ϕ^3 Lagrangian as a basic Lagrangian plus a counterterm Lagrangian:

$$\mathscr{L} = \partial\phi^2/2 - m^2\phi^2/2 - g\phi^3/3! - \delta Z\partial\phi^2/2 - \delta m^2\phi^2/2 - \delta g\phi^3/3!. \tag{3.4.1}$$

But we avoid identifying the renormalized mass with the mass m_{ph} of a particle. Similarly we do not identify the renormalized coupling, g, with any specific measured quantity, and we do not define Z by requiring that the residue of the propagator's pole be unity.

Consider the calculation of the one-loop self-energy at $d = 4$. We can choose $\delta Z = 0$ for this case; the only divergence is in δm^2. The renormalized self-energy (3.1.7) is

$$\Sigma_{1R} = \Sigma_{1\text{fin}} + (\delta m^2 + \Sigma_{1\text{div}}).$$

We must choose δm^2 to have a divergent part to cancel the divergence at $a = 0$ of $\Sigma_{1\text{div}}$. But the finite part of δm^2 is not determined; the arbitrariness is the same as that of the integration constant when obtaining Σ_{1R} by integrating $\partial\Sigma_1/\partial p^\mu$.

At first sight it might appear that the arbitrariness ruins the theory unless one pins down m to be the physical mass m_{ph}. This is in fact not so; the arbitrariness is more like the arbitrariness in choosing a coordinate system. Suppose one first computes the propagator with the mass-shell condition $\Sigma_{1R}(p^2 = m_{\text{ph}}^2) = 0$. Then

$$\bar{G}_2 = \mathrm{i}/[p^2 - m_{\text{ph}}^2 - \Sigma_{1\text{fin}}(p^2) + \Sigma_{1\text{fin}}(p^2 = m_{\text{ph}}^2) + O(g^4)], \quad (3.4.2a)$$

$$m_0^2 = m_{\text{ph}}^2 - \Sigma_{1\text{div}} - \Sigma_{1\text{fin}}(m_{\text{ph}}^2) + O(g^4). \quad (3.4.2b)$$

One could also compute with a different finite part to δm^2, with a result

$$\bar{G}_2 = \mathrm{i}/[p^2 - m^2 - \Sigma_{1\text{fin}}(p^2, m^2) + g^2 C + O(g^4)], \quad (3.4.3a)$$

$$m_0^2 = m^2 - \Sigma_{1\text{div}} - g^2 C + O(g^4), \quad (3.4.3b)$$

where C is any chosen number. The self-energy is now

$$\Sigma_{1R}^{(C)}(p^2, m^2) = \Sigma_{1\text{fin}} + g^2 C.$$

Evidently the two ways of renormalizing the theory give the same results if we require that the bare mass m_0^2 is the same in both of (3.4.2) and (3.4.3). In the complete solution of the theory, say by the functional integral, it is only m_0^2 that matters, not the partition into a renormalized mass squared m^2 and a counterterm $-\Sigma_{1\text{div}} - g^2 C$. Clearly we have

$$m^2 = m_{\text{ph}}^2 + g^2 C - \Sigma_{1\text{fin}}(m_{\text{ph}}^2) + O(g^4), \quad (3.4.4)$$

with the $O(g^4)$ terms depending on the renormalization of higher-order self-energy graphs.

We come then to the central idea of the renormalization group. The arbitrariness in the definition of Σ_{1R} is physically irrelevant, for a change in the arbitrary constant C can be exactly compensated by a change in m^2. A change in C merely gives a different parametrization of the set of theories that can be obtained by varying the mass parameter m. The renormalization group is the set of transformations on the parametrizations of the theory. The transformations are accomplished by moving parts of the terms in $\mathscr{L}$ from the basic Lagrangian to the counterterm Lagrangian. In the case of m it is a move from the free Lagrangian to the interaction Lagrangian. This of course gives a rearrangement of the perturbation series, which is the key to the many practical applications of the renormalization group.

It might be objected that

$$p^2 - m^2 - \Sigma_{1\text{fin}}(p^2, m^2) + g^2 C$$

is not equal to

$$p^2 - m_{\rm ph}^2 - \Sigma_{1\rm fin}(p^2, m_{\rm ph}^2) + \Sigma_{1\rm fin}(m_{\rm ph}^2, m_{\rm ph}^2),$$

since the mass parameters in $\Sigma_{1\rm fin}(p^2)$ are different, whereas the theory parametrized in either way is the same. But since $m_{\rm ph}^2 - m^2$ is $O(g^2)$ the difference in the two expressions is in fact $O(g^4)$. Thus the rearrangement of the perturbation series does not leave the p^2-dependence of the coefficients invariant. The $O(g^4)$ terms will cancel the difference (up to even higher-order terms), etc.

The utility of the renormalization group is precisely in its ability to reorganize the perturbation series. Since one effect of the interaction is to induce dynamical contributions to the mass and couplings, it is evidently a good idea to arrange that these contributions are small. The result is to reduce the values of higher-order corrections and thus improve the reliability of a perturbative calculation.

Now the effective size of the dynamical mass or coupling must be treated as dependent on the situation under consideration. This can be seen by examining Σ_{1R} given in (3.1.12) at large p^2:

$$\Sigma_{1R} \sim (g^2/32\pi^2)[\ln(-p^2/m_{\rm ph}^2) + \text{constant} + \cdots]. \tag{3.4.5}$$

If $|p^2|$ is large enough this can be large. Since the graph occurs as a subgraph of higher-order graphs, it is likely (and often is true) that higher-order graphs are as important as low-order graphs at large enough p^2. This situation is undesirable and can be remedied by a renormalization-group transformation.

We absorb the large part of Σ_{1R} into a redefinition of the renormalized mass m^2. We must examine higher-order graphs at large p^2 to demonstrate that there are no further sources of large coefficients. We will do this systematically in Chapter 7.

3.4.2 *Renormalization prescriptions*

There are infinitely many ways of resolving the ambiguity in constructing the counterterms for a given theory, each of these ways corresponding to a particular parametrization. It is essential that, whenever a particular divergent graph occurs as a subgraph of a bigger graph, the ambiguity is resolved in the same way at each occurrence, since the corresponding counterterm vertex is generated by a single term in the Lagrangian. So to perform concrete calculations one adopts some rule to resolve the ambiguity. Such a rule is called a renormalization prescription or renormalization scheme.

Of the infinitely many possible renormalization prescriptions, a few have become standard, because they are especially convenient either for practical use or for theoretical considerations. In this section we will explain two of the standard ones with the aid of the example of the one-loop graph Fig. 3.1.1.

We have already encountered the mass-shell, or physical, scheme. The renormalized mass is defined to be the physical mass, i.e., the position of the propagator pole. Wave-function renormalization is fixed by requiring the residue of the pole to be unity (see (3.3.9)). Couplings can be defined by specifying the value of a suitable S-matrix element.

A possibility that is much used in discussions of renormalization theory is the BPH or BPHZ scheme (Bogoliubov–Parasiuk–Hepp–Zimmermann), otherwise known as zero-momentum subtraction. Let Γ be a one-particle-irreducible (1PI) graph that is divergent, i.e., it has $\delta(\Gamma) \geq 0$. The prescription is that at zero external momentum its renormalized value $R(\Gamma)$ and its first $\delta(\Gamma)$ derivatives with respect to external momentum are zero. The BPHZ scheme is to implement this by subtracting off the first $\delta(\Gamma)$ terms in the Taylor expansion of the integrand (Zimmermann (1970)), that is, the renormalization is performed before the integration over loop momenta. No explicit UV cut-off is needed. In this scheme the self-energy already discussed is, at $d = 4$,

$$\Sigma_{1R}^{(BPHZ)} = \frac{ig^2}{32\pi^4}\int d^4k \left\{ \frac{1}{[(k^2 - m^2)((p+k)^2 - m^2)]} - \frac{1}{(k^2 - m^2)^2} \right\}, \quad (3.4.6)$$

while at $d = 6$, we have:

$$\Sigma_{1R}^{(BPHZ)} = \frac{ig^2}{128\pi^6}\int d^6k \left\{ \frac{1}{(k^2 - m^2)[(p+k)^2 - m^2]} - \frac{1}{(k^2 - m^2 + i\varepsilon)^2} + \frac{2p\cdot k}{(k^2 - m^2)^3} - \frac{[4p\cdot k^2 - (k^2 - m^2)p^2]}{(k^2 - m^2)^4} \right\}. \quad (3.4.7)$$

3.5 Dimensional regularization

In our initial treatment of UV divergences we used the lattice as a cut-off, or regulator. However, what we are really interested in is the renormalized theory with no cut-off. We could equally well use some other kind of regulator. For example, a Pauli–Villars type of cut-off is achieved by a higher derivative term in the Lagrangian. For example, from

$$\mathscr{L} = \partial A^2/2 - m^2 A^2/2 - [(\Box + m^2)A]^2/2(M^2 - m^2) - gA^3/3! + \text{counterterms} \quad (3.5.1)$$

we obtain the free propagator (2.4.1). When the cut-off M goes to infinity, the propagator is the ordinary one, but when p^2 is much bigger than M^2, it is smaller by a factor p^2/M^2. Thus UV divergences are cut-off for the theory in six or fewer space-time dimensions; the divergences reappear when we take the limit $M \to \infty$.

If we defined the theory by a functional integral the lattice would appear as an intermediate step, but the $a \to 0$ limit would give no divergences, if M is finite. Although the Euclidean Green's functions for the cut-off theory (3.5.1) exist, the Minkowski space field theory is not physical. A symptom of this is that the pole of the free propagator at $p^2 = M^2$ has the wrong sign of residue; it implies a particle with negative metric.

A theory with no cut-off can be obtained by adding counterterms with appropriate M-dependences to cancel the divergences and then taking the $M \to \infty$ limit. As an example we showed that counterterms cancelled the divergences of the one-loop self-energy graph. Although we assumed a lattice regulator, we used no properties of the lattice propagator that are not true for the Pauli–Villars case. We assumed only that:

(1) If the cut-off is taken away (i.e., $M \to \infty$ or $a \to 0$) with p and m fixed, then the propagator goes to $i/(p^2 - m^2)$.
(2) If $p^2 \to \infty$ with fixed cut-off, then the propagator is sufficiently much smaller than $1/p^2$ that the graph is not UV divergent.
(3) In the Euclidean region there are no propagator poles.

In principle, any method of imposing a UV cut-off is good enough, but in practice some methods are more convenient than others. For most purposes dimensional regularization is the most convenient. The method starts from the observation that UV divergences are eliminated by going to a small enough space-time dimension d. We can use the space-time dimension as a regulator provided we treat d as a continuous variable (so that the cut-off can be removed by taking the limit $d \to 4$). This idea has a long history, but its popularity roughly started after the papers by Wilson (1973) in statistical mechanics and by 't Hooft & Veltman (1972a), Bollini & Giambiagi (1972), Ashmore (1972), and Cicuta & Montaldi (1972) in field theory (especially non-abelian gauge theories).

Since vector spaces of non-integer dimension do not exist as such, it is not obvious that the concept has any consistency, let alone validity, even in a purely formal sense. This we will remedy in the next chapter. For the present we will assume uniqueness and existence, and apply standard manipulations to the integral

$$\Sigma_1(p^2, m^2, d) = \frac{ig^2}{2(2\pi)^d} \int d^d k \frac{1}{(k^2 - m^2 + i\varepsilon)[(p+k)^2 - m^2 + i\varepsilon]} \qquad (3.5.2)$$

until it is of a form where we can sensibly assign a value. We will (following Wilson (1973)) express (3.5.2) in terms of a standard Gaussian integral:

$$\int \mathrm{d}^d k \exp(k^2) = \mathrm{i} \int \mathrm{d}\omega \int \mathrm{d}^{d-1} k \exp(-\omega^2 - k^2). \tag{3.5.3}$$

It is sensible to give this the value $\mathrm{i}\pi^{d/2}$, which is correct if d is an integer. In the following calculation of the value of (3.5.2) with non-integer d, the assumed properties of the integration are italicized. All the manipulations are valid for any integer value of d for which the integral converges.

We use the Schwinger representation for each propagator:

$$1/(m^2 - k^2 - \mathrm{i}\varepsilon) = \int_0^\infty \mathrm{d}a \exp[-a(m^2 - k^2 - \mathrm{i}\varepsilon)]. \tag{3.5.4}$$

Observe that because of the Wick rotation we treat k^2 as negative. Then we *exchange the order of integration* to obtain

$$\Sigma_1 = \frac{\mathrm{i}g^2}{2(2\pi)^d} \int_0^\infty \mathrm{d}a \int_0^\infty \mathrm{d}b \int \mathrm{d}^d k \exp[-(a+b)m^2 + bp^2 + 2bp\cdot k + (a+b)k^2]. \tag{3.5.5}$$

We *shift* k^μ by an amount $p^\mu b/(a+b)$ and change variables to $z = a + b$, $x = a/z$ to get:

$$\Sigma_1 = \frac{\mathrm{i}g^2}{2(2\pi)^d} \int_0^1 \mathrm{d}x \int_0^\infty \mathrm{d}z\, z \int \mathrm{d}^d k \exp\{-z[m^2 - p^2 x(1-x)] + zk^2\}.$$

After *scaling* k by a factor $z^{1/2}$ we find that

$$\Sigma_1 = \frac{\mathrm{i}g^2}{2(2\pi)^d} \int_0^1 \mathrm{d}x \int_0^\infty \mathrm{d}z\, z^{1-d/2} \exp\{-z[m^2 - p^2 x(1-x)]\} \int \mathrm{d}^d k \exp(k^2). \tag{3.5.6}$$

It is this stage which brings in the dimensionality d. We have now reduced the d-dimensional integral to the form (3.5.3), which we defined to be $\mathrm{i}\pi^{d/2}$. The z-integral in (3.5.6) gives a Γ-function, so we finally obtain:

$$\Sigma_1 = \frac{-g^2}{2(4\pi)^{d/2}} \Gamma(2 - d/2) \int_0^1 \mathrm{d}x [m^2 - p^2 x(1-x)]^{d/2-2}. \tag{3.5.7}$$

This result is unique (except possibly for an overall normalization, which is universal – the same for every d-dimensional integral). The divergences now reside in the Γ-function which has simple poles at $d = 4, 6, 8, \ldots$. The residue of each pole is a polynomial in p of degree equal to the degree of divergence.

One of the main advantages of dimensional regularization is immediately apparent. Not only was the integral unchanged from its form in a theory in

integer dimensional space-time with no cut-off, but the method of calculation was unchanged. Use of the representation (3.5.4) is an efficient way of obtaining a parametric representation like (3.5.7) for a Feynman graph.

A second advantage that we will see later is that it preserves not only Poincaré invariance in the regulated theory, but also gauge symmetries. This was a main motivation for its use by 't Hooft & Veltman (1972a) and many others. Most methods of introducing a cut-off fail in this respect. (For example, gauge invariance is preserved on the lattice but full Poincaré invariance is lost.)

A third advantage – also of great importance in practice – is that a continuous space-time dimension is also a gauge-invariant cut-off for infrared divergences in theories with massless fields (Gastmans & Meuldermans (1973), Gastmans, Verwaest & Meuldermans (1976), and Marciano & Sirlin (1975)). A trivial example, but without any gauge invariance, is given by the ϕ^3 self-energy with $m=0$ at $d=2$. It is

$$\frac{ig^2}{8\pi^2}\int d^2k \frac{1}{(k^2+i\varepsilon)[(p+k)^2+i\varepsilon]},$$

which is divergent at $k=0$ and at $p=-k$. The divergence is regulated by increasing d. Care is required in using this method if UV divergences are present in the same graph, for they are regulated by reducing d.

3.6 Minimal subtraction

3.6.1 *Definition*

From the unrenormalized self-energy (3.5.7) we compute the renormalized self-energy Σ_{1R} at $d=4$ by adding a mass counterterm $\delta m^2(g,m^2,d)$ and then letting $d\to 4$. Suppose we choose m to be the physical mass. Then

$$\begin{aligned}\Sigma_{1R}^{(ph)} &= \Sigma_1(p^2,m_{ph}^2,g,d)-\Sigma_1(m_{ph}^2,m_{ph}^2,g,d)\\ &= \frac{-g^2}{2(4\pi)^{d/2}}\Gamma(2-d/2)\int_0^1 dx\{[m_{ph}^2-p^2x(1-x)]^{d/2-2}\\ &\qquad -[m_{ph}^2(1-x+x^2)]^{d/2-2}\}. \qquad (3.6.1)\end{aligned}$$

Now $\Gamma(z)$ has a pole at $z=0$:

$$\Gamma(z)=1/z-\gamma_E+O(z), \qquad (3.6.2)$$

where $\gamma_E=0.5772\ldots$ is Euler's constant. So at $d=4$

$$\Sigma_{1R}^{(ph)}=\frac{g^2}{32\pi^2}\int_0^1 dx \ln\left\{\frac{m_{ph}^2-p^2x(1-x)}{m_{ph}^2(1-x+x^2)}\right\} \qquad (3.6.3)$$

in agreement with our earlier calculation.

Since the divergence in Σ_1 amounts to a simple pole at $d=4$, a rather obvious way of renormalizing it is to define δm^2 to cancel just the singularity, i.e., the pole ('t Hooft (1973)). This, of course, means that we are changing our renormalization prescription. By referring back to (3.5.7), we see that δm^2 in this scheme is:

$$\delta m^2 = (g^2/32\pi^2)/(2-d/2), \tag{3.6.4}$$

from which we get Σ_{1R} by expanding Σ_1 in a power series in $d-4$. We find

$$\Sigma_{1R} = \frac{g^2}{32\pi^2}\int_0^1 dx \left\{ \ln\left[\frac{m^2 - p^2x(1-x)}{4\pi}\right] + \gamma_E \right\}. \tag{3.6.5}$$

Unfortunately this contains the logarithm of a dimensional quantity. The reason is that in the expansion in powers of $d-4$ we did not allow for the fact that g has a dimension dependent on d. Therefore we implicitly introduced a mass scale.

To make this scale explicit, we rewrite the coupling

$$g \to \mu^{2-d/2} g, \tag{3.6.6}$$

where we have introduced a parameter μ with the dimensions of mass, called the unit of mass ('t Hooft (1973)). The redefined coupling g now has fixed dimension equal to 1, and the renormalized self-energy becomes

$$\Sigma_{1R}^{(MS)} = \frac{g^2}{32\pi^2}\int_0^1 dx \left\{ \ln\left[\frac{m^2 - p^2x(1-x)}{4\pi\mu^2}\right] + \gamma_E \right\}. \tag{3.6.7}$$

We derived this by observing that

$$\begin{aligned}\mu^{2-d/2} &= e^{(2-d/2)\ln\mu} \\ &= 1 + (2-d/2)\ln\mu + \tfrac{1}{2}(2-d/2)^2\ln^2\mu + \cdots.\end{aligned}$$

This renormalization prescription, where counterterms are pure poles at the physical value of d, is called minimal subtraction (MS).

The unit of mass μ is entirely arbitrary. Thus the self-energy (3.6.7) now depends on three parameters instead of two. However a change of μ amounts to a change of renormalization prescription, so the change can be compensated, in this case, by a change in m. In effect minimal subtraction is a one-parameter family of renormalization prescriptions.

3.6.2 $d=6$

We can also apply minimal subtraction to the six-dimensional theory. There we define

$$\left.\begin{aligned}\delta m^2 &= \text{poles at } d=6, \\ \delta z &= \text{poles at } d=6, \\ \delta g &= \mu^{3-d/2}(\text{poles at } d=6).\end{aligned}\right\} \tag{3.6.8}$$

Note that the renormalized coupling is now dimensionless. Since

$$\Gamma(z-1) = -1/z + \gamma_E - 1 + O(z) \tag{3.6.9}$$

as $z \to 0$, we find that

$$\Sigma_{1R}^{(MS)} = \frac{-g^2}{128\pi^3}\left\{(m^2 - p/6)(\gamma_E - 1) + \int_0^1 dx\,[m^2 - p^2x(1-x)]\ln\left[\frac{m^2 - p^2x(1-x)}{4\pi\mu^2}\right]\right\} \tag{3.6.10}$$

and

$$\delta m^2 = \frac{g^2}{64\pi^3}\frac{m^2}{d-6} + O(g^4), \tag{3.6.11}$$

$$\delta Z = \frac{g^2}{6 \times 64\pi^3}\frac{1}{d-6} + O(g^4). \tag{3.6.12}$$

The counterterm δg for the coupling can also be calculated. From the graph of Fig. 3.6.1, we find (Macfarlane & Woo (1974))

$$\delta g = \mu^{3-d/2}\frac{g^3}{64\pi^3(d-6)} + O(g^5). \tag{3.6.13}$$

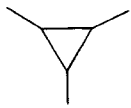

Fig. 3.6.1. One-loop vertex graph in ϕ^3 theory.

3.6.3 *Renormalization group and minimal subtraction*

When we discuss the renormalization group in Chapter 7, we will focus on one particular subgroup. The transformations in this subgroup consist of multiplying μ by a factor and making compensating changes in the renormalized coupling and mass. As a group it is trivial – being a representation of the group of positive real numbers under multiplication. What is non-trivial is the way in which it is represented in relation to the parametrization of the theory by a renormalized coupling and mass.

The renormalization group can be exploited in calculating high-energy behavior. While a full treatment will be made in Chapter 7, the basic idea can be seen by examining the one-loop self-energy. Let p^2 get large (with m and μ fixed), and consider the propagator defined using minimal subtraction at $d = 6$:

$$\bar{G}_2 \sim i\Big/\left(p^2\left\{1 - \frac{g^2}{768\pi^3}\left[\ln\left(\frac{-p^2}{\mu^2}\right) + \text{const.}\right] + O(g^4)\right\}\right). \tag{3.6.14}$$

To be able to make use of a perturbation expansion we must keep higher-order corrections small. But this is not so in (3.6.14) if p^2 is too large. The large correction can be avoided by setting μ^2 to be of order $|p^2|$. The theory is unchanged if we make suitable changes in g, in m, and in the scale of the field. We will learn how to do this in Chapter 7, with the result that the large corrections are effectively moved from higher-order terms in the perturbation series to the lowest-order graphs.

3.6.4 *Massless theories*

Let us return (for simplicity) to the self-energy of the four-dimensional theory. Consider the limit $m \to 0$. If we use mass-shell subtraction, we have (3.6.3), which diverges as $m_{\text{ph}}^2 \to 0$.

The divergence is an artifact of the mass-shell scheme, for which

$$\delta m_{\text{ph}}^2 = \frac{g^2}{2(4\pi)^{d/2}}\Gamma(2-d/2)m_{\text{ph}}^{d-4}\int_0^1 \mathrm{d}x(1-x+x^2)^{d/2-2}$$
$$= \frac{g^2}{32\pi^2}\left[\frac{1}{2-d/2} - \ln(m_{\text{ph}}^2) + \text{finite (as } m_{\text{ph}} \to 0, d \to 4)\right]. \quad (3.6.15)$$

In addition to the pole needed to cancel the UV divergence, there is a $\ln(m_{\text{ph}}^2)$ term.

Physically what happens is that in a massless theory there are long-range forces. These mean that separated particles are never completely free of each others' influence. Thus, for example, the singularity in the propagator is not a simple pole, for the self-energy (with MS subtraction) is, from (3.6.5),

$$\frac{g^2}{32\pi^2}\left[\ln\left(\frac{-p^2}{4\pi\mu^2}\right) + \gamma_{\text{E}} - 2\right]. \quad (3.6.16)$$

The mass-shell renormalization prescription relies on the assumption of a simple propagator pole to generate counterterms, so it must fail. However, the nature of the propagator's singularity is an infra-red problem, so it is irrelevant to the question of whether an ultra-violet divergence can be renormalized. Some other renormalization scheme, like minimal subtraction, must be used in the massless theory.

3.7 Coordinate space

A good way to understand the infinite renormalizations is to work in coordinate space, as was emphasized by Bogoliubov & Shirkov (1980). This point of view is especially useful in treating field theories at finite

temperature or on a curved space-time background, as we will see in Chapter 11. There we will see why the counterterms are the same as at zero temperature in flat space-time.

For most ordinary calculations, it is cumbersome to work in coordinate space, because the propagator $S_F(x)$ for a free massive field is a Bessel function. In momentum space the propagator is simple: $i/(p^2 - m^2 + i\varepsilon)$. However the asymptotic behavior of $S_F(x)$ as $x \to 0$ is simple.

We have

$$S_F = \frac{\Gamma(d/2-1)}{4\pi^{d/2}(-x^2)^{d/2-1}} + \text{less singular as } x \to 0. \tag{3.7.1}$$

The one-loop correction to the propagator $G_2(x, y)$ is

$$\begin{aligned} G_{2,2} &= -(g^2/2)\int d^d z \int d^d w S_F(x-z) S_F(z-w)^2 S_F(w-y) \\ &= \int d^d z \int d^d w S_F(x-z)\hat{\Sigma}_1(z-w) S_F(w-y), \end{aligned} \tag{3.7.2}$$

where $\hat{\Sigma}_1(z-w)$ is the self-energy in coordinate space:

$$\hat{\Sigma}_1(z-w) = -(g^2/2) S_F(z-w)^2. \tag{3.7.3}$$

This is singular at $z = w$, and causes a logarithmic divergence in (3.7.2), where we integrate over all z and w. The fact that the divergence is from the region $z \simeq w$ means that it is in fact a δ-function:

$$\begin{aligned} G_{2,2} = \int d^d z S_F(x-z) S_F(z-y) \\ \cdot \left\{ \int_{w \sim 0} d^d w \left(\frac{-g^2}{32\pi^4} \right) \frac{\Gamma(d/2-1)^2}{[-(w-z)^2]^{d-2}} \right\} + \text{finite.} \end{aligned} \tag{3.7.4}$$

By Wick rotating the w^0-integral and using the following result (next chapter) for the d-dimensional integral of a Lorentz-invariant function

$$\int d^d w f(w^2) = -i[2\pi^{d/2}/\Gamma(d/2)] \int_0^\infty dw w^{d-1} f(w^2), \tag{3.7.5}$$

we find

$$G_{2,2} = \int d^d z S_F(x-z) S_F(z-y) \frac{ig^2}{16\pi^2(4-d)} + \text{finite}, \tag{3.7.6}$$

as $d \to 4$.

Evidently, the divergence is cancelled by adding a δ-function to $\hat{\Sigma}_1$:

$$\hat{\Sigma}_{1R} = \Sigma_1 - \frac{ig^2}{16\pi^2(4-d)} \delta^{(d)}(w-z), \tag{3.7.7}$$

which is exactly the mass counterterm computed earlier by momentum-space methods.

The important point, which is in fact true for an arbitrary graph, is that any UV divergence comes from a region in coordinate space where several interactions occur very close to each other. The divergence can then be cancelled by a counterterm which is a δ-function in the positions of these interactions. If the divergence is worse than logarithmic, then the counterterm will include derivatives of the δ-function. In any event the fact that it is a δ-function means that the counterterm can be included as a local interaction in the action. The locality means that it is a product of fields at the same point.

Since the singularity at $x = 0$ of the free propagator $S_F(x)$ is independent of the boundary conditions used to define it, we should expect, for example, that the counterterms used in thermal field theory are the same as at zero temperature. At non-zero temperature, the vacuum is replaced by a mixed state, and the boundary conditions for $S_F(x)$ change. The momentum-space proof that the counterterms are temperature-independent is therefore made difficult, but the coordinate-space proof is unchanged.