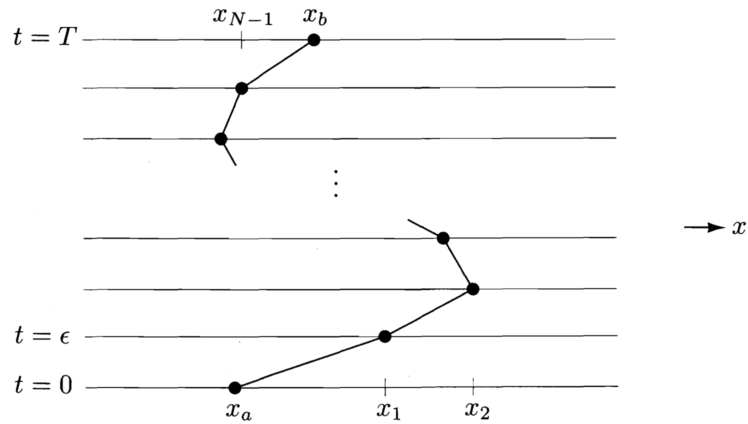


Figure 9.2. We define the path integral by dividing the time interval into small slices of duration ϵ, then integrating over the coordinate x_k of each slice.

We then define the path integral by

$$\int \mathcal{D}x(t) \equiv \frac{1}{C(\epsilon)} \int \frac{dx_1}{C(\epsilon)} \int \frac{dx_2}{C(\epsilon)} \cdots \int \frac{dx_{N-1}}{C(\epsilon)} = \frac{1}{C(\epsilon)} \prod_k \int_{-\infty}^{\infty} \frac{dx_k}{C(\epsilon)}, \tag{9.4}$$

where $C(\epsilon)$ is a constant, to be determined later. (We have included one factor of $C(\epsilon)$ for each of the N time slices, for reasons that will be clear below.) At the end of the calculation we take the limit $\epsilon \to 0$. (As in Sections 4.5 and 6.2, the $\prod$ symbol is an instruction to write what follows once for each k.)

Using (9.4) as the definition of the right-hand side of (9.3), we will now demonstrate the validity of (9.3) for a general one-particle potential problem. To do this, we will show that the left- and right-hand sides of (9.3) are obtained by integrating the same differential equation, with the same initial condition. In the process, we will determine the constant $C(\epsilon)$.

To derive the differential equation satisfied by (9.4), consider the addition of the very last time slice in Fig. 9.2. According to (9.3) and the definition (9.4), we should have

$$U(x_a, x_b; T) = \int_{-\infty}^{\infty} \frac{dx'}{C(\epsilon)} \exp\left[\frac{i}{\hbar} \frac{m(x_b - x')^2}{2\epsilon} - \frac{i}{\hbar} \epsilon V\left(\frac{x_b + x'}{2}\right)\right] U(x_a, x'; T - \epsilon).$$

The integral over x' is just the contribution to $\int \mathcal{D}x$ from the last time slice, while the exponential factor is the contribution to $e^{iS/\hbar}$ from that slice. All contributions from previous slices are contained in $U(x_a, x'; T - \epsilon)$.

As we send $\epsilon \to 0$, the rapid oscillation of the first term in the exponential constrains x' to be very close to x_b. We can therefore expand the above expression in powers of $(x'-x_b)$:

$$U(x_a,x_b;T) = \int_{-\infty}^{\infty} \frac{dx'}{C} \exp\left(\frac{i}{\hbar}\frac{m}{2\epsilon}(x_b-x')^2\right)\left[1 - \frac{i\epsilon}{\hbar}V(x_b) + \cdots\right]$$
$$\times \left[1 + (x'-x_b)\frac{\partial}{\partial x_b} + \frac{1}{2}(x'-x_b)^2\frac{\partial^2}{\partial x_b^2} + \cdots\right]U(x_a,x_b;T-\epsilon). \tag{9.5}$$

We can now perform the x' integral by treating the exponential factor as a Gaussian. (Properly, we should introduce a small real term in the exponent for convergence; we will ignore this term until the next section, when we derive Feynman rules using functional methods.) Recall the Gaussian integration formulae

$$\int d\xi\, e^{-b\xi^2} = \sqrt{\frac{\pi}{b}}, \qquad \int d\xi\,\xi\, e^{-b\xi^2} = 0, \qquad \int d\xi\,\xi^2\, e^{-b\xi^2} = \frac{1}{2b}\sqrt{\frac{\pi}{b}}.$$

Applying these identities to (9.5), we find

$$U(x_a,x_b;T) = \left(\frac{1}{C}\sqrt{\frac{2\pi\hbar\epsilon}{-im}}\right)\left[1 - \frac{i\epsilon}{\hbar}V(x_b) + \frac{i\epsilon\hbar}{2m}\frac{\partial^2}{\partial x_b^2} + \mathcal{O}(\epsilon^2)\right]U(x_a,x_b,T-\epsilon).$$

This expression makes no sense in the limit $\epsilon \to 0$ unless the factor in parentheses is equal to 1. We can therefore identify the correct definition of C:

$$C(\epsilon) = \sqrt{\frac{2\pi\hbar\epsilon}{-im}}. \tag{9.6}$$

Given this definition, we can compare terms of order ϵ and multiply by $i\hbar$ to obtain

$$\begin{aligned} i\hbar\frac{\partial}{\partial T}U(x_a,x_b;T) &= \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_b^2} + V(x_b)\right]U(x_a,x_b;T) \\ &= HU(x_a,x_b,T). \end{aligned} \tag{9.7}$$

This is the Schrödinger equation. But it is easy to show that the time-evolution operator U, as originally defined in (9.1), satisfies the same equation.

As $T \to 0$, the left-hand side of (9.3) tends to $\delta(x_a - x_b)$. Compare this to the value of (9.4) in the case of one time slice:

$$\frac{1}{C(\epsilon)}\exp\left[\frac{i}{\hbar}\frac{m(x_b-x_a)^2}{2\epsilon} + \mathcal{O}(\epsilon)\right].$$

This is just the peaked exponential of (9.5), and it also tends to $\delta(x_a - x_b)$ as $\epsilon \to 0$. Thus the left- and right-hand sides of (9.3) satisfy the same differential equation with the same initial condition. We conclude that the Hamiltonian definition of the time evolution operator (9.1) and the path-integral definition (9.3) are equivalent, at least for the case of this simple one-dimensional system.

To conclude this section, let us generalize our path-integral formula to more complicated quantum systems. Consider a very general quantum system, described by an arbitrary set of coordinates q^i, conjugate momenta p^i, and Hamiltonian $H(q,p)$. We will give a direct proof of the path-integral formula for transition amplitudes in this system.

The transition amplitude that we would like to compute is

$$U(q_a, q_b; T) = \langle q_b | \, e^{-iHT} \, | q_a \rangle \, . \tag{9.8}$$

(When q or p appears without a superscript, it will denote the set of all coordinates $\{q^i\}$ or momenta $\{p^i\}$. Also, for convenience, we now set $\hbar = 1$.) To write this amplitude as a functional integral, we first break the time interval into N short slices of duration ϵ. Thus we can write

$$e^{-iHT} = e^{-iH\epsilon} \, e^{-iH\epsilon} \, e^{-iH\epsilon} \cdots e^{-iH\epsilon} \qquad (N \text{ factors}).$$

The trick is to insert a complete set of intermediate states between each of these factors, in the form

$$\mathbf{1} = \Big(\prod_i \int dq_k^i \Big) \, |q_k\rangle \, \langle q_k| \, .$$

Inserting such factors for $k = 1 \ldots (N-1)$, we are left with a product of factors of the form

$$\langle q_{k+1} | \, e^{-iH\epsilon} \, | q_k \rangle \underset{\epsilon \to 0}{\longrightarrow} \langle q_{k+1} | \, \big(1 - iH\epsilon + \cdots \big) \, | q_k \rangle \, . \tag{9.9}$$

To express the first and last factors in this form, we define $q_0 = q_a$ and $q_N = q_b$.

Now we must look inside H and consider what kinds of terms it might contain. The simplest kind of term to evaluate would be a function only of the coordinates, not of the momenta. The matrix element of such a term would be

$$\langle q_{k+1} | \, f(q) \, | q_k \rangle = f(q_k) \prod_i \delta(q_k^i - q_{k+1}^i).$$

It will be convenient to rewrite this as

$$\langle q_{k+1} | \, f(q) \, | q_k \rangle = f\Big(\frac{q_{k+1} + q_k}{2}\Big) \Big(\prod_i \int \frac{dp_k^i}{2\pi} \Big) \exp\Big[i \sum_i p_k^i (q_{k+1}^i - q_k^i) \Big],$$

for reasons that will soon be apparent.

Next consider a term in the Hamiltonian that is purely a function of the momenta. We introduce a complete set of momentum eigenstates to obtain

$$\langle q_{k+1} | \, f(p) \, | q_k \rangle = \Big(\prod_i \int \frac{dp_k^i}{2\pi} \Big) f(p_k) \exp\Big[i \sum_i p_k^i (q_{k+1}^i - q_k^i) \Big].$$

Thus if H contains only terms of the form $f(q)$ and $f(p)$, its matrix element can be written

$$\langle q_{k+1} | \, H(q,p) \, | q_k \rangle = \Big(\prod_i \int \frac{dp_k^i}{2\pi} \Big) H\Big(\frac{q_{k+1} + q_k}{2}, p_k \Big) \exp\Big[i \sum_i p_k^i (q_{k+1}^i - q_k^i) \Big]. \tag{9.10}$$

It would be nice if Eq. (9.10) were true even when H contains products of p's and q's. In general this formula must be false, since the order of a product pq matters on the left-hand side (where H is an operator) but not on the right-hand side (where H is just a function of the numbers p_k and q_k). But for one specific ordering, we can preserve (9.10). For example, the combination

$$\langle q_{k+1}|\, \tfrac{1}{4}\big(q^2p^2 + 2qp^2q + p^2q^2\big)\, |q_k\rangle = \left(\frac{q_{k+1}+q_k}{2}\right)^2 \langle q_{k+1}|\, p^2\, |q_k\rangle$$

works out as desired, since the q's appear symmetrically on the left and right in just the right way. When this happens, the Hamiltonian is said to be *Weyl ordered.* Any Hamiltonian can be put into Weyl order by commuting p's and q's; in general this procedure will introduce some extra terms, and those extra terms must appear on the right-hand side of (9.10).

Assuming from now on that H is Weyl ordered, our typical matrix element from (9.9) can be expressed as

$$\langle q_{k+1}|\, e^{-i\epsilon H}\, |q_k\rangle = \left(\prod_i \int \frac{dp_k^i}{2\pi}\right) \exp\left[-i\epsilon H\left(\frac{q_{k+1}+q_k}{2}, p_k\right)\right] \times \exp\left[i\sum_i p_k^i (q_{k+1}^i - q_k^i)\right].$$

(We have again used the fact that ϵ is small, writing $1 - i\epsilon H$ as $e^{-i\epsilon H}$.) To obtain $U(q_a, q_b; T)$, we multiply N such factors, one for each k, and integrate over the intermediate coordinates q_k:

$$U(q_0, q_N; T) = \left(\prod_{i,k} \int dq_k^i \int \frac{dp_k^i}{2\pi}\right) \times \exp\left[i\sum_k \left(\sum_i p_k^i (q_{k+1}^i - q_k^i) - \epsilon H\left(\frac{q_{k+1}+q_k}{2}, p_k\right)\right)\right]. \tag{9.11}$$

There is one momentum integral for each k from 0 to $N-1$, and one coordinate integral for each k from 1 to $N-1$. This expression is therefore the discretized form of

$$U(q_a, q_b; T) = \left(\prod_i \int \mathcal{D}q(t)\, \mathcal{D}p(t)\right) \exp\left[i\int_0^T dt\left(\sum_i p^i \dot{q}^i - H(q,p)\right)\right], \tag{9.12}$$

where the functions $q(t)$ are constrained at the endpoints, but the functions $p(t)$ are not. Note that the integration measure $\mathcal{D}q$ contains no peculiar constants, as it did in (9.4). The functional measure in (9.12) is just the product of the standard integral over phase space

$$\prod_i \int \frac{dq^i\, dp^i}{2\pi\hbar}$$

at each point in time. Equation (9.12) is the most general formula for computing transition amplitudes via functional integrals.

For a nonrelativistic particle, the Hamiltonian is simply $H = p^2/2m + V(q)$. In this case we can evaluate the p-integrals by completing the square in the exponent:

$$\int \frac{dp_k}{2\pi} \exp\Big[i\big(p_k(q_{k+1}-q_k) - \epsilon p_k^2/2m\big)\Big] = \frac{1}{C(\epsilon)} \exp\Big[\frac{im}{2\epsilon}(q_{k+1} - q_k)^2\Big],$$

where $C(\epsilon)$ is just the factor (9.6). Notice that we have one such factor for each time slice. Thus we recover expression (9.3), in discretized form, including the proper factors of C:

$$U(q_a, q_b; T) = \left(\frac{1}{C(\epsilon)} \prod_k \int \frac{dq_k}{C(\epsilon)}\right) \exp\left[i\sum_k \left(\frac{m}{2}\frac{(q_{k+1}-q_k)^2}{\epsilon} - \epsilon V\left(\frac{q_{k+1}+q_k}{2}\right)\right)\right]. \tag{9.13}$$

9.2 Functional Quantization of Scalar Fields

In this section we will apply the functional integral formalism to the quantum theory of a real scalar field $\phi(x)$. Our goal is to derive the Feynman rules for such a theory directly from functional integral expressions.

The general functional integral formula (9.12) derived in the last section holds for any quantum system, so it should hold for a quantum field theory. In the case of a real scalar field, the coordinates q^i are the field amplitudes $\phi(\mathbf{x})$, and the Hamiltonian is

$$H = \int d^3x \left[\tfrac{1}{2}\pi^2 + \tfrac{1}{2}(\nabla\phi)^2 + V(\phi)\right].$$

Thus our formula becomes

$$\langle\phi_b(\mathbf{x})|\, e^{-iHT}\, |\phi_a(\mathbf{x})\rangle = \int \mathcal{D}\phi\,\mathcal{D}\pi \exp\left[i\int_0^T d^4x\left(\pi\dot{\phi} - \tfrac{1}{2}\pi^2 - \tfrac{1}{2}(\nabla\phi)^2 - V(\phi)\right)\right],$$

where the functions $\phi(x)$ over which we integrate are constrained to the specific configurations $\phi_a(\mathbf{x})$ at $x^0 = 0$ and $\phi_b(\mathbf{x})$ at $x^0 = T$. Since the exponent is quadratic in π, we can complete the square and evaluate the $\mathcal{D}\pi$ integral to obtain

$$\langle\phi_b(\mathbf{x})|\, e^{-iHT}\, |\phi_a(\mathbf{x})\rangle = \int \mathcal{D}\phi \exp\left[i\int_0^T d^4x\,\mathcal{L}\right], \tag{9.14}$$

where

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu\phi)^2 - V(\phi)$$

is the Lagrangian density. The integration measure $\mathcal{D}\phi$ in (9.14) again involves an awkward constant, which we will not write explicitly.

The time integral in the exponent of (9.14) goes from 0 to T, as determined by our choice of what transition function to compute; in all other respects this formula is manifestly Lorentz invariant. Any other symmetries that the Lagrangian may have are also explicitly preserved by the functional integral. As we proceed in our study of quantum field theory, symmetries and their associated conservation laws will play an increasingly central role. We therefore propose to take a rash step: Abandon the Hamiltonian formalism, and take Eq. (9.14) to *define* the Hamiltonian dynamics. Any such formula corresponds to some Hamiltonian; to find it, one can always differentiate with respect to T and derive the Schrödinger equation as in the previous section. We thus consider the Lagrangian $\mathcal{L}$ to be the most fundamental specification of a quantum field theory. We will see next that one can use the functional integral to compute from $\mathcal{L}$ directly, without invoking the Hamiltonian at all.

Correlation Functions

To make direct use of the functional integral, we need a functional formula for computing correlation functions. To find such an expression, consider the object

$$\int \mathcal{D}\phi(x)\, \phi(x_1)\phi(x_2)\, \exp\left[i \int_{-T}^{T} d^4x\, \mathcal{L}(\phi)\right], \tag{9.15}$$

where the boundary conditions on the path integral are $\phi(-T, \mathbf{x}) = \phi_a(\mathbf{x})$ and $\phi(T, \mathbf{x}) = \phi_b(\mathbf{x})$ for some ϕ_a, ϕ_b. We would like to relate this quantity to the two-point correlation function, $\langle\Omega| T\phi_H(x_1)\phi_H(x_2) |\Omega\rangle$. (To distinguish operators from ordinary numbers, we write the Heisenberg picture operator with an explicit subscript: $\phi_H(x)$. Similarly, we will write $\phi_S(\mathbf{x})$ for the Schrödinger picture operator.)

First we break up the functional integral in (9.15) as follows:

$$\int \mathcal{D}\phi(x) = \int \mathcal{D}\phi_1(\mathbf{x}) \int \mathcal{D}\phi_2(\mathbf{x}) \int\limits_{\substack{\phi(x_1^0,\mathbf{x})=\phi_1(\mathbf{x}) \\ \phi(x_2^0,\mathbf{x})=\phi_2(\mathbf{x})}} \mathcal{D}\phi(x). \tag{9.16}$$

The main functional integral $\int \mathcal{D}\phi(x)$ is now constrained at times x_1^0 and x_2^0 (in addition to the endpoints $-T$ and T), but we must integrate separately over the intermediate configurations $\phi_1(\mathbf{x})$ and $\phi_2(\mathbf{x})$. After this decomposition, the extra factors $\phi(x_1)$ and $\phi(x_2)$ in (9.15) become $\phi_1(\mathbf{x}_1)$ and $\phi_2(\mathbf{x}_2)$, and can be taken outside the main integral. The main integral then factors into three pieces, each being a simple transition amplitude according to (9.14). The times x_1^0 and x_2^0 automatically fall in order; for example, if $x_1^0 < x_2^0$, then (9.15) becomes

$$\begin{aligned}\int \mathcal{D}\phi_1(\mathbf{x}) \int \mathcal{D}\phi_2(\mathbf{x})\; \phi_1(\mathbf{x}_1)\phi_2(\mathbf{x}_2)\; &\langle\phi_b| e^{-iH(T-x_2^0)} |\phi_2\rangle \\ &\times \langle\phi_2| e^{-iH(x_2^0-x_1^0)} |\phi_1\rangle\, \langle\phi_1| e^{-iH(x_1^0+T)} |\phi_a\rangle\,.\end{aligned}$$

We can turn the field $\phi_1(\mathbf{x}_1)$ into a Schrödinger operator using $\phi_S(\mathbf{x}_1)\,|\phi_1\rangle = \phi_1(\mathbf{x}_1)\,|\phi_1\rangle$. The completeness relation $\int \mathcal{D}\phi_1\,|\phi_1\rangle\,\langle\phi_1| = \mathbf{1}$ then allows us to eliminate the intermediate state $|\phi_1\rangle$. Similar manipulations work for ϕ_2, yielding the expression

$$\langle\phi_b|\,e^{-iH(T-x_2^0)}\,\phi_S(\mathbf{x}_2)\,e^{-iH(x_2^0-x_1^0)}\,\phi_S(\mathbf{x}_1)\,e^{-iH(x_1^0+T)}\,|\phi_a\rangle\,.$$

Most of the exponential factors combine with the Schrödinger operators to make Heisenberg operators. In the case $x_1^0 > x_2^0$, the order of x_1 and x_2 would simply be interchanged. Thus expression (9.15) is equal to

$$\langle\phi_b|\,e^{-iHT}\,T\{\phi_H(x_1)\phi_H(x_2)\}\,e^{-iHT}\,|\phi_a\rangle\,. \tag{9.17}$$

This expression is almost equal to the two-point correlation function. To make it more nearly equal, we take the limit $T \to \infty(1-i\epsilon)$. Just as in Section 4.2, this trick projects out the vacuum state $|\Omega\rangle$ from $|\phi_a\rangle$ and $|\phi_b\rangle$ (provided that these states have some overlap with $|\Omega\rangle$, which we assume). For example, decomposing $|\phi_a\rangle$ into eigenstates $|n\rangle$ of H, we have

$$e^{-iHT}\,|\phi_a\rangle = \sum_n e^{-iE_nT}\,|n\rangle\,\langle n|\phi_a\rangle \underset{T\to\infty(1-i\epsilon)}{\longrightarrow} \langle\Omega|\phi_a\rangle\,e^{-iE_0\cdot\infty(1-i\epsilon)}\,|\Omega\rangle\,.$$

As in Section 4.2, we obtain some awkward phase and overlap factors. But these factors cancel if we divide by the same quantity as (9.15) but without the two extra fields $\phi(x_1)$ and $\phi(x_2)$. Thus we obtain the simple formula

$$\langle\Omega|\,T\phi_H(x_1)\phi_H(x_2)\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\,\phi(x_1)\phi(x_2)\exp\left[i\int_{-T}^{T} d^4x\,\mathcal{L}\right]}{\int \mathcal{D}\phi\,\exp\left[i\int_{-T}^{T} d^4x\,\mathcal{L}\right]}\,. \tag{9.18}$$

This is our desired formula for the two-point correlation function in terms of functional integrals. For higher correlation functions, just insert additional factors of ϕ on both sides.

Feynman Rules

Our next task is to compute various correlation functions directly from the right-hand side of formula (9.18). In other words, we will now use (9.18) to derive the Feynman rules for a scalar field theory. We will begin by computing the two-point function in the free Klein-Gordon theory, then generalize to higher correlation functions in the free theory. Finally, we will consider ϕ^4 theory, in which we can perform a perturbation expansion to obtain the same Feynman rules as in Section 4.4.

Consider first a noninteracting real-valued scalar field:

$$S_0 = \int d^4x\,\mathcal{L}_0 = \int d^4x\left[\tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2\right]. \tag{9.19}$$

Since $\mathcal{L}_0$ is quadratic in ϕ, the functional integrals in (9.18) take the form of generalized, infinite-dimensional Gaussian integrals. We will therefore be able to evaluate the functional integrals exactly.

Since this is our first functional integral computation, we will do it in a very explicit, but ugly, way. We must first define the integral $\mathcal{D}\phi$ over field configurations. To do this, we use the method of Eq. (9.4) in considering the continuous integral as a limit of a large but finite number of integrals. We thus replace the variables $\phi(x)$ defined on a continuum of points by variables $\phi(x_i)$ defined at the points x_i of a square lattice. Let the lattice spacing be ϵ, let the four-dimensional spacetime volume be L^4, and define

$$\mathcal{D}\phi = \prod_i d\phi(x_i), \tag{9.20}$$

up to an irrelevant overall constant.

The field values $\phi(x_i)$ can be represented by a discrete Fourier series:

$$\phi(x_i) = \frac{1}{V}\sum_n e^{-ik_n\cdot x_i}\phi(k_n), \tag{9.21}$$

where $k_n^\mu = 2\pi n^\mu/L$, with n^μ an integer, $|k^\mu| < \pi/\epsilon$, and $V = L^4$. The Fourier coefficients $\phi(k)$ are complex. However, $\phi(x)$ is real, and so these coefficients must obey the constraint $\phi^*(k) = \phi(-k)$. We will consider the real and imaginary parts of the $\phi(k_n)$ with $k_n^0 > 0$ as independent variables. The change of variables from the $\phi(x_i)$ to these new variables $\phi(k_n)$ is a unitary transformation, so we can rewrite the integral as

$$\mathcal{D}\phi(x) = \prod_{k_n^0>0} d\,\mathrm{Re}\,\phi(k_n)\, d\,\mathrm{Im}\,\phi(k_n).$$

Later, we will take the limit $L \to \infty$, $\epsilon \to 0$. The effect of this limit is to convert discrete, finite sums over k_n to continuous integrals over k:

$$\frac{1}{V}\sum_n \to \int \frac{d^4k}{(2\pi)^4}. \tag{9.22}$$

In the following discussion, this limit will produce Feynman perturbation theory in the form derived in Part I. We will not eliminate the infrared and ultraviolet divergences of Feynman diagrams that we encountered in Chapter 6, but at least the functional integral introduces no new types of singular behavior.

Having defined the measure of integration, we now compute the functional integral over ϕ. The action (9.19) can be rewritten in terms of the Fourier coefficients as

$$\begin{aligned}\int d^4x\left[\tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2\right] &= -\frac{1}{V}\sum_n \tfrac{1}{2}(m^2 - k_n^2)\,|\phi(k_n)|^2 \\ &= -\frac{1}{V}\sum_{k_n^0>0}(m^2-k_n^2)\left[(\mathrm{Re}\,\phi_n)^2 + (\mathrm{Im}\,\phi_n)^2\right],\end{aligned}$$

where we have abbreviated $\phi(k_n)$ as ϕ_n in the second line. The quantity $(m^2 - k_n^2) = (m^2 + |\mathbf{k}_n|^2 - k_n^{02})$ is positive as long as k_n^0 is not too large. In the following discussion, we will treat this quantity as if it were positive. More precisely, we evaluate it by analytic continuation from the region where $|\mathbf{k}_n| > k_n^0$.

The denominator of formula (9.18) now takes the form of a product of Gaussian integrals:

$$\begin{aligned}
\int \mathcal{D}\phi\, e^{iS_0} &= \Big(\prod_{k_n^0>0} \int d\,\mathrm{Re}\,\phi_n\, d\,\mathrm{Im}\,\phi_n\Big) \exp\Big[-\frac{i}{V}\sum_{k_n^0>0}(m^2-k_n^2)|\phi_n|^2\Big] \\
&= \prod_{k_n^0>0}\Big(\int d\,\mathrm{Re}\,\phi_n \exp\Big[-\frac{i}{V}(m^2-k_n^2)(\mathrm{Re}\,\phi_n)^2\Big]\Big) \\
&\qquad\qquad \times \Big(\int d\,\mathrm{Im}\,\phi_n \exp\Big[-\frac{i}{V}(m^2-k_n^2)(\mathrm{Im}\,\phi_n)^2\Big]\Big) \\
&= \prod_{k_n^0>0}\sqrt{\frac{-i\pi V}{m^2-k_n^2}}\sqrt{\frac{-i\pi V}{m^2-k_n^2}} \\
&= \prod_{\text{all } k_n}\sqrt{\frac{-i\pi V}{m^2-k_n^2}}.
\end{aligned} \tag{9.23}$$

To justify using Gaussian integration formulae when the exponent appears to be purely imaginary, recall that the time integral in (9.18) is along a contour that is rotated clockwise in the complex plane: $t \to t(1-i\epsilon)$. This means that we should change $k^0 \to k^0(1+i\epsilon)$ in (9.21) and all subsequent equations; in particular, we should replace $(k^2 - m^2) \to (k^2 - m^2 + i\epsilon)$. The $i\epsilon$ term gives the necessary convergence factor for the Gaussian integrals. It also defines the direction of the analytic continuation that might be needed to define the square roots in (9.23).

To understand the result of (9.23), consider as an analogy the general Gaussian integral

$$\Big(\prod_k \int d\xi_k\Big) \exp\big[-\xi_i B_{ij}\xi_j\big],$$

where B is a symmetric matrix with eigenvalues b_i. To evaluate this integral we write $\xi_i = O_{ij}x_j$, where O is the orthogonal matrix of eigenvectors that diagonalizes B. Changing variables from ξ_i to the coefficients x_i, we have

$$\begin{aligned}
\Big(\prod_k \int d\xi_k\Big) \exp\big[-\xi_i B_{ij}\xi_j\big] &= \Big(\prod_k \int dx_k\Big) \exp\Big[-\sum_i b_i x_i^2\Big] \\
&= \prod_i\Big(\int dx_i \exp\big[-b_i x_i^2\big]\Big)
\end{aligned}$$

$$= \prod_i \sqrt{\frac{\pi}{b_i}}$$

$$= \text{const} \times \left[\det B\right]^{-1/2}. \tag{9.24}$$

The analogy is clearer if we perform an integration by parts to write the Klein-Gordon action as

$$S_0 = \tfrac{1}{2}\int d^4x\, \phi\,(-\partial^2 - m^2)\,\phi \;+\; (\text{surface term}).$$

Thus the matrix B corresponds to the operator $(m^2+\partial^2)$, and we can formally write our result as

$$\int \mathcal{D}\phi\, e^{iS_0} = \text{const} \times \left[\det(m^2+\partial^2)\right]^{-1/2}. \tag{9.25}$$

This object is called a *functional determinant.* The actual result (9.23) looks quite ill-defined, and in fact all of these factors will cancel in Eq. (9.18). However, in many circumstances, the functional determinant itself has physical meaning. We will see examples of this in Sections 9.5 and 11.4.

Now consider the numerator of formula (9.18). We need to Fourier-expand the two extra factors of ϕ:

$$\phi(x_1)\phi(x_2) = \frac{1}{V}\sum_m e^{-ik_m\cdot x_1}\phi_m\, \frac{1}{V}\sum_l e^{-ik_l\cdot x_2}\phi_l.$$

Thus the numerator is

$$\begin{aligned} \frac{1}{V^2}\sum_{m,l} e^{-i(k_m\cdot x_1 + k_l\cdot x_2)} &\left(\prod_{k_n^0>0}\int d\operatorname{Re}\phi_n\, d\operatorname{Im}\phi_n\right) \\ &\times (\operatorname{Re}\phi_m + i\operatorname{Im}\phi_m)(\operatorname{Re}\phi_l + i\operatorname{Im}\phi_l) \\ &\times \exp\Bigl[-\frac{i}{V}\sum_{k_n^0>0}(m^2 - k_n^2)\bigl[(\operatorname{Re}\phi_n)^2 + (\operatorname{Im}\phi_n)^2\bigr]\Bigr]. \end{aligned} \tag{9.26}$$

For most values of k_m and k_l this expression is zero, since the extra factors of ϕ make the integrand odd. The situation is more complicated when $k_m = \pm k_l$. Suppose, for example, that $k_m^0 > 0$. Then if $k_l = +k_m$, the term involving $(\operatorname{Re}\phi_m)^2$ is nonzero, but is exactly canceled by the term involving $(\operatorname{Im}\phi_m)^2$. If $k_l = -k_m$, however, the relation $\phi(-k) = \phi^*(k)$ gives an extra minus sign on the $(\operatorname{Im}\phi_m)^2$ term, so the two terms add. When $k_m^0 < 0$ we obtain the same expression, so the numerator is

$$\text{Numerator} = \frac{1}{V^2}\sum_m e^{-ik_m\cdot(x_1-x_2)}\left(\prod_{k_n^0>0}\frac{-i\pi V}{m^2 - k_n^2}\right)\frac{-iV}{m^2 - k_m^2 - i\epsilon}.$$

The factor in parentheses is identical to the denominator (9.23), while the rest of this expression is the discretized form of the Feynman propagator. Taking the continuum limit (9.22), we find

$$\langle 0|\, T\phi(x_1)\phi(x_2)\, |0\rangle = \int \frac{d^4k}{(2\pi)^4} \frac{i\, e^{-ik\cdot(x_1-x_2)}}{k^2 - m^2 + i\epsilon} = D_F(x_1-x_2). \tag{9.27}$$

This is exactly right, including the $+i\epsilon$.

Next we would like to compute higher correlation functions in the free Klein-Gordon theory.

Inserting an extra factor of ϕ in (9.18), we see that the three-point function vanishes, since the integrand of the numerator is odd. All other odd correlation functions vanish for the same reason.

The four-point function has four factors of ϕ in the numerator. Fourier-expanding the fields, we obtain an expression similar to Eq. (9.26), but with a quadruple sum over indices that we will call m, l, p, and q. The integrand contains the product

$$(\mathrm{Re}\,\phi_m + i\,\mathrm{Im}\,\phi_m)(\mathrm{Re}\,\phi_l + i\,\mathrm{Im}\,\phi_l)(\mathrm{Re}\,\phi_p + i\,\mathrm{Im}\,\phi_p)(\mathrm{Re}\,\phi_q + i\,\mathrm{Im}\,\phi_q).$$

Again, most of the terms vanish because the integrand is odd. One of the nonvanishing terms occurs when $k_l = -k_m$ and $k_q = -k_p$. After the Gaussian integrations, this term of the numerator is

$$\frac{1}{V^4}\sum_{m,p} e^{-ik_m\cdot(x_1-x_2)}e^{-ik_p\cdot(x_3-x_4)}\left(\prod_{k_n^0>0}\frac{-i\pi V}{m^2-k_n^2}\right)\frac{-iV}{m^2-k_m^2-i\epsilon}\,\frac{-iV}{m^2-k_p^2-i\epsilon}$$

$$\underset{V\to\infty}{\longrightarrow}\left(\prod_{k_n^0>0}\frac{-i\pi V}{m^2-k_n^2}\right)D_F(x_1-x_2)D_F(x_3-x_4).$$

The factor in parentheses is again canceled by the denominator. We obtain similar terms for each of the other two ways of grouping the four momenta in pairs. To keep track of the groupings, let us define the *contraction* of two fields as

$$\overline{\phi(x_1)\phi(x_2)} = \frac{\int \mathcal{D}\phi\, e^{iS_0}\phi(x_1)\phi(x_2)}{\int \mathcal{D}\phi\, e^{iS_0}} = D_F(x_1 - x_2). \tag{9.28}$$

Then the four-point function is simply

$$\begin{aligned}\langle 0|\, T\phi_1\phi_2\phi_3\phi_4\, |0\rangle &= \text{sum of all full contractions} \\ &= D_F(x_1 - x_2)D_F(x_3 - x_4) \\ &\quad + D_F(x_1 - x_3)D_F(x_2 - x_4) \\ &\quad\quad + D_F(x_1 - x_4)D_F(x_2 - x_3),\end{aligned} \tag{9.29}$$

the same expression that we obtained using Wick's theorem in Eq. (4.40).

The same method allows us to compute still higher correlation functions. In each case the answer is just the sum of all possible full contractions of the fields. This result, identical to that obtained from Wick's theorem in Section 4.3, arises here from the simple rules of Gaussian integration.

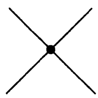
We are now ready to move from the free Klein-Gordon theory to ϕ^4 theory. Add to $\mathcal{L}_0$ a ϕ^4 interaction:

$$\mathcal{L} = \mathcal{L}_0 - \frac{\lambda}{4!}\phi^4.$$

Assuming that λ is small, we can expand

$$\exp\left[i\int d^4x\,\mathcal{L}\right] = \exp\left[i\int d^4x\,\mathcal{L}_0\right]\left(1 - i\int d^4x\,\frac{\lambda}{4!}\phi^4 + \cdots\right).$$

Making this expansion in both the numerator and the denominator of (9.18), we see that each is (aside from the constant factor (9.23), which again cancels) expressed entirely in terms of free-field correlation functions. Moreover, since $i\int d^3x\,\mathcal{L}_{\rm int} = -iH_{\rm int}$, we obtain exactly the same expansion as in Eq. (4.31). We can express both the numerator and the denominator in terms of Feynman diagrams, with the fundamental interaction again given by the vertex



$$= -i\lambda\,(2\pi)^4\delta^{(4)}(\textstyle\sum p). \qquad (9.30)$$

All of the combinatorics work the same as in Section 4.4. In particular, the disconnected vacuum bubble diagrams exponentiate and factor from the numerator of (9.18), and are canceled by the denominator, just as in Eq. (4.31).

The vertex rule for ϕ^4 theory follows from the Lagrangian in an exceedingly simple way, and this simple procedure will turn out to be valid for other quantum field theories as well. Once the quadratic terms in the Lagrangian are properly understood and the propagators of the theory are computed, the vertices can be read directly from the Lagrangian as the coefficients of the cubic and higher-order terms.

Functional Derivatives and the Generating Functional

To conclude this section, we will now introduce a slicker, more formal, method for computing correlation functions. This method, based on an object called the *generating functional*, avoids the awkward Fourier expansions of the preceding derivation.

First we define the *functional derivative*, $\delta/\delta J(x)$, as follows. The functional derivative obeys the basic axiom (in four dimensions)

$$\frac{\delta}{\delta J(x)}J(y) = \delta^{(4)}(x-y) \qquad \text{or} \qquad \frac{\delta}{\delta J(x)}\int d^4y\,J(y)\phi(y) = \phi(x). \qquad (9.31)$$

This definition is the natural generalization, to continuous functions, of the rule for discrete vectors,

$$\frac{\partial}{\partial x_i}x_j = \delta_{ij} \qquad \text{or} \qquad \frac{\partial}{\partial x_i}\sum_j x_j k_j = k_i.$$

To take functional derivatives of more complicated functionals we simply use the ordinary rules for derivatives of composite functions. For example,

$$\frac{\delta}{\delta J(x)} \exp\Big[i\int d^4y\, J(y)\phi(y)\Big] = i\phi(x)\exp\Big[i\int d^4y\, J(y)\phi(y)\Big]. \tag{9.32}$$

When the functional depends on the derivative of J, we integrate by parts before applying the functional derivative:

$$\frac{\delta}{\delta J(x)} \int d^4y\, \partial_\mu J(y) V^\mu(y) = -\partial_\mu V^\mu(x). \tag{9.33}$$

The basic object of this formalism is the *generating functional* of correlation functions, $Z[J]$. (Some authors call it $W[J]$.) In a scalar field theory, $Z[J]$ is defined as

$$Z[J] \equiv \int \mathcal{D}\phi\, \exp\Big[i\int d^4x\big[\mathcal{L} + J(x)\phi(x)\big]\Big]. \tag{9.34}$$

This is a functional integral over ϕ in which we have added to $\mathcal{L}$ in the exponent a *source term*, $J(x)\phi(x)$.

Correlation functions of the Klein-Gordon field theory can be simply computed by taking functional derivatives of the generating functional. For example, the two-point function is

$$\langle 0|\, T\phi(x_1)\phi(x_2)\,|0\rangle = \frac{1}{Z_0}\left(-i\frac{\delta}{\delta J(x_1)}\right)\left(-i\frac{\delta}{\delta J(x_2)}\right) Z[J]\Big|_{J=0}, \tag{9.35}$$

where $Z_0 = Z[J=0]$. Each functional derivative brings down a factor of ϕ in the numerator of $Z[J]$; setting $J=0$, we recover expression (9.18). To compute higher correlation functions we simply take more functional derivatives.

Formula (9.35) is useful because, in a free field theory, $Z[J]$ can be rewritten in a very explicit form. Consider the exponent of (9.34) in the free Klein-Gordon theory. Integrating by parts, we obtain

$$\int d^4x\big[\mathcal{L}_0(\phi) + J\phi\big] = \int d^4x\big[\tfrac{1}{2}\phi(-\partial^2 - m^2 + i\epsilon)\phi + J\phi\big]. \tag{9.36}$$

(The $i\epsilon$ is a convergence factor for the functional integral, as we discussed below Eq. (9.23).) We can complete the square by introducing a shifted field,

$$\phi'(x) \equiv \phi(x) - i\int d^4y\, D_F(x-y)J(y).$$

Making this substitution and using the fact that D_F is a Green's function of the Klein-Gordon operator, we find that (9.36) becomes

$$\begin{aligned}\int d^4x\big[\mathcal{L}_0(\phi) + J\phi\big] &= \int d^4x\big[\tfrac{1}{2}\phi'(-\partial^2 - m^2 + i\epsilon)\phi'\big] \\ &\quad - \int d^4x\, d^4y\, \tfrac{1}{2}J(x)[-iD_F(x-y)]J(y).\end{aligned}$$

More symbolically, we could write the change of variables as

$$\phi' \equiv \phi + (-\partial^2 - m^2 + i\epsilon)^{-1} J, \tag{9.37}$$

and the result

$$\int d^4x \big[\mathcal{L}_0(\phi) + J\phi\big] = \int d^4x \big[\tfrac{1}{2}\phi'(-\partial^2 - m^2 + i\epsilon)\phi' - \tfrac{1}{2}J(-\partial^2 - m^2 + i\epsilon)^{-1} J\big]. \tag{9.38}$$

Now change variables from ϕ to ϕ' in the functional integral of (9.34). This is just a shift, and so the Jacobian of the transformation is 1. The result is

$$\int \mathcal{D}\phi' \exp\Big[i\int d^4x\, \mathcal{L}_0(\phi')\Big] \exp\Big[-i\int d^4x\, d^4y\, \tfrac{1}{2}J(x)[-iD_F(x-y)]J(y)\Big].$$

The second exponential factor is independent of ϕ', while the remaining integral over ϕ' is precisely Z_0. Thus the generating functional of the free Klein-Gordon theory is simply

$$Z[J] = Z_0 \exp\Big[-\tfrac{1}{2}\int d^4x\, d^4y\, J(x)D_F(x-y)J(y)\Big]. \tag{9.39}$$

Let us use Eqs. (9.39) and (9.35) to compute some correlation functions. The two-point function is

$$\begin{aligned}
&\langle 0|T\phi(x_1)\phi(x_2)\,|0\rangle \\
&= -\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)} \exp\Big[-\tfrac{1}{2}\int d^4x\, d^4y\, J(x)D_F(x-y)J(y)\Big]\Big|_{J=0} \\
&= -\frac{\delta}{\delta J(x_1)}\Big[-\tfrac{1}{2}\int d^4y\, D_F(x_2-y)J(y) - \tfrac{1}{2}\int d^4x\, J(x)D_F(x-x_2)\Big]\frac{Z[J]}{Z_0}\Big|_{J=0} \\
&= D_F(x_1 - x_2).
\end{aligned} \tag{9.40}$$

Taking one derivative brings down two identical terms; the second derivative gives several terms, but only when it acts on the outside factor do we get a term that survives when we set $J = 0$.

It is instructive to work out the four-point function by this method as well. In order to fit the computation in a reasonable amount of space, let us abbreviate arguments of functions as subscripts: $\phi_1 \equiv \phi(x_1)$, $J_x \equiv J(x)$, $D_{x4} \equiv D_F(x-x_4)$, and so on. Repeated subscripts will be integrated over implicitly. The four-point function is then

$$\begin{aligned}
\langle 0|\,T\phi_1\phi_2\phi_3\phi_4\,|0\rangle &= \frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\frac{\delta}{\delta J_3}\big[-J_xD_{x4}\big]e^{-\frac{1}{2}J_xD_{xy}J_y}\Big|_{J=0} \\
&= \frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\Big[-D_{34} + J_xD_{x4}J_yD_{y3}\Big]e^{-\frac{1}{2}J_xD_{xy}J_y}\Big|_{J=0} \\
&= \frac{\delta}{\delta J_1}\Big[D_{34}J_xD_{x2} + D_{24}J_yD_{y3} + J_xD_{x4}D_{23}\Big]e^{-\frac{1}{2}J_xD_{xy}J_y}\Big|_{J=0} \\
&= D_{34}D_{12} + D_{24}D_{13} + D_{14}D_{23},
\end{aligned} \tag{9.41}$$

in agreement with (9.29). The rules for differentiating the exponential give rise to the same familiar pattern: We get one term for each possible way of contracting the four points in pairs, with a factor of D_F for each contraction.

The generating functional method used just above to construct the correlations of a free field theory can be used as well to represent the correlation functions of an interacting field theory. Formula (9.35) is independent of whether the theory is free or interacting. The factor $Z[J=0]$ is nontrivial in the case of an interacting field theory, but it simply gives the denominator of Eq. (9.18), that is, the sum of vacuum diagrams. Again from this approach, the combinatoric issues in the evaluation of correlation functions are the same as in Section 4.4.

9.3 The Analogy Between Quantum Field Theory and Statistical Mechanics

Let us now pause from the technical aspects of this discussion to consider some implications of the formulae we have derived. To begin, let us summarize the formal conclusions of the previous section in the following way: For a field theory governed by the Lagrangian $\mathcal{L}$, the generating functional of correlation functions is

$$Z[J] = \int \mathcal{D}\phi \, \exp\Big[i\int d^4x \,(\mathcal{L} + J\phi)\Big]. \tag{9.42}$$

The time variable of integration in the exponent runs from $-T$ to T, with $T \to \infty(1-i\epsilon)$. A correlation function such as (9.18) is reproduced by writing

$$\langle 0| T\phi(x_1)\phi(x_2) |0\rangle = Z[J]^{-1}\left(-i\frac{\delta}{\delta J(x_1)}\right)\left(-i\frac{\delta}{\delta J(x_2)}\right) Z[J]\bigg|_{J=0}. \tag{9.43}$$

The generating functional (9.42) is reminiscent of the partition function of statistical mechanics. It has the same general structure of an integral over all possible configurations of an exponential statistical weight. The source $J(x)$ plays the role of an external field. In fact, our method of computing correlation functions by differentiating with respect to $J(x)$ mimics the trick often used in statistical mechanics of computing correlation functions by differentiating with respect to such variables as the pressure or the magnetic field.

This analogy can be made more precise by manipulating the time variable of integration in (9.42). The derivation of the functional integral formula implied that the time integration was slightly tipped into the complex plane, in just the direction to permit the contour to be rotated clockwise onto the imaginary axis. We have already noted (below (9.23)) that the original infinitesimal rotation gives the correct $i\epsilon$ prescription to produce the Feynman propagator. The finite rotation is the analogue in configuration space of the Wick rotation of the time component of momentum illustrated in Fig. 6.1. Like the Wick rotation in a momentum integral, this Wick rotation of the time coordinate $t \to -ix^0$ produces a Euclidean 4-vector product:

$$x^2 = t^2 - |\mathbf{x}|^2 \to -(x^0)^2 - |\mathbf{x}|^2 = -|x_E|^2. \tag{9.44}$$

It is possible to show, by manipulating the expression for each Feynman diagram, that the analytic continuation of the time variables in any Green's function of a quantum field theory produces a correlation function invariant under the rotational symmetry of four-dimensional Euclidean space. This Wick rotation inside the functional integral demonstrates this same conclusion in a more general way.

To understand what we have achieved by this rotation, consider the example of ϕ^4 theory. The action of ϕ^4 theory coupled to sources is

$$\int d^4x\,(\mathcal{L} + J\phi) = \int d^4x\left[\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 + J\phi\right]. \tag{9.45}$$

After the Wick rotation (9.44), this expression takes the form

$$i\int d^4x_E(\mathcal{L}_E - J\phi) = i\int d^4x_E\left[\frac{1}{2}(\partial_{E\mu}\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4 - J\phi\right]. \tag{9.46}$$

This expression is identical in form to the expression (8.8) for the Gibbs free energy of a ferromagnet in the Landau theory. The field $\phi(x_E)$ plays the role of the fluctuating spin field $s(\mathbf{x})$, and the source $J(\mathbf{x})$ plays the role of an external magnetic field. Note that the new ferromagnet lives in four, rather than three, spatial dimensions.

The Wick-rotated generating functional $Z[J]$ becomes

$$Z[J] = \int \mathcal{D}\phi\,\exp\left[-\int d^4x_E\,(\mathcal{L}_E - J\phi)\right]. \tag{9.47}$$

The functional $\mathcal{L}_E[\phi]$ has the form of an energy: It is bounded from below and becomes large when the field ϕ has large amplitude or large gradients. The exponential, then, is a reasonable statistical weight for the fluctuations of ϕ. In this new form, $Z[J]$ is precisely the partition function describing the statistical mechanics of a macroscopic system, described approximately by treating the fluctuating variable as a continuum field.

The Green's functions of $\phi(x_E)$ after Wick rotation can be calculated from the functional integral (9.47) exactly as we computed Minkowski Green's functions in the previous section. For the free theory ($\lambda = 0$), a set of manipulations analogous to those that produced (9.27) or (9.40) gives the correlation function of ϕ as

$$\langle\phi(x_{E1})\phi(x_{E2})\rangle = \int \frac{d^4k_E}{(2\pi)^4}\frac{e^{ik_E\cdot(x_{E1}-x_{E2})}}{k_E^2 + m^2}. \tag{9.48}$$

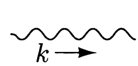
This is just the Feynman propagator evaluated in the spacelike region; according to Eq. (2.52), this function falls off as $\exp(-m|x_{E1} - x_{E2}|)$. That behavior is the four-dimensional analogue of the spin correlation function (8.15). We see that, in the Euclidean continuation of field theory Green's functions, the

Compton wavelength m^{-1} of the quanta becomes the correlation length of statistical fluctuations.

This correspondence between quantum field theory and statistical mechanics will play an important role in the developments of the next few chapters. In essence, it adds to our reserves of knowledge a completely new source of intuition about how field theory expectation values should behave. This intuition will be useful in imagining the general properties of loop diagrams and, as we have already discussed in Chapter 8, it will give important insights that will help us correctly understand the role of ultraviolet divergences in field theory calculations. In Chapter 13, we will see that field theory can also contribute to statistical mechanics by making profound predictions about the behavior of thermal systems from the properties of Feynman diagrams.

9.4 Quantization of the Electromagnetic Field

In Section 4.8 we stated without proof the Feynman rule for the photon propagator,



$$= \frac{-ig_{\mu\nu}}{k^2 + i\epsilon}. \tag{9.49}$$

Now that we have the functional integral quantization method at our command, let us apply it to the derivation of this expression.

Consider the functional integral

$$\int \mathcal{D}A\, e^{iS[A]}, \tag{9.50}$$

where $S[A]$ is the action for the free electromagnetic field. (The functional integral is over each of the four components: $\mathcal{D}A \equiv \mathcal{D}A^0\mathcal{D}A^1\mathcal{D}A^2\mathcal{D}A^3$.) Integrating by parts and expanding the field as a Fourier integral, we can write the action as

$$\begin{aligned} S &= \int d^4x\, \left[-\tfrac{1}{4}(F_{\mu\nu})^2\right] \\ &= \frac{1}{2}\int d^4x\, A_\mu(x)\big(\partial^2 g^{\mu\nu} - \partial^\mu\partial^\nu\big)A_\nu(x) \\ &= \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\, \widetilde{A}_\mu(k)\big(-k^2 g^{\mu\nu} + k^\mu k^\nu\big)\widetilde{A}_\nu(-k). \end{aligned} \tag{9.51}$$

This expression vanishes when $\widetilde{A}_\mu(k) = k_\mu\, \alpha(k)$, for any scalar function $\alpha(k)$. For this large set of field configurations the integrand of (9.50) is 1, and therefore the functional integral is badly divergent (there is no Gaussian damping). Equivalently, the equation

$$\begin{aligned} &\big(\partial^2 g_{\mu\nu} - \partial_\mu\partial_\nu\big)D_F^{\nu\rho}(x-y) = i\delta_\mu{}^\rho\, \delta^{(4)}(x-y) \\ \text{or}\qquad &\big(-k^2 g_{\mu\nu} + k_\mu k_\nu\big)\widetilde{D}_F^{\nu\rho}(k) = i\delta_\mu{}^\rho, \end{aligned} \tag{9.52}$$

which would define the Feynman propagator $D_F^{\nu\rho}$, has no solution, since the 4×4 matrix $(-k^2 g_{\mu\nu} + k_\mu k_\nu)$ is singular.

This difficulty is due to gauge invariance. Recall that $F_{\mu\nu}$, and hence $\mathcal{L}$, is invariant under a general gauge transformation of the form

$$A_\mu(x) \to A_\mu(x) + \frac{1}{e}\partial_\mu \alpha(x).$$

The troublesome modes are those for which $A_\mu(x) = \frac{1}{e}\partial_\mu \alpha(x)$, that is, those that are gauge-equivalent to $A_\mu(x) = 0$. The functional integral is badly defined because we are redundantly integrating over a continuous infinity of physically equivalent field configurations. To fix the problem, we would like to isolate the interesting part of the functional integral, which counts each physical configuration only once.

We can accomplish this by means of a trick, due to Faddeev and Popov.† Let $G(A)$ be some function that we wish to set equal to zero as a gauge-fixing condition; for example, $G(A) = \partial_\mu A^\mu$ corresponds to Lorentz gauge. We could constrain the functional integral to cover only the configurations with $G(A) = 0$ by inserting a functional delta function, $\delta\big(G(A)\big)$. (Think of this object as an infinite product of delta functions, one for each point x.) To do so legally, we insert 1 under the integral of (9.50), in the following form:

$$1 = \int \mathcal{D}\alpha(x)\, \delta\big(G(A^\alpha)\big) \det\Big(\frac{\delta G(A^\alpha)}{\delta \alpha}\Big), \tag{9.53}$$

where A^α denotes the gauge-transformed field,

$$A^\alpha_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu \alpha(x).$$

Equation (9.53) is the continuum generalization of the identity

$$1 = \Big(\prod_i \int da_i\Big)\, \delta^{(n)}(\mathbf{g}(\mathbf{a})) \det\Big(\frac{\partial g_i}{\partial a_j}\Big)$$

for discrete n-dimensional vectors. In Lorentz gauge we have $G(A^\alpha) = \partial^\mu A_\mu + (1/e)\partial^2 \alpha$, so the functional determinant $\det\big(\delta G(A^\alpha)/\delta\alpha\big)$ is equal to $\det(\partial^2/e)$. For the present discussion, the only relevant property of this determinant is that it is independent of A, so we can treat it as a constant in the functional integral.

After inserting (9.53), the functional integral (9.50) becomes

$$\det\Big(\frac{\delta G(A^\alpha)}{\delta\alpha}\Big) \int \mathcal{D}\alpha \int \mathcal{D}A\, e^{iS[A]}\, \delta\big(G(A^\alpha)\big).$$

Now change variables from A to A^α. This is a simple shift, so $\mathcal{D}A = \mathcal{D}A^\alpha$. Also, by gauge invariance, $S[A] = S[A^\alpha]$. Since A^α is now just a dummy

†L. D. Faddeev and V. N. Popov, *Phys. Lett.* **25B**, 29 (1967).

integration variable, we can rename it back to A, obtaining

$$\int \mathcal{D}A\, e^{iS[A]} = \det\left(\frac{\delta G(A^\alpha)}{\delta\alpha}\right) \int \mathcal{D}\alpha \int \mathcal{D}A\, e^{iS[A]}\, \delta\big(G(A)\big). \tag{9.54}$$

The functional integral over A is now restricted by the delta function to physically inequivalent field configurations, as desired. The divergent integral over $\alpha(x)$ simply gives an infinite multiplicative factor.

To go further we must specify a gauge-fixing function $G(A)$. We choose the general class of functions

$$G(A) = \partial^\mu A_\mu(x) - \omega(x), \tag{9.55}$$

where $\omega(x)$ can be any scalar function. Setting this $G(A)$ equal to zero gives a generalization of the Lorentz gauge condition. The functional determinant is the same as in Lorentz gauge, $\det\big(\delta G(A^\alpha)/\delta\alpha\big) = \det(\partial^2/e)$. Thus the functional integral becomes

$$\int \mathcal{D}A\, e^{iS[A]} = \det\left(\frac{1}{e}\partial^2\right)\left(\int \mathcal{D}\alpha\right) \int \mathcal{D}A\, e^{iS[A]}\, \delta\big(\partial^\mu A_\mu - \omega(x)\big).$$

This equality holds for any $\omega(x)$, so it will also hold if we replace the right-hand side with any properly normalized linear combination involving different functions $\omega(x)$. For our final trick, we will integrate over all $\omega(x)$, with a Gaussian weighting function centered on $\omega = 0$. The above expression is thus equal to

$$\begin{aligned} N(\xi)\int \mathcal{D}\omega \exp\left[-i\int d^4x\, \frac{\omega^2}{2\xi}\right] \det\left(\frac{1}{e}\partial^2\right)\left(\int \mathcal{D}\alpha\right)\int \mathcal{D}A\, e^{iS[A]}\, \delta\big(\partial^\mu A_\mu - \omega(x)\big) \\ = N(\xi)\det\left(\frac{1}{e}\partial^2\right)\left(\int \mathcal{D}\alpha\right)\int \mathcal{D}A\, e^{iS[A]} \exp\left[-i\int d^4x\, \frac{1}{2\xi}(\partial^\mu A_\mu)^2\right], \end{aligned} \tag{9.56}$$

where $N(\xi)$ is an unimportant normalization constant and we have used the delta function to perform the integral over ω. We can choose ξ to be any finite constant. Effectively, we have added a new term $-(\partial^\mu A_\mu)^2/2\xi$ to the Lagrangian.

So far we have worked only with the denominator of our formula for correlation functions,

$$\langle\Omega|\, T\,\mathcal{O}(A)\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}A\, \mathcal{O}(A) \exp\left[i\int_{-T}^{T} d^4x\, \mathcal{L}\right]}{\int \mathcal{D}A\, \exp\left[i\int_{-T}^{T} d^4x\, \mathcal{L}\right]}.$$

The same manipulations can also be performed on the numerator, provided that the operator $\mathcal{O}(A)$ is gauge invariant. (If it is not, the variable change from A to A^α preceding Eq. (9.54) does not work). Assuming that $\mathcal{O}(A)$ is gauge invariant, we find for its correlation function

$$\langle\Omega|\, T\,\mathcal{O}(A)\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}A\,\mathcal{O}(A)\exp\Big[i\int_{-T}^{T} d^4x \left[\mathcal{L} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2\right]\Big]}{\int \mathcal{D}A\, \exp\Big[i\int_{-T}^{T} d^4x \left[\mathcal{L} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2\right]\Big]}. \tag{9.57}$$

The awkward constant factors in (9.56) have canceled; the only trace left by this whole process is the extra ξ-term that is added to the action.

At the beginning of this section, in Eq. (9.52), we saw that we could not obtain a sensible photon propagator from the action $S[A]$. With the new ξ-term, however, that equation becomes

$$\big(-k^2 g_{\mu\nu} + (1-\tfrac{1}{\xi})k_\mu k_\nu\big)\widetilde{D}_F^{\nu\rho}(k) = i\delta_\mu{}^\rho,$$

which has the solution

$$\widetilde{D}_F^{\mu\nu}(k) = \frac{-i}{k^2+i\epsilon}\left(g^{\mu\nu} - (1-\xi)\frac{k^\mu k^\nu}{k^2}\right). \tag{9.58}$$

This is our desired expression for the photon propagator. The $i\epsilon$ term in the denominator arises exactly as in the Klein-Gordon case. Note the overall minus sign relative to the Klein-Gordon propagator, which was already evident in Eq. (9.52).

In practice one usually chooses a specific value of ξ when making computations. Two choices that are often convenient are

$$\begin{aligned} \xi &= 0 \qquad \text{Landau gauge;} \\ \xi &= 1 \qquad \text{Feynman gauge.} \end{aligned}$$

So far in this book we have always used Feynman gauge.‡

The Faddeev-Popov procedure guarantees that the value of any correlation function of gauge-invariant operators computed from Feynman diagrams will be independent of the value of ξ used in the calculation (as long as the same value of ξ is used consistently). In the case of QED, it is not difficult to prove this ξ-independence directly. Notice in Eq. (9.58) that ξ multiplies a term in the photon propagator proportional to $k^\mu k^\nu$. According to the Ward-Takahashi identity (7.68), the replacement in a Green's function of any photon propagator by $k^\mu k^\nu$ yields zero, except for terms involving external off-shell fermions. These terms are equal and opposite for particle and antiparticle and vanish when the fermions are grouped into gauge-invariant combinations.

To complete our treatment of the quantization of the electromagnetic field, we need one additional ingredient. In Chapters 5 and 6, we computed S-matrix

‡Other choices of ξ may be useful in specific applications; for example, in certain problems of bound states in QED, the *Yennie gauge*, $\xi = 3$, produces a cancellation that is otherwise difficult to make explicit. See H. M. Fried and D. R. Yennie, *Phys. Rev.* **112**, 1391 (1958).

elements for QED from the correlation functions of non-gauge-invariant operators $\psi(x)$, $\bar{\psi}(x)$, and $A_\mu(x)$. We will now argue that the S-matrix elements are given correctly by this procedure. Since the S-matrix is defined between asymptotic states, we can compute S-matrix elements in a formalism in which the coupling constant is turned off adiabatically in the far past and far future. In the zero coupling limit, there is a clean separation between gauge-invariant and gauge-variant states. Single-particle states containing one electron, one positron, or one transversely polarized photon are gauge-invariant, while states with timelike and longitudinal photon polarizations transform under gauge motions. We can thus define a gauge-invariant S-matrix in the following way: Let S_{FP} be the S-matrix between general asymptotic states, computed from the Faddeev-Popov procedure. This matrix is unitary but not gauge-invariant. Let P_0 be a projection onto the subspace of the space of asymptotic states in which all particles are either electrons, positrons, or transverse photons. Then let

$$S = P_0 \, S_{\text{FP}} \, P_0 \,. \tag{9.59}$$

This S-matrix is gauge invariant by construction, because it is projected onto gauge-invariant states. It is not obvious that it is unitary. However, we addressed this issue in Section 5.5. We showed there that any matrix element $\mathcal{M}^\mu \epsilon^*_\mu$ for photon emission satisfies

$$\sum_{i=1,2} \epsilon^*_{i\mu} \epsilon_{i\nu} \mathcal{M}^\mu \mathcal{M}^{*\nu} = (-g_{\mu\nu}) \mathcal{M}^\mu \mathcal{M}^{*\nu}, \tag{9.60}$$

where the sum on the left-hand side runs only over transverse polarizations. The same argument applies if $\mathcal{M}^\mu$ and $\mathcal{M}^{*\nu}$ are distinct amplitudes, as long as they satisfy the Ward identity. This is exactly the information we need to see that

$$SS^\dagger = P_0 \, S_{\text{FP}} \, P_0 \, S^\dagger_{\text{FP}} \, P_0 = P_0 \, S_{\text{FP}} \, S^\dagger_{\text{FP}} \, P_0 . \tag{9.61}$$

Now we can use the unitarity of S_{FP} to see that S is unitary, $SS^\dagger = 1$, on the subspace of gauge-invariant states. It is easy to check explicitly that the formula (9.59) for the S-matrix is independent of ξ: The Ward identity implies that any QED matrix element with all external fermions on-shell is unchanged if we add to the photon propagator $D^{\mu\nu}(q)$ any term proportional to q^μ.

9.5 Functional Quantization of Spinor Fields

The functional methods that we have used so far allow us to compute, using Eq. (9.18) or (9.35), correlation functions involving fields that obey canonical commutation relations. To generalize these methods to include spinor fields, which obey canonical anticommutation relations, we must do something different: We must represent even the classical fields by anticommuting numbers.

Anticommuting Numbers

We will define anticommuting numbers (also called *Grassmann numbers*) by giving algebraic rules for manipulating them. These rules are formal and might seem *ad hoc*. We will justify them by showing that they lead to the familiar quantum theory of the Dirac equation.

The basic feature of anticommuting numbers is that they *anticommute*. For any two such numbers θ and η,

$$\theta\eta = -\eta\theta. \tag{9.62}$$

In particular, the square of any Grassmann number is zero:

$$\theta^2 = 0.$$

(This fact makes algebra extremely easy.) A product $(\theta\eta)$ of two Grassmann numbers commutes with other Grassmann numbers. We will also wish to add Grassmann numbers, and to multiply them by ordinary numbers; these operations have all the properties of addition and scalar multiplication in any vector space.

The main thing we want to do with anticommuting numbers is integrate over them. To define functional integration, we do not need general definite integrals of these parameters, but only the analog of $\int_{-\infty}^{\infty} dx$. So let us define the integral of a general function f of a Grassmann variable θ, over the complete range of θ:

$$\int d\theta\, f(\theta) = \int d\theta\, \big(A + B\theta\big).$$

In general, $f(\theta)$ can be expanded in a Taylor series, which terminates after two terms since $\theta^2 = 0$. The integral should be linear in f; thus it must be a linear function of A and B. Its value is fixed by one additional property: In our analysis of bosonic functional integrals (for instance, in (9.38) and (9.54)), we made strong use of the invariance of the integral to shifts of the integration variable. We will see in Section 9.6 that this shift invariance of the functional integral plays a central role in the derivation of the quantum mechanical equations of motion and conservation laws, and thus must be considered a fundamental aspect of the formalism. We must, then, demand this same property for integrals over θ. Invariance under the shift $\theta \to \theta + \eta$ yields the condition

$$\int d\theta\, \big(A + B\theta\big) = \int d\theta\, \big((A + B\eta) + B\theta\big).$$

The shift changes the constant term, but leaves the linear term unchanged. The only linear function of A and B that has this property is a constant (conventionally taken to be 1) times B, so we define*

$$\int d\theta\,\big(A + B\theta\big) = B. \tag{9.63}$$

When we perform a multiple integral over more than one Grassmann variable, an ambiguity in sign arises; we adopt the convention

$$\int d\theta \int d\eta\;\eta\theta = +1, \tag{9.64}$$

performing the innermost integral first.

Since the Dirac field is complex-valued, we will work primarily with complex Grassmann numbers, which can be built out of real and imaginary parts in the usual way. It is convenient to define complex conjugation to reverse the order of products, just like Hermitian conjugation of operators:

$$(\theta\eta)^* \equiv \eta^*\theta^* = -\theta^*\eta^*. \tag{9.65}$$

To integrate over complex Grassmann numbers, let us define

$$\theta = \frac{\theta_1 + i\theta_2}{\sqrt{2}}, \qquad \theta^* = \frac{\theta_1 - i\theta_2}{\sqrt{2}}.$$

We can now treat θ and θ^* as independent Grassmann numbers, and adopt the convention $\int d\theta^* d\theta\,(\theta\theta^*) = 1$.

Let us evaluate a Gaussian integral over a complex Grassmann variable:

$$\int d\theta^*\,d\theta\,e^{-\theta^* b\theta} = \int d\theta^*\,d\theta\,\big(1 - \theta^* b\theta\big) = \int d\theta^*\,d\theta\,\big(1 + \theta\theta^* b\big) = b. \tag{9.66}$$

If θ were an ordinary complex number, this integral would equal $2\pi/b$. The factor of 2π is unimportant; the main difference with anticommuting numbers is that the b comes out in the numerator rather than the denominator. However, if there is an additional factor of $\theta\theta^*$ in the integrand, we obtain

$$\int d\theta^*\,d\theta\;\theta\theta^*\,e^{-\theta^* b\theta} = 1 = \frac{1}{b}\cdot b. \tag{9.67}$$

The extra $\theta\theta^*$ introduces a factor of $(1/b)$, just as it does in an ordinary Gaussian integral.

To perform general Gaussian integrals in higher dimensions, we must first prove that an integral over complex Grassmann variables is invariant under unitary transformations. Consider a set of n complex Grassmann variables θ_i, and a unitary matrix U. If $\theta_i' = U_{ij}\theta_j$, then

$$\prod_i \theta_i' = \frac{1}{n!}\epsilon^{ij\dots l}\theta_i'\theta_j'\dots\theta_l'$$

*This definition is due to F. A. Berezin, *The Method of Second Quantization*, Academic Press, New York, 1966.

$$
\begin{aligned}
&= \frac{1}{n!}\epsilon^{ij\ldots l} U_{ii'}\theta_{i'} U_{jj'}\theta_{j'} \ldots U_{ll'}\theta_{l'} \\
&= \frac{1}{n!}\epsilon^{ij\ldots l} U_{ii'} U_{jj'} \ldots U_{ll'} \epsilon^{i'j'\ldots l'} \Big(\prod_i \theta_i\Big) \\
&= (\det U)\Big(\prod_i \theta_i\Big).
\end{aligned} \tag{9.68}
$$

In a general integral

$$\Big(\prod_i \int d\theta_i^* \, d\theta_i\Big) f(\theta),$$

the only term of $f(\theta)$ that survives has exactly one factor of each θ_i and θ_i^*; it is proportional to $(\prod \theta_i)(\prod \theta_i^*)$. If we replace θ by $U\theta$, this term acquires a factor of $(\det U)(\det U)^* = 1$, so the integral is unchanged under the unitary transformation.

We can now evaluate a general Gaussian integral involving a Hermitian matrix B with eigenvalues b_i:

$$\Big(\prod_i \int d\theta_i^* \, d\theta_i\Big) e^{-\theta_i^* B_{ij} \theta_j} = \Big(\prod_i \int d\theta_i^* \, d\theta_i\Big) e^{-\Sigma_i \theta_i^* b_i \theta_i} = \prod_i b_i = \det B. \tag{9.69}$$

(If θ were an ordinary number, we would have obtained $(2\pi)^n/(\det B)$.) Similarly, you can show that

$$\Big(\prod_i \int d\theta_i^* \, d\theta_i\Big) \, \theta_k \theta_l^* \, e^{-\theta_i^* B_{ij} \theta_j} = (\det B)(B^{-1})_{kl}. \tag{9.70}$$

Inserting another pair $\theta_m \theta_n^*$ in the integrand would yield a second factor $(B^{-1})_{mn}$, and a second term in which the indices l and n are interchanged (the sum of all possible pairings). In general, except for the determinant being in the numerator rather than the denominator, Gaussian integrals over Grassmann variables behave exactly like Gaussian integrals over ordinary variables.

The Dirac Propagator

A Grassmann *field* is a function of spacetime whose values are anticommuting numbers. More precisely, we can define a Grassmann field $\psi(x)$ in terms of any set of orthonormal basis functions:

$$\psi(x) = \sum_i \psi_i \phi_i(x). \tag{9.71}$$

The basis functions $\phi_i(x)$ are ordinary c-number functions, while the coefficients ψ_i are Grassmann numbers. To describe the Dirac field, we take the ϕ_i to be a basis of four-component spinors.

We now have all the machinery needed to evaluate functional integrals, and hence correlation functions, involving fermions. For example, the Dirac two-point function is given by

$$\langle 0| T\psi(x_1)\bar{\psi}(x_2) |0\rangle = \frac{\int \mathcal{D}\bar{\psi}\mathcal{D}\psi \, \exp\left[i\int d^4x\, \bar{\psi}(i\partial\!\!\!/ - m)\psi\right] \psi(x_1)\bar{\psi}(x_2)}{\int \mathcal{D}\bar{\psi}\mathcal{D}\psi \, \exp\left[i\int d^4x\, \bar{\psi}(i\partial\!\!\!/ - m)\psi\right]}.$$

(We write $\mathcal{D}\bar{\psi}$ instead of $\mathcal{D}\psi^*$ for convenience; the two are unitarily equivalent. We also leave the limits on the time integrals implicit; they are the same as in Eq. (9.18), and will yield an $i\epsilon$ term in the propagator as usual.) The denominator of this expression, according to (9.69), is $\det(i\partial\!\!\!/ - m)$. The numerator, according to (9.70), is this same determinant times the inverse of the operator $-i(i\partial\!\!\!/ - m)$. Evaluating this inverse in Fourier space, we find the familiar result for the Feynman propagator,

$$\langle 0| T\psi(x_1)\bar{\psi}(x_2) |0\rangle = S_F(x_1 - x_2) = \int \frac{d^4k}{(2\pi)^4} \frac{ie^{-ik\cdot(x_1-x_2)}}{k\!\!\!/ - m + i\epsilon}. \tag{9.72}$$

Higher correlation functions of free Dirac fields can be evaluated in a similar manner. The answer is always just the sum of all possible full contractions of the operators, with a factor of S_F for each contraction, as we found from Wick's theorem in Chapter 4.

Generating Functional for the Dirac Field

As with the Klein-Gordon field, we can alternatively derive the Feynman rules for the free Dirac theory by means of a generating functional. In analogy with (9.34), we define the Dirac generating functional as

$$Z[\bar{\eta}, \eta] = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \, \exp\left[i\int d^4x \left[\bar{\psi}(i\partial\!\!\!/ - m)\psi + \bar{\eta}\psi + \bar{\psi}\eta\right]\right], \tag{9.73}$$

where $\eta(x)$ is a Grassmann-valued source field. You can easily shift $\psi(x)$ to complete the square, to derive the simpler expression

$$Z[\bar{\eta}, \eta] = Z_0 \cdot \exp\left[-\int d^4x \, d^4y \, \bar{\eta}(x) S_F(x-y) \eta(y)\right], \tag{9.74}$$

where, as before, Z_0 is the value of the generating functional with the external sources set to zero.

To obtain correlation functions, we will differentiate Z with respect to η and $\bar{\eta}$. First, however, we must adopt a sign convention for derivatives with respect to Grassmann numbers. If η and θ are anticommuting numbers, let us define

$$\frac{d}{d\eta}\theta\eta = -\frac{d}{d\eta}\eta\theta = -\theta. \tag{9.75}$$

Then referring to the definition (9.73) of Z, we see that the two-point function, for example, is given by

$$\langle 0| T\psi(x_1)\bar{\psi}(x_2) |0\rangle = Z_0^{-1} \left(-i\frac{\delta}{\delta\bar{\eta}(x_1)}\right)\left(+i\frac{\delta}{\delta\eta(x_2)}\right) Z[\bar{\eta}, \eta]\bigg|_{\bar{\eta},\eta=0}.$$

Plugging in formula (9.74) for $Z[\bar{\eta}, \eta]$ and carefully keeping track of the signs, we find that this expression is equal to the Feynman propagator, $S_F(x_1 - x_2)$. Higher correlation functions can be evaluated in a similar way.

QED

As we saw in Section 9.2 for the case of scalar fields, the functional integral method allows us to read the Feynman rules for vertices directly from the Lagrangian for an interacting field theory. For the theory of Quantum Electrodynamics, the full Lagrangian is

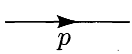
$$\begin{aligned}\mathcal{L}_{\text{QED}} &= \bar{\psi}(i\not{D} - m)\psi - \tfrac{1}{4}(F_{\mu\nu})^2 \\ &= \bar{\psi}(i\not{\partial} - m)\psi - \tfrac{1}{4}(F_{\mu\nu})^2 - e\bar{\psi}\gamma^\mu\psi A_\mu \\ &= \mathcal{L}_0 - e\bar{\psi}\gamma^\mu\psi A_\mu,\end{aligned}$$

where $D_\mu = \partial_\mu + ieA_\mu$ is the gauge-covariant derivative.

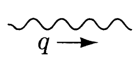
To evaluate correlation functions, we expand the exponential of the interaction term:

$$\exp\left[i\textstyle\int\mathcal{L}\right] = \exp\left[i\textstyle\int\mathcal{L}_0\right]\left[1 - ie\displaystyle\int d^4x\,\bar{\psi}\gamma^\mu\psi A_\mu + \cdots\right].$$

The two terms of the free Lagrangian yield the Dirac and electromagnetic propagators derived in this section and the last:

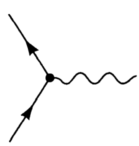


$$= \int \frac{d^4p}{(2\pi)^4}\,\frac{i\,e^{-ip\cdot(x-y)}}{\not{p} - m + i\epsilon};$$



$$= \int \frac{d^4q}{(2\pi)^4}\,\frac{-i\,g_{\mu\nu}\,e^{-iq\cdot(x-y)}}{q^2 + i\epsilon} \qquad \text{(Feynman gauge).}$$

The interaction term gives the QED vertex,



$$\mu = -ie\gamma^\mu \int d^4x.$$

As in Chapter 4, we can rearrange these rules, performing the integrations over vertex positions to obtain momentum-conserving delta functions, and using these delta functions to perform most of the propagator momentum integrals.

The only remaining aspect of the QED Feynman rules is the placement of various minus signs. These signs are also built into the functional integral; for example, interchanging θ_k and θ_l^* in Eq. (9.70) would introduce a factor of -1. We will see another example of a fermion minus sign in the computation that follows.

Functional Determinants

Throughout this chapter we have encountered expressions that we wrote formally as functional determinants. To end this section, let us investigate one of these objects more closely. We will find that, at least in this case, we can write the determinant explicitly as a sum of Feynman diagrams.

Consider the object

$$\int \mathcal{D}\bar{\psi}\,\mathcal{D}\psi\, \exp\Big[i\int d^4x\, \bar{\psi}(i\not{D}-m)\psi\Big], \tag{9.76}$$

where $D_\mu = \partial_\mu + ieA_\mu$ and $A_\mu(x)$ is a given external background field. Formally, this expression is a functional determinant:

$$\begin{aligned} &= \det(i\not{D}-m) = \det(i\not{\partial}-m-e\not{A}) \\ &\qquad\qquad = \det(i\not{\partial}-m)\cdot \det\Big(1-\frac{i}{i\not{\partial}-m}(-ie\not{A})\Big). \end{aligned}$$

In the last form, the first term is an infinite constant. The second term contains the dependence of the determinant on the external field A. We will now show that this dependence is well defined and, in fact, is exactly equivalent to the sum of vacuum diagrams.

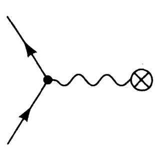
To demonstrate this, we need only apply standard identities from linear algebra. First notice that, if a matrix B has eigenvalues b_i, we can write its determinant as

$$\det B = \prod_i b_i = \exp\Big[\sum_i \log b_i\Big] = \exp\Big[\mathrm{Tr}(\log B)\Big], \tag{9.77}$$

where the logarithm of a matrix is defined by its power series. Applying this identity to our determinant, and writing out the power series of the logarithm, we obtain†

$$\det\Big(1-\frac{i}{i\not{\partial}-m}(-ie\not{A})\Big) = \exp\Big[\sum_{n=1}^{\infty} -\frac{1}{n}\,\mathrm{Tr}\Big[\Big(\frac{i}{i\not{\partial}-m}(-ie\not{A})\Big)^n\Big]\Big]. \tag{9.78}$$

Alternatively, we can evaluate this determinant by returning to expression (9.76) and using Feynman diagrams. Expanding the interaction term, we obtain the vertex rule

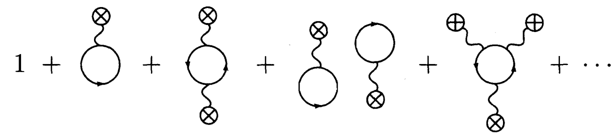


$$= -ie\gamma^\mu \int d^4x\, A_\mu(x).$$

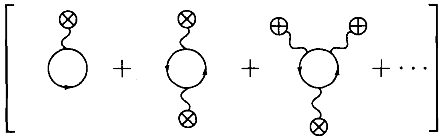
†We use Tr() to denote operator traces, and tr() to denote Dirac traces.

Our determinant is then equal to a sum of Feynman diagrams,

$$\det\left(1-\frac{i(-ie\not{A})}{i\not{\partial}-m}\right) = 1 +$$

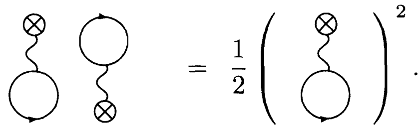


$$= \exp\Bigg[$$

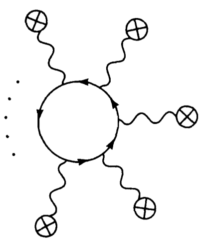


$$\Bigg] . \qquad (9.79)$$

The series exponentiates, since the disconnected diagrams are products of connected pieces (with appropriate symmetry factors when a piece is repeated). For example,



Now let us evaluate the nth diagram in the exponent of (9.79). There is a factor of -1 from the fermion loop, and a symmetry factor of $1/n$ since we could rotate the interactions around the diagram up to n times without changing it. (The factor is not $1/n!$, because the cyclic order of the interaction points is significant.) The diagram is therefore



$$\begin{aligned} &= -\frac{1}{n}\int dx_1\cdots dx_n \; \mathrm{tr}\Big[\big(-ie\not{A}(x_1)\big)S_F(x_2-x_1)\cdots \\ &\qquad\qquad \big(-ie\not{A}(x_n)\big)S_F(x_1-x_n)\Big] \\ &= -\frac{1}{n}\,\mathrm{Tr}\left[\left(\frac{i}{i\not{\partial}-m}(-ie\not{A})\right)^n\right], \end{aligned} \qquad (9.80)$$

in exact agreement with (9.78), including the minus sign and the symmetry factor.

The computation of functional determinants using Feynman diagrams is an important tool, as we will see in Chapter 11.

9.6 Symmetries in the Functional Formalism

We have now seen that the quantum field theoretic correlation functions of scalar, vector, and spinor fields can be computed from the functional integral, completely bypassing the construction of the Hamiltonian, the Hilbert space of states, and the equations of motion. The functional integral formalism makes the symmetries of the problem manifest; any invariance of the Lagrangian will be an invariance of the quantum dynamics.‡ However, we would like to be able to appeal also to the conservation laws that follow from the quantum equations of motion, or to these equations of motion themselves. For example, the Ward identity, which played a major role in our discussion of photons in QED (Section 5.5), is essentially the conservation law of the electric charge current. Since, as we saw in Section 2.2, the conservation laws follow from symmetries of the Lagrangian, one might guess that it is not difficult to derive these conservation laws from the functional integral. In this section we will see how to do that. We will see that the functional integral gives, in a most direct way, a quantum generalization of Noether's theorem. This result will lead to the analogue of the Ward-Takahashi identity for any symmetry of a general quantum field theory.

Equations of Motion

To prepare for this discussion, we should determine how the quantum equations of motion follow from the functional integral formalism. As a first problem to study, let us examine the Green's functions of the free scalar field. To be specific, consider the three-point function:

$$\langle\Omega| T\phi(x_1)\phi(x_2)\phi(x_3)\,|\Omega\rangle = Z^{-1}\int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}[\phi]}\phi(x_1)\phi(x_2)\phi(x_3), \qquad (9.81)$$

where $\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2$ and Z is a shorthand for $Z[J=0]$, the functional integral over the exponential. In classical mechanics, we would derive the equations of motion by insisting that the action be stationary under an infinitesimal variation

$$\phi(x) \to \phi'(x) = \phi(x) + \epsilon(x). \qquad (9.82)$$

The appropriate generalization is to consider (9.82) as an infinitesimal change of variables. A change of variables does not alter the value of the integral. Nor does a shift of the integration variable alter the measure: $\mathcal{D}\phi' = \mathcal{D}\phi$. Thus we can write

$$\int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}[\phi]}\phi(x_1)\phi(x_2)\phi(x_3) = \int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}[\phi']}\phi'(x_1)\phi'(x_2)\phi'(x_3),$$

‡There are some subtle exceptions to this rule, which we will treat in Chapter 19.

where $\phi' = \phi + \epsilon$. Expanding this equation to first order in ϵ, we find

$$0 = \int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}} \bigg\{ \Big(i\int d^4x\, \epsilon(x) \Big[(-\partial^2 - m^2)\phi(x)\Big] \phi(x_1)\phi(x_2)\phi(x_3) \Big) + \epsilon(x_1)\phi(x_2)\phi(x_3) + \phi(x_1)\epsilon(x_2)\phi(x_3) + \phi(x_1)\phi(x_2)\epsilon(x_3) \bigg\}. \tag{9.83}$$

The last three terms can be combined with the first by writing, for instance, $\epsilon(x_1) = \int d^4x\, \epsilon(x)\delta(x - x_1)$. Noting that the right-hand side must vanish for any possible variation $\epsilon(x)$, we then obtain

$$0 = \int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}} \Big[(\partial^2 + m^2)\phi(x)\, \phi(x_1)\phi(x_2)\phi(x_3) + i\delta(x - x_1)\phi(x_2)\phi(x_3) + i\phi(x_1)\delta(x - x_2)\phi(x_3) + i\phi(x_1)\phi(x_2)\delta(x - x_3) \Big]. \tag{9.84}$$

A similar equation holds for any number of fields $\phi(x_i)$.

To see the implications of (9.84), let us specialize to the case of one field $\phi(x_1)$ in (9.81). Notice that the derivatives acting on $\phi(x)$ can be pulled outside the functional integral. Then, dividing (9.84) by Z yields the identity

$$(\partial^2 + m^2) \langle \Omega | T\phi(x)\phi(x_1) |\Omega\rangle = -i\delta(x - x_1). \tag{9.85}$$

The left-hand side of this relation is the Klein-Gordon operator acting on a correlation function of $\phi(x)$. The right-hand side is zero unless $x = x_1$; that is, the correlation function satisfies the Klein-Gordon equation except at the point where the arguments of the two ϕ fields coincide. The modification of the Klein-Gordon equation at this point is called a *contact term*. In this simple case, the modification is hardly unfamiliar to us; Eq. (9.85) merely says that the Feynman propagator is a Green's function of the Klein-Gordon operator, as we originally showed in Section 2.4. We saw there that the delta function arises when the time derivative in ∂^2 acts on the time-ordering symbol. We will see below that, quite generally in quantum field theory, the classical equations of motion for fields are satisfied by all quantum correlation functions of those fields, up to contact terms.

As an example, consider the identity that follows from (9.84) for an $(n+1)$-point correlation function of scalar fields:

$$(\partial^2 + m^2) \langle \Omega | T\phi(x)\phi(x_1)\cdots\phi(x_n) |\Omega\rangle = \sum_{i=1}^{n} \langle \Omega | T\phi(x_1)\cdots \big(-i\delta(x - x_i)\big) \cdots \phi(x_n) |\Omega\rangle . \tag{9.86}$$

This identity says that the Klein-Gordon equation is obeyed by $\phi(x)$ inside any expectation value, up to contact terms associated with the time ordering. The result can also be derived from the Hamiltonian formalism using the methods of Section 2.4, or, using the special properties of free-field theory, by evaluating both sides of the equation using Wick's theorem.

As long as the functional measure is invariant under a shift of the integration variable, we can repeat this argument and obtain the quantum equations of motion for Green's functions for any theory of scalar, vector, and spinor fields. This is the reason why, in Eq. (9.63), we took the shift invariance to be the fundamental, defining property of the Grassmann integral.

For a general field theory of a field $\varphi(x)$, governed by the Lagrangian $\mathcal{L}[\varphi]$, the manipulations leading to (9.83) give the identity

$$\begin{aligned} 0 = \int \mathcal{D}\varphi\, e^{i\int d^4x\,\mathcal{L}} \Big\{ i \int d^4x\, \epsilon(x) \frac{\delta}{\delta\varphi(x)} \Big(\int d^4x' \mathcal{L} \Big) \cdot \varphi(x_1)\varphi(x_2) \\ + \epsilon(x_1)\varphi(x_2) + \varphi(x_1)\epsilon(x_2) \Big\}, \end{aligned} \tag{9.87}$$

and similar identities for correlation functions of n fields. By the rule for functional differentiation (9.31), the derivative of the action is

$$\frac{\delta}{\delta\varphi(x)} \Big(\int d^4x' \mathcal{L} \Big) = \frac{\partial \mathcal{L}}{\partial\varphi} - \partial_\mu \Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \Big);$$

this is the quantity that equals zero by the Euler-Lagrange equation of motion (2.3) for φ. Formula (9.87) and its generalizations lead to the set of identities

$$\Big\langle \Big(\frac{\delta}{\delta\varphi(x)} \int d^4x' \mathcal{L} \Big) \varphi(x_1)\cdots\varphi(x_n) \Big\rangle = \sum_{i=1}^{n} \big\langle \varphi(x_1)\cdots \big(i\delta(x-x_i)\big)\cdots\varphi(x_n) \big\rangle . \tag{9.88}$$

In this equation, the angle-brackets denote a time-ordered correlation function in which derivatives on $\varphi(x)$ are placed outside the time-ordering symbol, as in Eq. (9.86). Relation (9.88) states that the classical Euler-Lagrange equations of the field φ are obeyed for all Green's functions of φ, up to contact terms arising from the nontrivial commutation relations of field operators. These quantum equations of motion for Green's functions, including the proper contact terms, are called *Schwinger-Dyson equations.*

Conservation Laws

In classical field theory, Noether's theorem says that, to each symmetry of a local Lagrangian, there corresponds a conserved current. In Section 2.2 we proved Noether's theorem by subjecting the Lagrangian to an infinitesimal symmetry variation. In the spirit of the above discussion of equations of motion, we should find the quantum analogue of this theorem by subjecting the functional integral to an infinitesimal change of variables along the symmetry direction.

Again, it will be most instructive to begin with an example. Let us consider the theory of a free, complex-valued scalar field, with the Lagrangian

$$\mathcal{L} = |\partial_\mu\phi|^2 - m^2|\phi|^2. \tag{9.89}$$

This Lagrangian is invariant under the transformation $\phi \to e^{i\alpha}\phi$. The classical consequences of this invariance were discussed in Section 2.2, below Eq. (2.14). To find the quantum formulae, consider the infinitesimal change of variables

$$\phi(x) \to \phi'(x) = \phi(x) + i\alpha(x)\phi(x). \tag{9.90}$$

Note that we have made the infinitesimal angle of rotation a function of x; the reason for this will be clear in a moment.

The measure of functional integration is invariant under the transformation (9.90), since this is a unitary transformation of the variables $\phi(x)$. Thus, for the case of two fields,

$$\int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}[\phi]}\phi(x_1)\phi^*(x_2) = \int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}[\phi']}\phi'(x_1)\phi'^*(x_2)\bigg|_{\phi'=(1+i\alpha)\phi}.$$

Expanding this equation to first order in α, we find

$$\begin{aligned}0 = \int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}} \bigg\{ i\int d^4x \Big[(\partial_\mu\alpha)\cdot i(\phi\partial^\mu\phi^* - \phi^*\partial^\mu\phi)\Big]\phi(x_1)\phi^*(x_2) \\ + \big[i\alpha(x_1)\phi(x_1)\big]\phi^*(x_2) + \phi(x_1)\big[-i\alpha(x_2)\phi^*(x_2)\big]\bigg\}.\end{aligned}$$

Notice that the variation of the Lagrangian contains only terms proportional to $\partial_\mu\alpha$, since the substitution (9.90) with a constant α leaves the Lagrangian invariant. To put this relation into a familiar form, integrate the term involving $\partial_\mu\alpha$ by parts. Then taking the coefficient of $\alpha(x)$ and dividing by Z gives

$$\begin{aligned}\big\langle \partial_\mu j^\mu(x)\phi(x_1)\phi^*(x_2)\big\rangle = (-i)\Big\langle \big(i\phi(x_1)\delta(x-x_1)\big)\phi^*(x_2) \\ + \phi(x_1)\big(-i\phi^*(x_2)\delta(x-x_2)\big)\Big\rangle,\end{aligned} \tag{9.91}$$

where

$$j^\mu = i(\phi\partial^\mu\phi^* - \phi^*\partial^\mu\phi) \tag{9.92}$$

is the Noether current identified in Eq. (2.16). As in Eq. (9.88), the correlation function denotes a time-ordered product with the derivative on $j^\mu(x)$ placed outside the time-ordering symbol. Relation (9.91) is the classical conservation law plus contact terms, that is, the Schwinger-Dyson equation associated with current conservation.

It is not much more difficult to discuss current conservation in more general situations. Consider a local field theory of a set of fields $\varphi_a(x)$, governed by a Lagrangian $\mathcal{L}[\varphi]$. An infinitesimal symmetry transformation on the fields φ_a will be of the general form

$$\varphi_a(x) \to \varphi_a(x) + \epsilon\Delta\varphi_a(x). \tag{9.93}$$

We assume that the action is invariant under this transformation. Then, as in Eq. (2.10), if the parameter ϵ is taken to be a constant, the Lagrangian must be invariant up to a total divergence:

$$\mathcal{L}[\varphi] \to \mathcal{L}[\varphi] + \epsilon\partial_\mu\mathcal{J}^\mu. \tag{9.94}$$

If the symmetry parameter ϵ depends on x, as in the analysis of the previous paragraph, the variation of the Lagrangian will be slightly more complicated:

$$\mathcal{L}[\varphi] \to \mathcal{L}[\varphi] + (\partial_\mu \epsilon)\Delta\varphi_a \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_a)} + \epsilon \partial_\mu \mathcal{J}^\mu.$$

Summation over the index a is understood. Then

$$\frac{\delta}{\delta\epsilon(x)} \int d^4x\, \mathcal{L}[\varphi + \epsilon\Delta\varphi] = -\partial_\mu j^\mu(x), \tag{9.95}$$

where j^μ is the Noether current of Eq. (2.12),

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_a)} \Delta\varphi_a - \mathcal{J}^\mu. \tag{9.96}$$

Using result (9.95) and carrying through the steps leading up to (9.91), we find the Schwinger-Dyson equation:

$$\begin{aligned} \big\langle \partial_\mu j^\mu(x)\varphi_a(x_1)\varphi_b(x_2)\big\rangle = (-i)\Big\langle &(\Delta\varphi_a(x_1)\delta(x-x_1))\varphi_b(x_2) \\ &+ \varphi_a(x_1)(\Delta\varphi_b(x_2)\delta(x-x_2))\Big\rangle. \end{aligned} \tag{9.97}$$

A similar equation can be found for the correlator of $\partial_\mu j^\mu$ with n fields $\varphi(x)$. These give the full set of Schwinger-Dyson equations associated with the classical Noether theorem.

As an example of the use of this variational procedure to obtain the Noether current, consider the symmetry of the Lagrangian with respect to spacetime translations. Under the transformation

$$\varphi_a \to \varphi_a + a^\mu(x)\partial_\mu \varphi_a \tag{9.98}$$

the Lagrangian transforms as

$$\mathcal{L} \to \mathcal{L} + \partial_\nu a^\mu \partial_\mu \varphi_a \frac{\partial \mathcal{L}}{\partial(\partial_\nu \varphi_a)} + a^\mu \partial_\mu \mathcal{L}.$$

The variation of $\int d^4x\, \mathcal{L}$ with respect to a^μ then gives rise to the conservation equation for the energy-momentum tensor $\partial_\nu T^{\mu\nu} = 0$, with

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\nu \varphi_a)} \partial^\mu \varphi_a - g^{\mu\nu}\mathcal{L}, \tag{9.99}$$

in agreement with Eq. (2.17).

The trick we have used in this section, that of considering a symmetry transformation whose parameter is a function of spacetime, is reminiscent of a technical feature of our earlier discussion introducing the Lagrangian of QED. In Eq. (4.6), we noted that the minimal coupling prescription for coupling the photon to charged fields produces a Lagrangian invariant not only under the global symmetry transformation with ϵ constant, but also under a transformation in which the symmetry parameter depends on x. In Chapter 15, we will draw these two ideas together in a general discussion of field theories with *local* symmetries.

The Ward-Takahashi Identity

As a final application of the methods of this section, let us derive the Schwinger-Dyson equations associated with the global symmetry of QED. Consider making, in the QED functional integral, the change of variables

$$\psi(x) \to (1 + ie\alpha(x))\psi(x), \tag{9.100}$$

without the corresponding term in the transformation law for A_μ (which would make the Lagrangian invariant under the transformation). The QED Lagrangian (4.3) then transforms according to

$$\mathcal{L} \to \mathcal{L} - e\partial_\mu \alpha \bar{\psi}\gamma^\mu \psi. \tag{9.101}$$

The transformation (9.100) thus leads to the following identity for the functional integral over two fermion fields:

$$\begin{aligned} 0 = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}A\, e^{i\int d^4x\,\mathcal{L}} \bigg\{ &- i\int d^4x\, \partial_\mu\alpha(x) \Big[j^\mu(x)\psi(x_1)\bar{\psi}(x_2) \Big] \\ &+ \big(ie\alpha(x_1)\psi(x_1)\big)\bar{\psi}(x_2) + \psi(x_1)\big(-ie\alpha(x_2)\bar{\psi}(x_2)\big) \bigg\}, \end{aligned} \tag{9.102}$$

with $j^\mu = e\bar{\psi}\gamma^\mu\psi$. As in our other examples, an analogous equation holds for any number of fermion fields.

To understand the implications of this set of equations, consider first the specific case (9.102). Dividing this relation by Z, we find

$$\begin{aligned} i\partial_\mu \langle 0| T j^\mu(x)\psi(x_1)\bar{\psi}(x_2) |0\rangle = &- ie\delta(x - x_1) \langle 0| T\psi(x_1)\bar{\psi}(x_2) |0\rangle \\ &+ ie\delta(x - x_2) \langle 0| T\psi(x_1)\bar{\psi}(x_2) |0\rangle . \end{aligned} \tag{9.103}$$

To put this equation into a more familiar form, compute its Fourier transform by integrating:

$$\int d^4x\, e^{-ik\cdot x} \int d^4x_1\, e^{+iq\cdot x_1} \int d^4x_2\, e^{-ip\cdot x_2}. \tag{9.104}$$

Then the amplitudes in (9.103) are converted to the amplitudes $\mathcal{M}(k;p;q)$ and $\mathcal{M}(p;q)$ defined below (7.67) in our discussion of the Ward-Takahashi identity. Indeed, (9.103) falls directly into the form

$$-ik_\mu \mathcal{M}^\mu(k;p;q) = -ie\mathcal{M}_0(p;q-k) + ie\mathcal{M}_0(p+k;q). \tag{9.105}$$

This is exactly the Ward-Takahashi identity for two external fermions, which we derived diagrammatically in Section 7.4. It is not difficult to check that the more general relations involving n fermion fields lead to the general Ward-Takahashi identity presented in (7.68). Because of this relation, the formula (9.97) associated with the arbitrary symmetry (9.93) is usually also referred to as a Ward-Takahashi identity, the one associated with the symmetry and its Noether current.

We have now arrived at a more general understanding of the terms on the right-hand side of the Ward-Takahashi identity. These are the contact terms that we now expect to find when we convert classical equations of motion to Schwinger-Dyson equations for quantum Green's functions. The functional integral formalism allows a simple and elegant derivation of these quantum-mechanical terms.

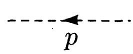
Problems

9.1 Scalar QED. This problem concerns the theory of a complex scalar field ϕ interacting with the electromagnetic field A^μ. The Lagrangian is

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}^2 + (D_\mu\phi)^*(D^\mu\phi) - m^2\phi^*\phi,$$

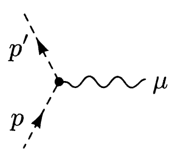
where $D_\mu = \partial_\mu + ieA_\mu$ is the usual gauge-covariant derivative.

(a) Use the functional method of Section 9.2 to show that the propagator of the complex scalar field is the same as that of a real field:

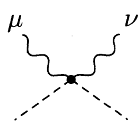


$$= \frac{i}{p^2 - m^2 + i\epsilon}.$$

Also derive the Feynman rules for the interactions between photons and scalar particles; you should find



$$= -ie(p+p')^\mu;$$



$$= 2ie^2 g^{\mu\nu}.$$

(b) Compute, to lowest order, the differential cross section for $e^+e^- \to \phi\phi^*$. Ignore the electron mass (but not the scalar particle's mass), and average over the electron and positron polarizations. Find the asymptotic angular dependence and total cross section. Compare your results to the corresponding formulae for $e^+e^- \to \mu^+\mu^-$.

(c) Compute the contribution of the charged scalar to the photon vacuum polarization, using dimensional regularization. Note that there are two diagrams. To put the answer into the expected form,

$$\Pi^{\mu\nu}(q^2) = (g^{\mu\nu}q^2 - q^\mu q^\nu)\Pi(q^2),$$

it is useful to add the two diagrams at the beginning, putting both terms over a common denominator before introducing a Feynman parameter. Show that, for $-q^2 \gg m^2$, the charged boson contribution to $\Pi(q^2)$ is exactly 1/4 that of a virtual electron-positron pair.

9.2 Quantum statistical mechanics.

(a) Evaluate the quantum statistical partition function

$$Z = \mathrm{tr}\left[e^{-\beta H}\right]$$

(where $\beta = 1/kT$) using the strategy of Section 9.1 for evaluating the matrix elements of e^{-iHt} in terms of functional integrals. Show that one again finds a functional integral, over functions defined on a domain that is of length β and periodically connected in the time direction. Note that the Euclidean form of the Lagrangian appears in the weight.

(b) Evaluate this integral for a simple harmonic oscillator,

$$L_E = \tfrac{1}{2}\dot{x}^2 + \tfrac{1}{2}\omega^2 x^2,$$

by introducing a Fourier decomposition of x(t):

$$x(t) = \sum_n x_n \cdot \frac{1}{\sqrt{\beta}} e^{2\pi i n t/\beta}.$$

The dependence of the result on β is a bit subtle to obtain explicitly, since the measure for the integral over $x(t)$ depends on β in any discretization. However, the dependence on ω should be unambiguous. Show that, up to a (possibly divergent and β-dependent) constant, the integral reproduces exactly the familiar expression for the quantum partition function of an oscillator. [You may find the identity

$$\sinh z = z \cdot \prod_{n=1}^{\infty}\left(1 + \frac{z^2}{(n\pi)^2}\right)$$

useful.]

(c) Generalize this construction to field theory. Show that the quantum statistical partition function for a free scalar field can be written in terms of a functional integral. The value of this integral is given formally by

$$\left[\det(-\partial^2 + m^2)\right]^{-1/2},$$

where the operator acts on functions on Euclidean space that are periodic in the time direction with periodicity β. As before, the β dependence of this expression is difficult to compute directly. However, the dependence on m^2 is unambiguous. (More generally, one can usually evaluate the variation of a functional determinant with respect to any explicit parameter in the Lagrangian.) Show that the determinant indeed reproduces the partition function for relativistic scalar particles.

(d) Now let $\psi(t)$, $\bar{\psi}(t)$ be two Grassmann-valued coordinates, and define a fermionic oscillator by writing the Lagrangian

$$L_E = \bar{\psi}\dot{\psi} + \omega\bar{\psi}\psi.$$

This Lagrangian corresponds to the Hamiltonian

$$H = \omega\bar{\psi}\psi, \quad \text{with } \{\bar{\psi}, \psi\} = 1;$$

that is, to a simple two-level system. Evaluate the functional integral, assuming that the fermions obey antiperiodic boundary conditions: $\psi(t+\beta) = -\psi(t)$. (Why is this reasonable?) Show that the result reproduces the partition function of a quantum-mechanical two-level system, that is, of a quantum state with Fermi statistics.

(e) Define the partition function for the photon field as the gauge-invariant functional integral

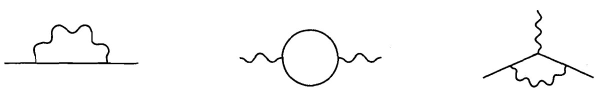
$$Z = \int \mathcal{D}A \, \exp\Big(-\int d^4x_E \left[\tfrac{1}{4}(F_{\mu\nu})^2\right]\Big)$$

over vector fields A_μ that are periodic in the time direction with period β. Apply the gauge-fixing procedure discussed in Section 9.4 (working, for example, in Feynman gauge). Evaluate the functional determinants using the result of part (c) and show that the functional integral does give the correct quantum statistical result (including the correct counting of polarization states).

Chapter 10

Systematics of Renormalization

While computing radiative corrections in Chapters 6 and 7, we encountered three QED diagrams with ultraviolet divergences:



In each case we saw that the divergence could be regulated and canceled, yielding finite expressions for measurable quantities. In Chapter 8, we pointed out that such ultraviolet divergences occur commonly and, in fact, naturally in quantum field theory calculations. We sketched a physical interpretation of these divergences, with implications both in quantum field theory and in the statistical theory of phase transitions. In the next few chapters, we will convert this sketchy picture into a quantitative theory that allows precise calculations.

In this chapter, we begin this study by developing a classification of the ultraviolet divergences that can appear in a quantum field theory. Rather than stumbling across these divergences one by one and repairing them case by case, we now set out to determine once and for all which diagrams are divergent, and in which theories these divergences can be eliminated systematically. As examples we will consider both QED and scalar field theories.

10.1 Counting of Ultraviolet Divergences

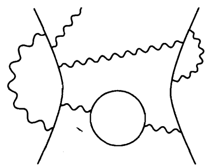
In this section we will use elementary arguments to determine, tentatively, when a Feynman diagram contains an ultraviolet divergence. We begin by analyzing quantum electrodynamics.

First we introduce the following notation, to characterize a typical diagram in QED:

$$
\begin{aligned}
N_e &= \text{number of external electron lines;}\\
N_\gamma &= \text{number of external photon lines;}\\
P_e &= \text{number of electron propagators;}\\
P_\gamma &= \text{number of photon propagators;}\\
V &= \text{number of vertices;}\\
L &= \text{number of loops.}
\end{aligned}
$$

(This analysis applies to correlation functions as well as scattering amplitudes. In the former case, propagators that are connected to external points should be counted as external lines, not as propagators.)

The expression corresponding to a typical diagram looks like this:



$$\sim \int \frac{d^4k_1\, d^4k_2 \cdots d^4k_L}{(\not{k}_i - m)\cdots(k_j^2)\cdots(k_n^2)}.$$

For each loop there is a potentially divergent 4-momentum integral, but each propagator aids the convergence of this integral by putting one or two powers of momentum into the denominator. Very roughly speaking, the diagram diverges unless there are more powers of momentum in the denominator than in the numerator. Let us therefore define the *superficial degree of divergence*, D, as the difference:

$$\begin{aligned} D &\equiv (\text{power of } k \text{ in numerator}) - (\text{power of } k \text{ in denominator}) \\ &= 4L - P_e - 2P_\gamma. \end{aligned} \tag{10.1}$$

Naively, we expect a diagram to have a divergence proportional to Λ^D, where Λ is a momentum cutoff, when $D > 0$. We expect a divergence of the form $\log\Lambda$ when $D = 0$, and no divergence when $D < 0$.

This naive expectation is often wrong, for one of three reasons (see Fig. 10.1). When a diagram contains a divergent subdiagram, its actual divergence may be worse than that indicated by D. When symmetries (such as the Ward identity) cause certain terms to cancel, the divergence of a diagram may be reduced or even eliminated. Finally, a trivial diagram with no propagators and no loops has $D = 0$ but no divergence.

Despite all of these complications, D is still a useful quantity. To see why, let us rewrite it in terms of the number of external lines (N_e, N_γ) and vertices (V). Note that the number of loop integrations in a diagram is

$$L = P_e + P_\gamma - V + 1, \tag{10.2}$$

since in our original Feynman rules each propagator has a momentum integral, each vertex has a delta function, and one delta function merely enforces overall momentum conservation. Furthermore, the number of vertices is

$$V = 2P_\gamma + N_\gamma = \tfrac{1}{2}(2P_e + N_e), \tag{10.3}$$

since each vertex involves exactly one photon line and two electron lines. (The propagators count twice since they have two ends on vertices.) Putting these relations together, we find that D can be expressed as

$$\begin{aligned} D &= 4(P_e + P_\gamma - V + 1) - P_e - 2P_\gamma \\ &= 4 - N_\gamma - \tfrac{3}{2}N_e, \end{aligned} \tag{10.4}$$

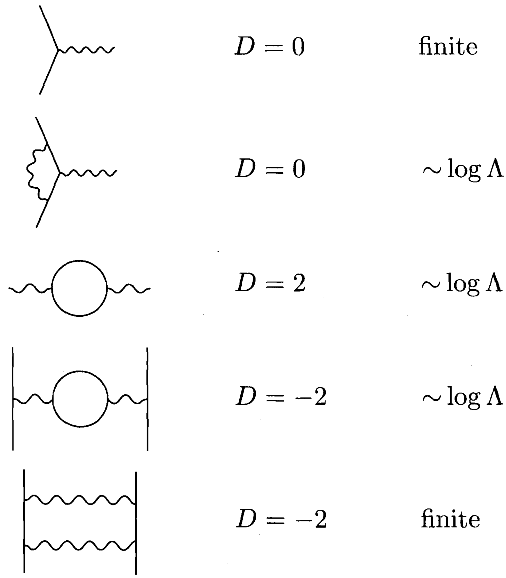


Figure 10.1. Some simple QED diagrams that illustrate the superficial degree of divergence. The first diagram is finite, even though $D = 0$. The third diagram has $D = 2$ but only a logarithmic divergence, due to the Ward identity (see Section 7.5). The fourth diagram diverges, even though $D < 0$, since it contains a divergent subdiagram. Only in the second and fifth diagrams does the superficial degree of divergence coincide with the actual degree of divergence.

independent of the number of vertices. The superficial degree of divergence of a QED diagram depends only on the number of external legs of each type.

According to result (10.4), only diagrams with a small number of external legs have $D \geq 0$; those seven types of diagrams are shown in Fig. 10.2. Since external legs do not enter the potentially divergent integral, we can restrict our attention to amputated diagrams. We can also restrict our attention to one-particle-irreducible diagrams, since reducible diagrams are simple products of the integrals corresponding to their irreducible parts. Thus the task of enumerating all of the divergent QED diagrams reduces to that of analyzing the seven types of amputated, one-particle-irreducible amplitudes shown in Fig. 10.2. Other diagrams may diverge, but only when they contain one of these seven as a subdiagram. Let us therefore consider each of these seven amplitudes in turn.

The zero-point function, Fig. 10.2a, is very badly divergent. But this object merely causes an unobservable shift of the vacuum energy; it never contributes to S-matrix elements.

To analyze the photon one-point function (Fig. 10.2b), note that the external photon must be attached to a QED vertex. Neglecting the external

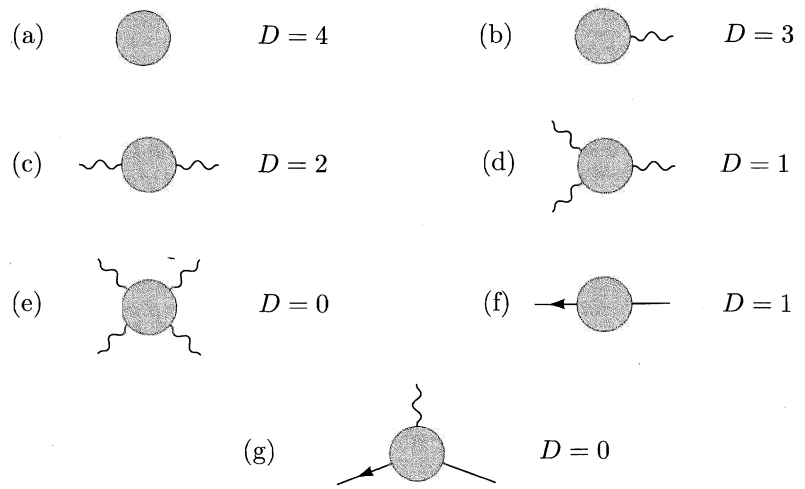
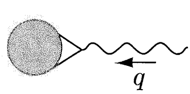


Figure 10.2. The seven QED amplitudes whose superficial degree of divergence (D) is ≥ 0. (Each circle represents the sum of all possible QED diagrams.) As explained in the text, amplitude (a) is irrelevant to scattering processes, while amplitudes (b) and (d) vanish because of symmetries. Amplitude (e) is nonzero, but its divergent parts cancel due to the Ward identity. The remaining amplitudes (c, f, and g) are all logarithmically divergent, even though $D > 0$ for (c) and (f).

photon propagator, this amplitude is therefore



$$= -ie \int d^4x\, e^{-iq\cdot x} \langle\Omega|\, T\, j_\mu(x)\, |\Omega\rangle\,, \tag{10.5}$$

where $j^\mu = \bar{\psi}\gamma^\mu\psi$ is the electromagnetic current operator. But the vacuum expectation value of j^μ must vanish by Lorentz invariance, since otherwise it would be a preferred 4-vector.

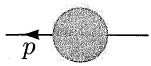
The photon one-point function also vanishes for a second reason: charge-conjugation invariance. Recall that C is a symmetry of QED, so $C\,|\Omega\rangle = |\Omega\rangle$. But $j^\mu(x)$ changes sign under charge conjugation, $Cj^\mu(x)C^\dagger = -j^\mu(x)$, so its vacuum expectation value must vanish:

$$\langle\Omega|\, Tj^\mu(x)\,|\Omega\rangle = \langle\Omega|\, C^\dagger C j^\mu(x) C^\dagger C\,|\Omega\rangle = -\,\langle\Omega|\, Tj^\mu(x)\,|\Omega\rangle = 0.$$

The same argument applies to any vacuum expectation value of an odd number of electromagnetic currents. In particular, the photon three-point function, Fig. 10.2d, vanishes. (This result is known as Furry's theorem.) It is not hard to check explicitly that the photon one- and three-point functions vanish in the leading order of perturbation theory (see Problem 10.1).

The remaining amplitudes in Fig. 10.2 are all nonzero, so we must analyze their structures in more detail. Consider, for example, the electron self-energy

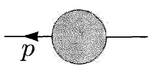
(Fig. 10.2f). This amplitude is a function of the electron momentum p, so let us expand it in a Taylor series about $p = 0$:



$$= A_0 + A_1 \not{p} + A_2 p^2 + \cdots ,$$

where each coefficient is independent of p:

$$A_n = \frac{1}{n!} \frac{d^n}{d\not{p}^n} \Big(\quad \Big) \Big|_{\not{p}=0} .$$

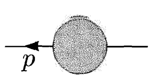


(These coefficients are infrared divergent; to compute them explicitly we would need an infrared regulator, as in Chapter 6.) The diagrams contributing to the electron self-energy depend on p through the denominators of propagators. To compute the coefficients A_n, we differentiate these propagators, giving expressions like

$$\frac{d}{d\not{p}} \left(\frac{1}{\not{k} + \not{p} - m} \right) = - \frac{1}{(\not{k} + \not{p} - m)^2} .$$

That is, each derivative with respect to the external momentum p lowers the superficial degree of divergence by 1. Since the constant term A_0 has (superficially) a linear divergence, A_1 can have only a logarithmic divergence; all the remaining A_n are finite. (This argument breaks down when the divergence is in a subdiagram, since then not all propagators involve the large momentum k. We will face this problem in Section 10.4.)

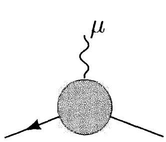
The electron self-energy amplitude has one additional subtlety. If the constant term A_0 were proportional to Λ (the ultraviolet cutoff), the electron mass shift would, according to the analysis in Section 7.1, also have a term proportional to Λ. But the electron mass shift must actually be proportional to m, since chiral symmetry would forbid a mass shift if m were zero. At worst, the constant term can be proportional to $m \log \Lambda$. We therefore expect the entire self-energy amplitude to have the form



$$= a_0 m \log \Lambda + a_1 \not{p} \log \Lambda + (\text{finite terms}), \qquad (10.6)$$

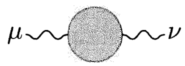
exactly what we found for the term of order α in Eq. (7.19).

Let us analyze the exact electron-photon vertex, Fig. 10.2g, in the same way. (Again we implicitly assume that infrared divergences have been regulated.) Expanding in powers of the three external momenta, we immediately see that only the constant term is divergent, since differentiating with respect to any external momentum would lower the degree of divergence to -1. This amplitude therefore contains only one divergent constant:



$$\propto -ie\gamma^\mu \log \Lambda + \text{finite terms}. \qquad (10.7)$$

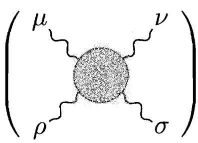
As discussed in Section 7.5, the photon self energy (Fig. 10.2c) is constrained by the Ward identity to have the form



$$= (g^{\mu\nu}q^2 - q^\mu q^\nu)\Pi(q^2). \tag{10.8}$$

Viewing this expression as a Taylor series in q, we see that the constant and linear terms both vanish, lowering the superficial degree of divergence from 2 to 0. The only divergence, therefore, is in the constant term of $\Pi(q^2)$, and this divergence is only logarithmic. This result is exactly what we found for the lowest-order contribution to $\Pi(q^2)$ in Eq. (7.90).

Finally, consider the photon-photon scattering amplitude, Fig. 10.2e. The Ward identity requires that if we replace any external photon by its momentum vector, the amplitude vanishes:

k^μ 

$$= 0. \tag{10.9}$$

By exhaustion one can show that this condition is satisfied only if the amplitude is proportional to $(g^{\mu\nu}k^\sigma - g^{\mu\sigma}k^\nu)$, with a similar factor for each of the other three legs. Each of these factors involves one power of momentum, so all terms with less than four powers of momentum in the Taylor series of this amplitude must vanish. The first nonvanishing term has $D = 0 - 4 = -4$, and therefore this amplitude is finite.

In summary, we have found that there are only three "primitively" divergent amplitudes in QED: the three that we already found in Chapters 6 and 7. (Other amplitudes may also be divergent, but only because of diagrams that contain these primitive amplitudes as components.) Furthermore, the dependence of these divergent amplitudes on external momenta is extremely simple. If we expand each amplitude as a power series in its external momenta, there are altogether only four divergent coefficients in the expansions. In other words, QED contains only four divergent numbers. In the next section we will see how these numbers can be absorbed into unobservable Lagrangian parameters, so that observable scattering amplitudes are always finite.

For the remainder of this section, let us try to understand the superficial degree of divergence from a more general viewpoint. The theory of QED in four spacetime dimensions is rather special, so let us first generalize to QED in d dimensions. In this case, D is given by

$$D \equiv dL - P_e - 2P_\gamma, \tag{10.10}$$

since each loop contributes a d-dimensional momentum integral. Relations (10.2) and (10.3) still hold, so we can again rewrite D in terms of V, N_e, and N_γ. This time the result is

$$D = d + \left(\frac{d-4}{2}\right)V - \left(\frac{d-2}{2}\right)N_\gamma - \left(\frac{d-1}{2}\right)N_e. \tag{10.11}$$

The cancellation of V in this expression is special to the case $d = 4$. For $d < 4$, diagrams with more vertices have a lower degree of divergence, so the total number of divergent *diagrams* is finite. For $d > 4$, diagrams with more vertices have a higher degree of divergence, so every amplitude becomes superficially divergent at a sufficiently high order in perturbation theory.

These three possible types of ultraviolet behavior will also occur in other quantum field theories. We will refer to them as follows:

Super-Renormalizable theory:	Only a finite number of Feynman diagrams superficially diverge.
Renormalizable theory:	Only a finite number of amplitudes superficially diverge; however, divergences occur at all orders in perturbation theory.
Non-Renormalizable theory:	All amplitudes are divergent at a sufficiently high order in perturbation theory.

Using this nomenclature, we would say that QED is renormalizable in four dimensions, super-renormalizable in less than four dimensions, and non-renormalizable in more than four dimensions.

These superficial criteria give a correct picture of the true divergence structure of the theory for most cases that have been studied in detail. Examples are known in which the true behavior is better than this picture suggests, when powerful symmetries set to zero some or all of the superficially divergent amplitudes.* On the other hand, as we will explain in Section 10.4, it is always true that the divergences of superficially renormalizable theories can be absorbed into a finite number of Lagrangian parameters. For theories containing fields of spin 1 and higher, loop diagrams can produce additional problems, including violation of unitarity; we will discuss this difficulty in Chapter 16.

As another example of the counting of ultraviolet divergences, consider a pure scalar field theory, in d dimensions, with a ϕ^n interaction term:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{n!}\phi^n. \tag{10.12}$$

Let N be the number of external lines in a diagram, P the number of propagators, and V the number of vertices. The number of loops in a diagram is $L = P - V + 1$. There are n lines meeting at each vertex, so $nV = N + 2P$.

*Some exotic four-dimensional field theories are actually free of divergences; see, for example, the article by P. West in *Shelter Island II*, R. Jackiw, N. N. Khuri, S. Weinberg, and E. Witten, eds. (MIT Press, Cambridge, 1985).

Combining these relations, we find that the superficial degree of divergence of a diagram is

$$\begin{aligned} D &= dL - 2P \\ &= d + \left[n\left(\frac{d-2}{2}\right) - d\right]V - \left(\frac{d-2}{2}\right)N. \end{aligned} \tag{10.13}$$

In four dimensions a ϕ^4 coupling is renormalizable, while higher powers of ϕ are non-renormalizable. In three dimensions a ϕ^6 coupling becomes renormalizable, while ϕ^4 is super-renormalizable. In two spacetime dimensions any coupling of the form ϕ^n is super-renormalizable.

Expression (10.13) can also be derived in a somewhat different way, from dimensional analysis. In any quantum field theory, the action $S = \int d^dx\,\mathcal{L}$ must be dimensionless, since we work in units where $\hbar = 1$. In this system of units, the integral d^dx has units $(\text{mass})^{-d}$, and so the Lagrangian has units $(\text{mass})^d$. Since all units can be expressed as powers of mass, it is unambiguous to say simply that the Lagrangian has "dimension d". Using this result, we can infer from the explicit form of (10.12) the dimensions of the field ϕ and the coupling constant λ. From the kinetic term in $\mathcal{L}$ we see that ϕ has dimension $(d-2)/2$. Note that the parameter m consistently has dimensions of mass. From the interaction term and the dimension of ϕ, we infer that the λ has dimension $d - n(d-2)/2$.

Now consider an arbitrary diagram with N external lines. One way that such a diagram could arise is from an interaction term $\eta\phi^N$ in the Lagrangian. The dimension of η would then be $d - N(d-2)/2$, and therefore we conclude that any (amputated) diagram with N external lines has dimension $d - N(d-2)/2$. In our theory with only the $\lambda\phi^n$ vertex, if the diagram has V vertices, its divergent part is proportional to $\lambda^V\Lambda^D$, where Λ is a high-momentum cutoff and D is the superficial degree of divergence. (This is the "generic" case; all the exceptions noted above also apply here.) Applying dimensional analysis, we find

$$d - N\left(\frac{d-2}{2}\right) = V\left[d - n\left(\frac{d-2}{2}\right)\right] + D,$$

in agreement with (10.13).

Note that the quantity that multiplies V in this expression is just the dimension of the coupling constant λ. This analysis can be carried out for QED and other field theories, with the same result. Thus we can characterize the three degrees of renormalizability in a second way:

Super-Renormalizable:	Coupling constant has positive mass dimension.
Renormalizable:	Coupling constant is dimensionless.
Non-Renormalizable:	Coupling constant has negative mass dimension.

This is exactly the conclusion that we stated without proof in Section 4.1. In QED, the coupling constant e is dimensionless; thus QED is (at least superficially) renormalizable.

10.2 Renormalized Perturbation Theory

In the previous section we saw that a renormalizable quantum field theory contains only a small number of superficially divergent amplitudes. In QED, for example, there are three such amplitudes, containing four infinite constants. In Chapters 6 and 7 these infinities disappeared by the end of our computations: The infinity in the vertex correction diagram was canceled by the electron field-strength renormalization, while the infinity in the vacuum polarization diagram caused only an unobservable shift of the electron's charge. In fact, it is generally true that the divergences in a renormalizable quantum field theory never show up in observable quantities.

To obtain a finite result for an amplitude involving divergent diagrams, we have so far used the following procedure: Compute the diagrams using a regulator, to obtain an expression that depends on the bare mass (m_0), the bare coupling constant (e_0), and some ultraviolet cutoff (Λ). Then compute the physical mass (m) and the physical coupling constant (e), to whatever order is consistent with the rest of the calculation; these quantities will also depend on m_0, e_0, and Λ. To calculate an S-matrix element (rather than a correlation function), one must also compute the field-strength renormalization(s) Z (in accord with Eq. (7.45)). Combining all of these expressions, eliminate m_0 and e_0 in favor of m and e; this step is the "renormalization". The resulting expression for the amplitude should be finite in the limit $\Lambda \to \infty$.

The above procedure always works in a renormalizable quantum field theory. However, it can often be cumbersome, especially at higher orders in perturbation theory. In this section we will develop an alternative procedure which works more automatically. We will do this first for ϕ^4 theory, returning to QED in the next section.

The Lagrangian of ϕ^4 theory is

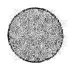
$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4.$$

We now write m_0 and λ_0, to emphasize that these are the bare values of the mass and coupling constant, not the values measured in experiments.

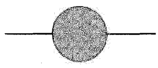
The superficial degree of divergence of a diagram with N external legs is, according to (10.13),

$$D = 4 - N.$$

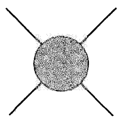
Since the theory is invariant under $\phi \to -\phi$, all amplitudes with an odd number of external legs vanish. The only divergent amplitudes are therefore



(unobservable vacuum energy shift);



$\sim\Lambda^2 + p^2 \log \Lambda +$ (finite terms);



$\sim \log \Lambda +$ (finite terms).

Ignoring the vacuum diagram, these amplitudes contain three infinite constants. Our goal is to absorb these constants into the three unobservable parameters of the theory: the bare mass, the bare coupling constant, and the field strength. To accomplish this goal, it is convenient to reformulate the perturbation expansion so that these unobservable quantities do not appear explicitly in the Feynman rules.

First we will eliminate the shift in the field strength. Recall from Section 7.1 that the exact two-point function has the form

$$\int d^4x \, \langle\Omega| T\phi(x)\phi(0) |\Omega\rangle \, e^{ip\cdot x} = \frac{iZ}{p^2 - m^2} + (\text{terms regular at } p^2 = m^2), \tag{10.14}$$

where m is the physical mass. We can eliminate the awkward residue Z from this equation by rescaling the field:

$$\phi = Z^{1/2}\phi_r. \tag{10.15}$$

This transformation changes the values of correlation functions by a factor of $Z^{-1/2}$ for each field. Thus, in computing S-matrix elements, we no longer need the factors of Z in Eq. (7.45); a scattering amplitude is simply the sum of all connected, amputated diagrams, exactly as we originally guessed in Eq. (4.103).

The Lagrangian is much uglier after the rescaling:

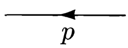
$$\mathcal{L} = \frac{1}{2}Z(\partial_\mu\phi_r)^2 - \frac{1}{2}m_0^2 Z\phi_r^2 - \frac{\lambda_0}{4!}Z^2\phi_r^4. \tag{10.16}$$

The bare mass and coupling constant still appear in $\mathcal{L}$, but they can be eliminated as follows. Define

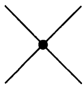
$$\delta_Z = Z - 1, \qquad \delta_m = m_0^2 Z - m^2, \qquad \delta_\lambda = \lambda_0 Z^2 - \lambda, \tag{10.17}$$

where m and λ are the physically measured mass and coupling constant. Then the Lagrangian becomes

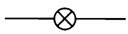
$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial_\mu\phi_r)^2 - \frac{1}{2}m^2\phi_r^2 - \frac{\lambda}{4!}\phi_r^4 \\ &\quad + \frac{1}{2}\delta_Z(\partial_\mu\phi_r)^2 - \frac{1}{2}\delta_m\phi_r^2 - \frac{\delta_\lambda}{4!}\phi_r^4.\end{aligned} \tag{10.18}$$



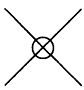
$$= \frac{i}{p^2 - m^2 + i\epsilon}$$



$$= -i\lambda$$



$$= i(p^2\delta_Z - \delta_m)$$

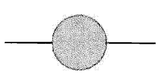


$$= -i\delta_\lambda$$

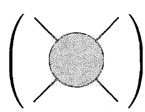
Figure 10.3. Feynman rules for ϕ^4 theory in renormalized perturbation theory.

The first line now looks like the familiar ϕ^4-theory Lagrangian, but is written in terms of the physical mass and coupling. The terms in the second line, known as *counterterms*, have absorbed the infinite but unobservable shifts between the bare parameters and the physical parameters. It is tempting to say that we have "added" these counterterms to the Lagrangian, but in fact we have merely split each term in (10.16) into two pieces.

The definitions in (10.17) are not useful unless we give precise definitions of the physical mass and coupling constant. Equation (10.14) defines m^2 as the location of the pole in the propagator. There is no obviously best definition of λ, but a perfectly good definition would be obtained by setting λ equal to the magnitude of the scattering amplitude at zero momentum. Thus we have the two defining relations,



$$= \frac{i}{p^2 - m^2} + (\text{terms regular at } p^2 = m^2);$$



amputated

$$= -i\lambda \qquad \text{at } s = 4m^2,\ t = u = 0. \tag{10.19}$$

These equations are called *renormalization conditions.* (The first equation actually contains two conditions, specifying both the location of the pole and its residue.)

Our new Lagrangian, Eq. (10.18), gives a new set of Feynman rules, shown in Fig. 10.3. The propagator and the first vertex come from the first line of (10.18), and are identical to the old rules except for the appearance of the physical mass and coupling in place of the bare values. The counterterms in the second line of (10.18) give two new vertices (also called counterterms).

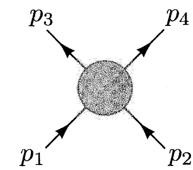
We can use these new Feynman rules to compute any amplitude in ϕ^4 theory. The procedure is as follows. Compute the desired amplitude as the sum of all possible diagrams created from the propagator and vertices shown in Fig. 10.3. The loop integrals in the diagrams will often diverge, so one must introduce a regulator. The result of this computation will be a function of the three unknown parameters δ_Z, δ_m, and δ_λ. Adjust (or "renormalize") these three parameters as necessary to maintain the renormalization conditions (10.19). After this adjustment, the expression for the amplitude should be finite and independent of the regulator.

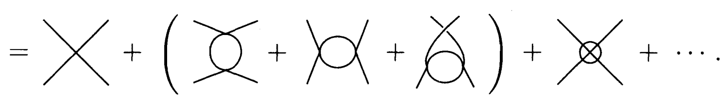
This procedure, using Feynman rules with counterterms, is known as *renormalized perturbation theory.* It should be contrasted with the procedure we used in Part 1, outlined at the beginning of this section, which is called *bare perturbation theory* (since the Feynman rules involve the bare mass and coupling constant). The two methods are completely equivalent. The differences between them are purely a matter of bookkeeping. You will get the same answers using either procedure, so you may choose whichever you find more convenient. In general, renormalized perturbation theory is technically easier to use, especially for multiloop diagrams; however, bare perturbation theory is sometimes easier for complicated one-loop calculations. We will use renormalized perturbation theory in most of the rest of this book.

One-Loop Structure of ϕ^4 Theory

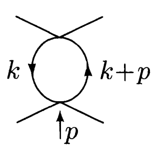
To make more sense of the renormalization procedure, let us carry it out explicitly at the one-loop level.

First consider the basic two-particle scattering amplitude,

$i\mathcal{M}(p_1p_2 \to p_3p_4) =$ 



If we define $p = p_1 + p_2$, then the second diagram is



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

$$\equiv (-i\lambda)^2 \cdot iV(p^2). \tag{10.20}$$

Note that p^2 is equal to the Mandelstam variable s. The next two diagrams are identical, except that s will be replaced by t and u. The entire amplitude is therefore

$$i\mathcal{M} = -i\lambda + (-i\lambda)^2\big[iV(s) + iV(t) + iV(u)\big] - i\delta_\lambda. \tag{10.21}$$

According to our renormalization condition (10.19), this amplitude should equal $-i\lambda$ at $s = 4m^2$ and $t = u = 0$. We must therefore set

$$\delta_\lambda = -\lambda^2\big[V(4m^2) + 2V(0)\big]. \tag{10.22}$$

(At higher orders, δ_λ will receive additional contributions.)

We can compute $V(p^2)$ explicitly using dimensional regularization. The procedure is exactly the same as in Section 7.5: Introduce a Feynman parameter, shift the integration variable, rotate to Euclidean space, and perform the momentum integral. We obtain

$$\begin{aligned}
V(p^2) &= \frac{i}{2}\int_0^1 dx \int \frac{d^dk}{(2\pi)^d}\frac{1}{\big[k^2 + 2xk\cdot p + xp^2 - m^2\big]^2} \\
&= \frac{i}{2}\int_0^1 dx \int \frac{d^d\ell}{(2\pi)^d}\frac{1}{\big[\ell^2 + x(1-x)p^2 - m^2\big]^2} \qquad (\ell = k + xp) \\
&= -\frac{1}{2}\int_0^1 dx \int \frac{d^d\ell_E}{(2\pi)^d}\frac{1}{\big[\ell_E^2 - x(1-x)p^2 + m^2\big]^2} \qquad (\ell_E^0 = -i\ell^0) \\
&= -\frac{1}{2}\int_0^1 dx\, \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}}\frac{1}{\big[m^2 - x(1-x)p^2\big]^{2-d/2}} \\
&\underset{d\to 4}{\longrightarrow} -\frac{1}{32\pi^2}\int_0^1 dx\Big(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log\big[m^2 - x(1-x)p^2\big]\Big),
\end{aligned} \tag{10.23}$$

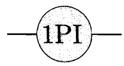
where $\epsilon = 4 - d$. The shift in the coupling constant (10.22) is therefore

$$\begin{aligned}
\delta_\lambda &= \frac{\lambda^2}{2}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}}\int_0^1 dx\Big(\frac{1}{[m^2 - x(1-x)4m^2]^{2-d/2}} + \frac{2}{[m^2]^{2-d/2}}\Big) \\
&\underset{d\to 4}{\longrightarrow} \frac{\lambda^2}{32\pi^2}\int_0^1 dx\Big(\frac{6}{\epsilon} - 3\gamma + 3\log(4\pi) - \log\big[m^2 - x(1-x)4m^2\big] - 2\log\big[m^2\big]\Big).
\end{aligned} \tag{10.24}$$

These expressions are divergent as $d \to 4$. But if we combine them according to (10.21), we obtain the finite (if rather complicated) result,

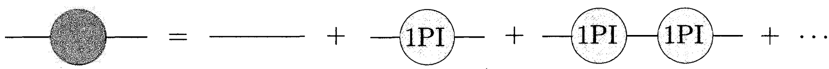
$$\begin{aligned}
i\mathcal{M} = -i\lambda - \frac{i\lambda^2}{32\pi^2}\int_0^1 dx\Big[&\log\Big(\frac{m^2 - x(1-x)s}{m^2 - x(1-x)4m^2}\Big) + \log\Big(\frac{m^2 - x(1-x)t}{m^2}\Big) \\
&+ \log\Big(\frac{m^2 - x(1-x)u}{m^2}\Big)\Big].
\end{aligned} \tag{10.25}$$

To determine δ_Z and δ_m we must compute the two-point function. As in Section 7.2, let us define $-iM^2(p^2)$ as the sum of all one-particle-irreducible insertions into the propagator:



$$= -iM^2(p^2). \tag{10.26}$$

Then the full two-point function is given by the geometric series,



$$= \frac{i}{p^2 - m^2 - M^2(p^2)}. \tag{10.27}$$

The renormalization conditions (10.19) require that the pole in this full propagator occur at $p^2 = m^2$ and have residue 1. These two conditions are equivalent, respectively, to

$$M^2(p^2)\big|_{p^2=m^2} = 0 \qquad \text{and} \qquad \frac{d}{dp^2}M^2(p^2)\big|_{p^2=m^2} = 0. \tag{10.28}$$

(To check the latter condition, expand M^2 about $p^2 = m^2$ in Eq. (10.27).)

Explicitly, to one-loop order,

$$-iM^2(p^2) =$$

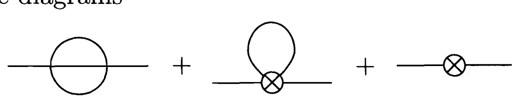
$$\begin{aligned} &= -i\lambda \cdot \frac{1}{2} \cdot \int \frac{d^dk}{(2\pi)^d} \frac{i}{k^2 - m^2} + i(p^2\delta_Z - \delta_m) \\ &= -\frac{i\lambda}{2} \frac{1}{(4\pi)^{d/2}} \frac{\Gamma(1-\frac{d}{2})}{(m^2)^{1-d/2}} + i(p^2\delta_Z - \delta_m). \end{aligned} \tag{10.29}$$

Since the first term is independent of p^2, the result is rather trivial: Setting

$$\delta_Z = 0 \qquad \text{and} \qquad \delta_m = -\frac{\lambda}{2(4\pi)^{d/2}} \frac{\Gamma(1-\frac{d}{2})}{(m^2)^{1-d/2}} \tag{10.30}$$

yields $M^2(p^2) = 0$ for all p^2, satisfying both of the conditions in (10.28).

The first nonzero contributions to $M^2(p^2)$ and δ_Z are proportional to λ^2, coming from the diagrams



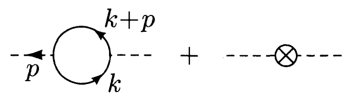
(10.31)

The second diagram contains the δ_λ counterterm, which we have already computed. It cancels ultraviolet divergences in the first diagram that occur when one of the loop momenta is large and the other is small. The third diagram is again the $(p^2\delta_Z - \delta_m)$ counterterm, and is fixed to order λ^2 by requiring that the remaining divergences (when both loop momenta become large) cancel. In Section 10.4 we will see an explicit example of the interplay of various counterterms in a two-loop calculation.

The vanishing of δ_Z at one-loop order is a special feature of ϕ^4 theory, which does not occur in more general theories of scalar fields. The Yukawa theory described in Section 4.7 gives an explicit example of a one-loop correction for which this counterterm is required.

In the Yukawa theory, the scalar field propagator receives corrections at order g^2 from a fermion loop diagram and the two propagator counterterms. Using the Feynman rules on p. 118 to compute the loop diagram, we find

$$-iM^2(p^2) =$$



$$\begin{aligned}
&= -(-ig)^2 \int \frac{d^dk}{(2\pi)^d} \operatorname{tr}\left[\frac{i(\not{k}+\not{p}+m_f)}{(k+p)^2 - m_f^2}\frac{i(\not{k}+m_f)}{k^2 - m_f^2}\right] + i(p^2\delta_Z - \delta_m) \\
&= -4g^2 \int \frac{d^dk}{(2\pi)^d} \frac{k\cdot(p+k) + m_f^2}{((p+k)^2 - m_f^2)(k^2 - m_f^2)} + i(p^2\delta_Z - \delta_m), \qquad (10.32)
\end{aligned}$$

where m_f is the mass of the fermion that couples to the Yukawa field. To evaluate the integral, combine denominators and shift as in Eq. (10.23). Then the first term in the last line becomes

$$\begin{aligned}
&-4g^2 \int_0^1 dx \int \frac{d^d\ell}{(2\pi)^d} \frac{\ell^2 - x(1-x)p^2 + m_f^2}{(\ell^2 + x(1-x)p^2 - m_f^2)^2} \\
&\qquad = -4g^2 \int_0^1 dx \frac{-i}{(4\pi)^{d/2}}\left(\frac{\frac{d}{2}\Gamma(1-\frac{d}{2})}{\Delta^{1-d/2}} - \frac{\Delta\,\Gamma(2-\frac{d}{2})}{\Delta^{2-d/2}}\right) \\
&\qquad = \frac{4ig^2(d-1)}{(4\pi)^{d/2}} \int_0^1 dx \frac{\Gamma(1-\frac{d}{2})}{\Delta^{1-d/2}}, \qquad (10.33)
\end{aligned}$$

where $\Delta = m_f^2 - x(1-x)p^2$.

Now we can see that both of the counterterms δ_m and δ_Z must take nonzero values in order to satisfy the renormalization conditions (10.28). To determine δ_m, we subtract the value of the loop diagram at $p^2 = m^2$ as before, so that

$$\delta_m = \frac{4g^2(d-1)}{(4\pi)^{d/2}} \int_0^1 dx \frac{\Gamma(1-\frac{d}{2})}{[m_f^2 - x(1-x)m^2]^{1-d/2}} + m^2\delta_Z. \qquad (10.34)$$

To determine δ_Z, we cancel also the first derivative with respect to p^2 of the loop integral (10.33). This gives

$$\begin{aligned}\delta_Z &= -\frac{4g^2(d-1)}{(4\pi)^{d/2}} \int_0^1 dx\, \frac{x(1-x)\Gamma(2-\frac{d}{2})}{[m_f^2 - x(1-x)m^2]^{2-d/2}} \\ &\underset{d\to 4}{\longrightarrow} -\frac{3g^2}{4\pi^2} \int_0^1 dx\, x(1-x)\left(\frac{2}{\epsilon} - \gamma - \frac{2}{3} + \log(4\pi) - \log\left[m_f^2 - x(1-x)m^2\right]\right).\end{aligned} \tag{10.35}$$

Thus, in Yukawa theory, the propagator corrections at one-loop order require a quadratically divergent mass renormalization and a logarithmically divergent field strength renormalization. This is the usual situation in scalar field theories.

10.3 Renormalization of Quantum Electrodynamics

The procedure we followed in the previous section, yielding a "renormalized" perturbation theory formulated in terms of physically measurable parameters, can be summarized as follows:

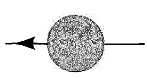
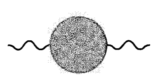
1. Absorb the field-strength renormalizations into the Lagrangian by rescaling the fields.
2. Split each term of the Lagrangian into two pieces, absorbing the infinite and unobservable shifts into counterterms.
3. Specify the renormalization conditions, which define the physical masses and coupling constants and keep the field-strength renormalizations equal to 1.
4. Compute amplitudes with the new Feynman rules, adjusting the counterterms as necessary to maintain the renormalization conditions.

Let us now use this procedure to construct a renormalized perturbation theory for Quantum Electrodynamics.

The original QED Lagrangian is

$$\mathcal{L} = -\tfrac{1}{4}(F_{\mu\nu})^2 + \bar{\psi}(i\partial\!\!\!/ - m_0)\psi - e_0\bar{\psi}\gamma^\mu\psi A_\mu.$$

Computing the electron and photon propagators with this Lagrangian, we would find expressions of the general form

 $= \dfrac{iZ_2}{p\!\!\!/ - m} + \cdots;$  $= \dfrac{-iZ_3 g_{\mu\nu}}{q^2} + \cdots.$

(We found just such expressions in the explicit one-loop calculations of Chapter 7.) To absorb Z_2 and Z_3 into $\mathcal{L}$, and hence eliminate them from formula (7.45) for the S-matrix, we substitute $\psi = Z_2^{1/2}\psi_r$ and $A^\mu = Z_3^{1/2}A_r^\mu$. Then the Lagrangian becomes

$$\mathcal{L} = -\tfrac{1}{4}Z_3(F_r^{\mu\nu})^2 + Z_2\bar{\psi}_r(i\partial\!\!\!/ - m_0)\psi_r - e_0 Z_2 Z_3^{1/2}\bar{\psi}_r\gamma^\mu\psi_r A_{r\mu}. \tag{10.36}$$

We can introduce the physical electric charge e, measured at large distances ($q = 0$), by defining a scaling factor Z_1 as follows:†

$$e_0 Z_2 Z_3^{1/2} = e Z_1. \tag{10.37}$$

If we let m be the physical mass (the location of the pole in the electron propagator), then we can split each term of the Lagrangian into two pieces as follows:

$$\begin{aligned}\mathcal{L} = &-\tfrac{1}{4}(F_r^{\mu\nu})^2 + \bar{\psi}_r(i\partial\!\!\!/ - m)\psi_r - e\bar{\psi}_r\gamma^\mu\psi_r A_{r\mu} \\ &- \tfrac{1}{4}\delta_3(F_r^{\mu\nu})^2 + \bar{\psi}_r(i\delta_2\partial\!\!\!/ - \delta_m)\psi_r - e\delta_1\bar{\psi}_r\gamma^\mu\psi_r A_{r\mu},\end{aligned} \tag{10.38}$$

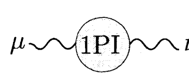
where

$$\delta_3 = Z_3 - 1, \qquad \delta_2 = Z_2 - 1,$$

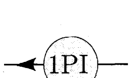
$$\delta_m = Z_2 m_0 - m, \qquad \text{and} \qquad \delta_1 = Z_1 - 1 = (e_0/e)Z_2 Z_3^{1/2} - 1.$$

The Feynman rules for renormalized QED are shown in Fig. 10.4. In addition to the familiar propagators and vertex, there are three counterterm vertices. The ee and $ee\gamma$ counterterm vertices can be read directly from the Lagrangian (10.38). To derive the two-photon counterterm, integrate $-\frac{1}{4}(F_{\mu\nu})^2$ by parts to obtain $-\frac{1}{2}A_\mu(-\partial^2 g^{\mu\nu} + \partial^\mu\partial^\nu)A_\nu$; this gives the expression shown in the figure. In the remainder of the book, when we set up renormalized perturbation theory, we will drop the subscript r used here to distinguish the rescaled fields.

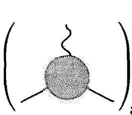
Each of the four counterterm coefficients must be fixed by a renormalization condition. The four conditions that we require have already been stated implicitly: Two of them fix the electron and photon field-strength renormalizations to 1, while the other two define the physical electron mass and charge. To write these conditions more explicitly, recall our notation from Chapters 6 and 7:



$$= i\Pi^{\mu\nu}(q) = i(g^{\mu\nu}q^2 - q^\mu q^\nu)\Pi(q^2),$$



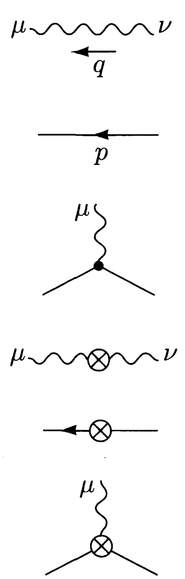
$$= -i\Sigma(p\!\!\!/), \tag{10.39}$$



amputated

$$= -ie\Gamma^\mu(p', p).$$

†Since we define e by the renormalization condition $\Gamma^\mu(q = 0) = \gamma^\mu$, the factor of Z_1 in the Lagrangian must cancel the multiplicative correction factor that arises from loop corrections. Therefore this definition of Z_1 is equivalent to that given in Eq. (7.47).



$$= \frac{-ig_{\mu\nu}}{q^2 + i\epsilon} \qquad \text{(Feynman gauge)}$$

$$= \frac{i}{\not{p} - m + i\epsilon}$$

$$= -ie\gamma^\mu$$

$$= -i(g^{\mu\nu}q^2 - q^\mu q^\nu)\delta_3$$

$$= i(\not{p}\delta_2 - \delta_m)$$

$$= -ie\gamma^\mu\delta_1$$

Figure 10.4. Feynman rules for Quantum Electrodynamics in renormalized perturbation theory.

These amplitudes are now to be computed in renormalized perturbation theory; that is, we are now redefining $\Pi(q^2)$, $\Sigma(\not{p})$, and $\Gamma(p', p)$ to include counterterm vertices. Furthermore, the new definition of Γ involves the physical electron charge. With this notation, the four conditions are

$$\begin{aligned}
\Sigma(\not{p} = m) &= 0; \\
\frac{d}{d\not{p}}\Sigma(\not{p})\Big|_{\not{p}=m} &= 0; \\
\Pi(q^2 = 0) &= 0; \\
-ie\Gamma^\mu(p' - p = 0) &= -ie\gamma^\mu.
\end{aligned} \tag{10.40}$$

The first condition fixes the electron mass at m, while the next two fix the residues of the electron and photon propagators at 1. Given these conditions, the final condition fixes the electron charge to be e.

One-Loop Structure of QED

The four conditions (10.40) allow us to determine the four countertems in (10.38) in terms of the values of loop diagrams. In Chapters 6 and 7 we computed all of the diagrams required to carry out this determination to one-loop order. We will now collect these results and find explicit expressions for the renormalization constants of QED to order α. For overall consistency, we will use dimensional regularization to control ultraviolet divergences, and a photon mass μ to control infrared divergences. In Part I, we computed the vertex and self-energy diagrams using the Pauli-Villars regularization scheme, before introducing dimensional regularization. Now we have an opportunity to quote the values of these diagrams as computed with dimensional regularization.

The first two conditions involve the electron self-energy. We evaluated the one-loop diagram contributing to $\Sigma(p)$, using a Pauli-Villars regulator, in Section 7.1; the result is given in Eq. (7.19). If we re-evaluate the diagram in dimensional regularization, we find some additional terms in the Dirac algebra from the modified contraction identities (7.89). Taking these terms into account, we find for this diagram ($\epsilon = 4 - d$)

$$-i\Sigma_2(p) = -i\frac{e^2}{(4\pi)^{d/2}}\int_0^1 dx\,\frac{\Gamma(2-\frac{d}{2})}{((1-x)m^2 + x\mu^2 - x(1-x)p^2)^{2-d/2}} \times ((4-\epsilon)m - (2-\epsilon)x\not{p}). \tag{10.41}$$

Therefore, according to the first of conditions (10.40),

$$m\delta_2 - \delta_m = \Sigma_2(m) = \frac{e^2 m}{(4\pi)^{d/2}}\int_0^1 dx\,\frac{\Gamma(2-\frac{d}{2})\cdot(4-2x-\epsilon(1-x))}{((1-x)^2m^2 + x\mu^2)^{2-d/2}}. \tag{10.42}$$

Similarly, the second of conditions (10.40) determines δ_2:

$$\begin{aligned}\delta_2 &= \frac{d}{d\not{p}}\Sigma_2(m)\\ &= -\frac{e^2}{(4\pi)^{d/2}}\int_0^1 dx\,\frac{\Gamma(2-\frac{d}{2})}{((1-x)^2m^2 + x\mu^2)^{2-d/2}}\\ &\qquad\times\left[(2-\epsilon)x - \frac{\epsilon}{2}\frac{2x(1-x)m^2}{(1-x)^2m^2 + x\mu^2}(4-2x-\epsilon(1-x))\right].\end{aligned} \tag{10.43}$$

Notice that the second term in the brackets gives a finite result as $\epsilon \to 0$, because it multiplies the divergent gamma function.

The third condition of (10.40) requires the value (7.90) of the photon self-energy diagram:

$$\Pi_2(q^2) = -\frac{e^2}{(4\pi)^{d/2}}\int_0^1 dx\,\frac{\Gamma(2-\frac{d}{2})}{(m^2 - x(1-x)q^2)^{2-d/2}}\big(8x(1-x)\big).$$

Then

$$\delta_3 = \Pi_2(0) = -\frac{e^2}{(4\pi)^{d/2}}\int_0^1 dx\,\frac{\Gamma(2-\frac{d}{2})}{(m^2)^{2-d/2}}\big(8x(1-x)\big). \tag{10.44}$$

The last condition requires the value of the electron vertex function, computed in Section 6.3. Again, we will rework the diagram in dimensional regularization. Then the shift in the form factor $F_1(q^2)$ (6.56) becomes

$$\delta F_1(q^2) = \frac{e^2}{(4\pi)^{d/2}} \int dx\, dy\, dz\, \delta(x+y+z-1) \left[\frac{\Gamma(2-\frac{d}{2})}{\Delta^{2-d/2}} \frac{(2-\epsilon)^2}{2} + \frac{\Gamma(3-\frac{d}{2})}{\Delta^{3-d/2}} \left(q^2[2(1-x)(1-y) - \epsilon xy] + m^2[2(1-4z+z^2) - \epsilon(1-z)^2] \right) \right], \tag{10.45}$$

where $\Delta = (1-z)^2 m^2 + z\mu^2 - xyq^2$ as before. The fourth renormalization condition then determines

$$\delta_1 = -\delta F_1(0) = -\frac{e^2}{(4\pi)^{d/2}} \int dz\, (1-z) \left[\frac{\Gamma(2-\frac{d}{2})}{((1-z)^2 m^2 + z\mu^2)^{2-d/2}} \frac{(2-\epsilon)^2}{2} + \frac{\Gamma(3-\frac{d}{2})}{((1-z)^2 m^2 + z\mu^2)^{3-d/2}} [2(1-4z+z^2) - \epsilon(1-z)^2] m^2 \right]. \tag{10.46}$$

Using an integration by parts similar to that following Eq. (7.32), one can show explicitly from (10.46) and (10.43) that $\delta_1 = \delta_2$, that is, that $Z_1 = Z_2$ to order α. As in our previous derivations, this formula follows from the Ward identity. The Lagrangian (10.38), with counterterms set to zero, is gauge invariant. If the regulator is also gauge invariant (and we do use dimensional regularization), this implies the Ward identity for diagrams without counterterm vertices. In particular, this implies that $\delta F_1(0) = -d\Sigma_2/d\not{p}|_m$. Then the counterterms δ_1 and δ_2, which are required to cancel these two factors, will be set equal.

By continuing this argument, it is straightforward to construct a full diagrammatic proof that $\delta_1 = \delta_2$, to all orders in renormalized perturbation theory, using the method we applied in Section 7.4 to prove the Ward-Takahashi identity in bare perturbation theory. With a generalization of the argument given there, one can show that the diagrammatic identity (7.68) holds for diagrams that include counterterm vertices in loops. Thus, if the counterterms δ_1 and δ_2 are determined up to order α^n, the unrenormalized vertex diagram at $q^2 = 0$ equals the derivative of the unrenormalized self-energy diagram on-shell in order α^{n+1}. To satisfy the renormalization conditions (10.40), we must then set the counterterms δ_1 and δ_2 equal to order α^{n+1}. This recursive argument gives yet another proof that $Z_1 = Z_2$ to all orders in QED perturbation theory.

The relation (10.37) between the bare and renormalized charge

$$e = \frac{Z_2}{Z_1} Z_3^{1/2} e_0 \tag{10.47}$$

gives a further physical interpretation of the identity $Z_1 = Z_2$. Using the identity, we can rewrite (10.47) as

$$e = \sqrt{Z_3} e_0,$$

which is just the relation (7.76) that we derived by a diagrammatic argument in Section 7.5. This says that the relation between the bare and renormalized electric charge depends only on the photon field strength renormalization, not on quantities particular to the electron. To see the importance of this observation, consider writing the renormalized quantum electrodynamics with two species of charged particles, say, electrons and muons. Then, in addition to (10.37), we will have a relation for the photon-muon vertex:

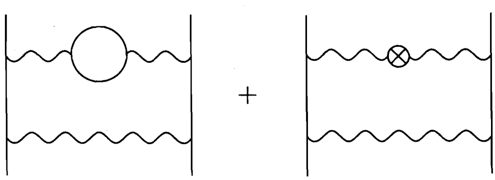
$$e Z_2'^{-1} Z_3^{-1/2} = e_0 Z_1'^{-1}, \tag{10.48}$$

where Z_1' and Z_2' are the vertex and field strength renormalizations for the muon. Each of these two constants depends on the mass of the muon, so (10.48) threatens to give a different relation between e_0 and e from the one written in (10.47). However, the Ward identity forces the factors Z_1' and Z_2' to cancel out of this relation, leaving over a universal electric charge which has the same value for all species.

10.4 Renormalization Beyond the Leading Order

In the last two sections we have developed an algorithm for computing scattering amplitudes to any order in a renormalizable field theory. We have seen explicitly that this algorithm yields finite results at the one-loop level in both ϕ^4 theory and QED. According to the naive analysis of Section 10.1, the algorithm should also work at higher orders. But that analysis ignored many of the intricacies of multiloop diagrams; specifically, it ignored the fact that diagrams can contain divergent subdiagrams.

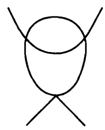
When an otherwise finite diagram contains a divergent subdiagram, the treatment of the divergence is relatively straightforward. For example, the sum of diagrams

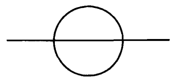


(10.49)

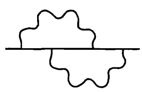
is finite: The divergence in the photon propagator cancels just as when this propagator occurs in a tree diagram. The finite sum of the two propagator diagrams gives an integrand for the outer loop that falls off fast enough that this integral still converges.

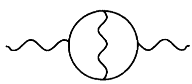
A more difficult situation occurs when we have *nested* or *overlapping* divergences, that is, when two divergent loops share a propagator. Some examples of diagrams with overlapping divergences are





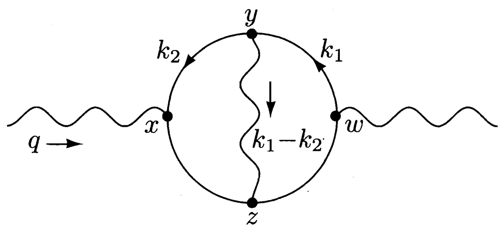
in ϕ^4 theory;



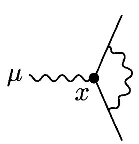


in QED.

To see the difficulty, consider the photon self-energy diagram:

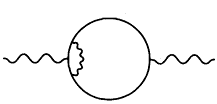


One contribution to this diagram comes from the region of momentum space where k_2 is very large. This means that, in position space, x, y, and z are very close together, while w can be farther away. In this region we can think of the virtual photon as giving a correction to the vertex at x. We saw in Section 6.3 that this vertex correction is logarithmically divergent, of the form



$$\sim -ie\gamma^\mu \cdot \alpha \log \Lambda^2$$

in the limit $\Lambda \to \infty$. Plugging this vertex into the rest of the diagram and integrating over k_1, we obtain an expression identical to the one-loop photon self-energy correction $\Pi_2(q^2)$, displayed in (7.90), multiplied by the additional logarithmic divergence:

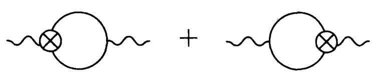


$$\begin{aligned} &\sim \alpha(g^{\mu\nu}q^2 - q^\mu q^\nu)\Pi_2(q^2) \cdot \alpha \log \Lambda^2 \\ &\sim \alpha(g^{\mu\nu}q^2 - q^\mu q^\nu)(\log \Lambda^2 + \log q^2) \cdot \alpha \log \Lambda^2 . \end{aligned} \tag{10.50}$$

The $\log^2 \Lambda^2$ term comes from the region where both k_1 and k_2 are large, while the $\log q^2 \log \Lambda^2$ term comes from the region where k_2 is large but k_1 is small. Another such term would come from the region where k_1 is large but k_2 is small.

The appearance of terms proportional to $\Pi_2(q^2) \cdot \log \Lambda^2$ in the two-loop vacuum polarization diagram contradicts our naive argument, based on the criterion of the superficial degree of divergence, that the divergent terms of a Feynman integral are always simple polynomials in q^2. We will refer to divergences multiplying only polynomials in q^2 as *local divergences*, since their Fourier transforms back to position space are delta functions or derivatives of delta functions. We will call the new, nonpolynomial, term a *nonlocal* divergence. Fortunately, our derivation of the nonlocal divergent term gave this term a physical interpretation: It is a local divergence surrounded by an ordinary, nondivergent, quantum field theory process.

If this picture accurately describes all of the divergent terms of the two-loop diagram, we should expect that these divergences are canceled by two types of counterterm diagrams. First, we can build diagrams of order α^2 by inserting the order-α counterterm vertex into the one-loop vacuum polarization diagram:



These diagrams should cancel the nonlocal divergence in (10.50) and the corresponding contribution from the region where k_1 is large and k_2 is small. In fact, a detailed analysis shows that the sum of the original diagram and these two counterterm diagrams contains only local divergences. Once these diagrams are added, the only divergence that remains is a local one, which can be canceled by the diagram

that is, by adding an order-α^2 term to δ_3.

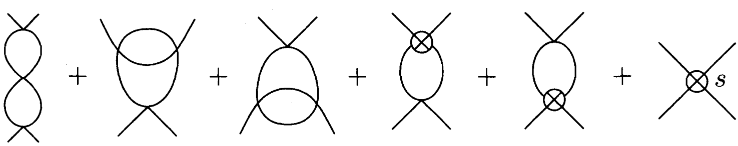
We can extend the lessons of this example to a general picture of the divergences of higher-loop Feynman diagrams and their cancellation. A given diagram may contain local divergences, as predicted by the analysis of Section 10.1. It may also contain nonlocal divergences due to divergent subgraphs embedded in loops carrying small momenta. These divergences are canceled by diagrams in which the divergent subgraphs are replaced by their counterterm vertices. One might still ask two questions: First, does this procedure remove all nonlocal divergences? Second, does this procedure preserve the finiteness of amplitudes, such as (10.49), that are not expected to be divergent by the superficial criteria of Section 10.1? To answer these questions requires an intricate study of nested Feynman integrals. The general analysis was begun by Bogoliubov and Parasiuk, completed by Hepp, and elegantly refined by

Zimmermann;[‡] they showed that the answer to both questions is yes. Their result, known as the BPHZ theorem, states that, for a general renormalizable quantum field theory, to any order in perturbation theory, all divergences are removed by the counterterm vertices corresponding to superficially divergent amplitudes. In other words, any superficially renormalizable quantum field theory is in fact rendered finite when one performs renormalized perturbation theory with the complete set of counterterms.

The proof of the BPHZ theorem is quite technical, and we will not include it in this book. Instead, we will investigate one detailed example of a two-loop calculation, which demonstrates explicitly the appearance and cancellation of nonlocal divergences.

10.5 A Two-Loop Example

To illustrate the issues discussed in the previous section, let us consider the two-loop contribution to the four-point function in ϕ^4 theory. There are 16 relevant diagrams, shown in Fig. 10.5. (There are also several diagrams involving the one-loop correction to the propagator. But each of these is exactly canceled by its counterterm, as we saw in Eq. (10.29), so we can just ignore them.) Fortunately, many of the diagrams are simply related to each other. Crossing symmetry reduces the number of distinct diagrams to only six,

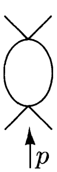


(10.51)

where the last diagram denotes only the s-channel piece of the second-order vertex counterterm. If this sum of diagrams is finite, then simply replacing s with t or u gives a finite result for the remaining diagrams.

The value of the last diagram in (10.51) is just a constant, which we can freely adjust to absorb any divergent terms that are independent of the external momenta. Our goal, therefore, is to show that all momentum-dependent divergent terms cancel among the remaining five diagrams.

The fourth and fifth diagrams in (10.51) involve the one-loop vertex counterterm, which we computed in Eq. (10.24). Let us briefly recall that computation. We defined $iV(p^2)$ as the fundamental loop integral,



$$= (-i\lambda)^2 \cdot iV(p^2) = (-i\lambda)^2\left[-\frac{i}{2}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}}\int_0^1 dx\frac{1}{\left[m^2 - x(1-x)p^2\right]^{2-d/2}}\right]. \tag{10.52}$$

[‡]N. N. Bogoliubov and O. S. Parasiuk, *Acta Math.* **97**, 227 (1957); K. Hepp, *Comm. Math. Phys.* **2**, 301 (1966); W. Zimmermann, in Deser, et. al. (1970).

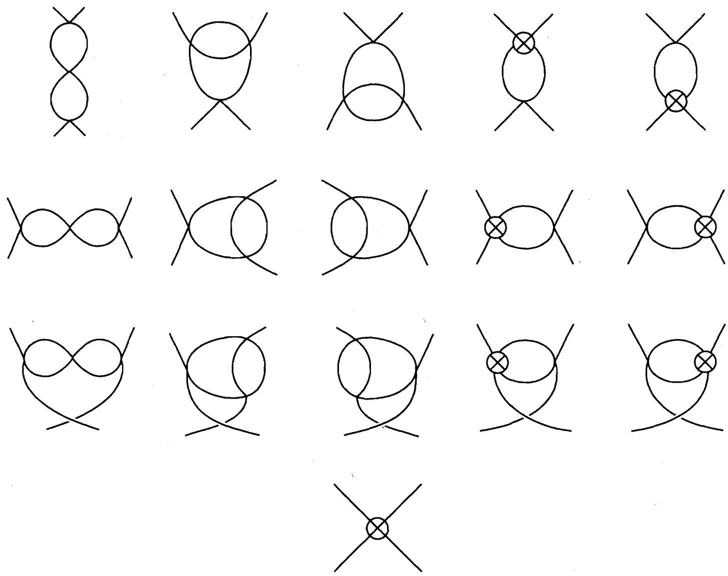
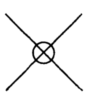


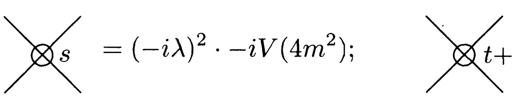
Figure 10.5. The two-loop contributions to the four-point function in ϕ^4 theory. Note that the diagrams in the first three lines are related to each other by crossing, being in the s-, t-, and u-channels, respectively. The last two diagrams in each of these lines involve the $\mathcal{O}(\lambda^2)$ vertex counterterm, while the final diagram is the $\mathcal{O}(\lambda^3)$ contribution to the vertex counterterm.

The counterterm, according to the renormalization condition (10.19), had to cancel the three one-loop diagrams (one for each channel) at threshold ($s = 4m^2$, $t = u = 0$); thus we found



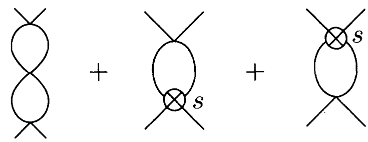
$$= -i\delta_\lambda = (-i\lambda)^2\big[-iV(4m^2) - 2iV(0)\big].$$

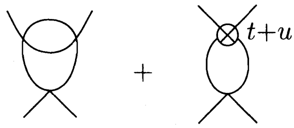
For our present purposes it will be convenient to separate the two terms of this expression. Let us therefore define

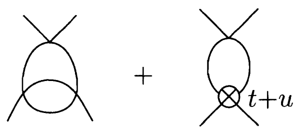


$$s \quad = (-i\lambda)^2 \cdot -iV(4m^2); \qquad t{+}u \quad = (-i\lambda)^2 \cdot -2iV(0).$$

We can now divide the first five diagrams in (10.51) into three groups, as follows:

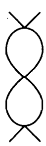
Group I: 

Group II: 

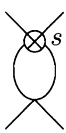
Group III: 

We will find that all divergent terms that depend on momentum cancel separately within each group. Since Groups II and III are related by a simple interchange of initial and final momenta, it suffices to demonstrate this cancellation for Groups I and II.

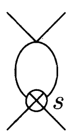
Group I is actually quite easy, since each diagram factors into a product of objects we have already computed. Referring to Eq. (10.52), we have



$$= (-i\lambda)^3 \cdot \left[iV(p^2)\right]^2;$$



$$=$$



$$= (-i\lambda)^3 \cdot iV(p^2) \cdot -iV(4m^2).$$

The sum of all three diagrams is therefore

$$\begin{aligned} &(-i\lambda)^3\Big(\left[iV(p^2)\right]^2 - 2iV(p^2)iV(4m^2)\Big) \\ &\qquad = (-i\lambda)^3\Big(-\left[V(p^2) - V(4m^2)\right]^2 + \left[V(4m^2)\right]^2\Big). \end{aligned} \tag{10.53}$$

But the difference $V(p^2) - V(4m^2)$ is finite, as was required for the cancellation of divergences in the one-loop calculation:

$$V(p^2) - V(4m^2) = \frac{1}{32\pi^2}\int_0^1 dx\, \log\left(\frac{m^2 - x(1-x)p^2}{m^2 - x(1-x)4m^2}\right).$$

The only remaining divergence is in the term $[V(4m^2)]^2$, which is independent of momentum and can therefore be absorbed into the second-order counterterm in (10.51).

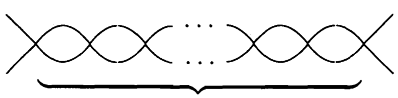
Two general properties of result (10.53) are worth noting. First, the divergent piece (and hence the $\mathcal{O}(\lambda^3)$ vertex counterterm) is proportional to

$$\left[V(4m^2)\right]^2 \propto \left[\Gamma(2-\tfrac{d}{2})\right]^2 \underset{d\to 4}{\longrightarrow} \left(\frac{2}{\epsilon}\right)^2 \qquad \text{for } d = 4-\epsilon.$$

This is a double pole, in contrast to the simple pole we found for the one-loop counterterm. Higher-loop diagrams will similarly have higher-order poles, but in all cases the divergent terms are momentum-independent constants. Second, consider the large-momentum limit,

$$V(p^2) - V(4m^2) \underset{p^2\to\infty}{\sim} \log\frac{p^2}{m^2}.$$

The two-loop vertex is proportional to $\log^2(p^2/m^2)$. A diagram of this structure with n loops will have the form

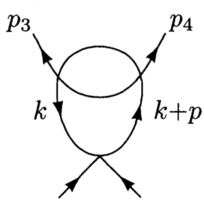


$$\sim \lambda^{n+1}\left(\log\frac{p^2}{m^2}\right)^n.$$

n loops

This asymptotic behavior is actually a generic property of multiloop diagrams, which we will explore in more detail in Chapter 12.

Now consider the more difficult diagram, from Group II:



$$= (-i\lambda)^3 \int \frac{d^dk}{(2\pi)^d} \frac{i}{k^2-m^2} \frac{i}{(k+p)^2-m^2}\, iV\big((k+p_3)^2\big). \qquad (10.54)$$

In evaluating this diagram, we will combine denominators in the manner that makes it most straightforward to extract the divergent terms, at the price of complicating the evaluation of the finite parts. Another approach to the calculation of this diagram is discussed in Problem 10.4.

To begin the evaluation of (10.54), combine the pair of denominators shown explicitly, and substitute expression (10.52) for $V(p^2)$. This gives the expression

$$-\frac{\lambda^3}{2}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \int_0^1 dx \int_0^1 dy \int \frac{d^dk}{(2\pi)^d} \frac{1}{\left[k^2 + 2yk\cdot p + yp^2 - m^2\right]^2}$$
$$\times \frac{1}{\left[m^2 - x(1-x)(k+p_3)^2\right]^{2-\frac{d}{2}}}. \qquad (10.55)$$

It is possible to combine this pair of denominators by using the identity

$$\frac{1}{A^\alpha B^\beta} = \int_0^1 dw\, \frac{w^{\alpha-1}(1-w)^{\beta-1}}{\big[wA+(1-w)B\big]^{\alpha+\beta}} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}. \tag{10.56}$$

This is a special case of the formula quoted in Section 6.3, Eq. (6.42). To prove it, change variables in the integral:

$$z \equiv \frac{wA}{wA+(1-w)B}, \qquad (1-z) = \frac{(1-w)B}{wA+(1-w)B}, \qquad dz = \frac{AB\,dw}{\big[wA+(1-w)B\big]^2},$$

so that

$$\int_0^1 dw\, \frac{w^{\alpha-1}(1-w)^{\beta-1}}{\big[wA+(1-w)B\big]^{\alpha+\beta}} = \frac{1}{A^\alpha B^\beta}\int_0^1 dz\, z^{\alpha-1}(1-z)^{\beta-1} = \frac{1}{A^\alpha B^\beta}B(\alpha,\beta),$$

where $B(\alpha,\beta)$ is the beta function, Eq. (7.82). The more general identity (6.42) can be proved by induction.

Applying identity (10.56) to (10.55), we obtain

$$= -\frac{\lambda^3\,\Gamma(4-\frac{d}{2})}{2\ (4\pi)^{d/2}} \int_0^1 dx \int_0^1 dy \int_0^1 dw \int \frac{d^dk}{(2\pi)^d}$$
$$\times \frac{w^{1-\frac{d}{2}}(1-w)}{\Big(w\big[m^2-x(1-x)(k+p_3)^2\big]+(1-w)\big[m^2-k^2-2yk\cdot p-yp^2\big]\Big)^{4-\frac{d}{2}}}. \tag{10.57}$$

Completing the square in the denominator yields a polynomial of the form

$$-\big[(1-w)+wx(1-x)\big]\ell^2 - P^2 + m^2, \tag{10.58}$$

where ℓ is a shifted momentum variable and P^2 is a rather complicated function of p, p_3, and the various Feynman parameters. It will only be important for this analysis that, as $w \to 0$,

$$P^2(w) = y(1-y)p^2 + \mathcal{O}(w), \tag{10.59}$$

and this can be seen easily from (10.57). Changing variables to ℓ, Wick-rotating, and performing the integral, we eventually obtain

$$= -\frac{i\lambda^3}{2(4\pi)^d}\int_0^1 dx \int_0^1 dy \int_0^1 dw\, \frac{w^{1-\frac{d}{2}}(1-w)}{\big[1-w+wx(1-x)\big]^{d/2}} \frac{\Gamma(4-d)}{(m^2-P^2)^{4-d}}. \tag{10.60}$$

This expression has one obvious pole as $d \to 4$, coming from the gamma function. However, it also has a less obvious pole, coming from the zero end of the w integral. Let us write (10.60) as

$$\int_0^1 dw\, w^{1-\frac{d}{2}}\, f(w),$$

where $f(w)$ incorporates all the factors not displayed explicitly. To isolate the pole at $w = 0$, we can add and subtract $f(0)$:

$$\int_0^1 dw\, w^{1-\frac{d}{2}}\, f(w) = \int_0^1 dw\, w^{1-\frac{d}{2}}\, f(0) + \int_0^1 dw\, w^{1-\frac{d}{2}}\, \big[f(w) - f(0)\big]. \tag{10.61}$$

The second piece is

$$-\frac{i\lambda^3\Gamma(4-d)}{2(4\pi)^d}\int_0^1 dx \int_0^1 dy \int_0^1 dw\, w^{1-\frac{d}{2}}$$
$$\times\left(\frac{(1-w)}{\big[1-w+wx(1-x)\big]^{d/2}}\,\frac{1}{\big[m^2-P^2(w)\big]^{4-d}} - \frac{1}{\big[m^2-P^2(0)\big]^{4-d}}\right).$$

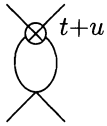
This term has only a simple pole as $d \to 4$; the residue of the pole is a momentum-independent constant, obtained by setting $d = 4$ everywhere except in $\Gamma(4-d)$. We can therefore absorb this divergence into the $\mathcal{O}(\lambda^3)$ vertex counterterm. (The finite part of this expression has a very complicated dependence on momentum, but we do not need to work this out to complete our argument.)

We are left with only the first term of (10.61). This expression contains only $P^2(0)$, which is given by (10.59). The w integral in this term is straightforward, and the x integral is trivial. With $\epsilon = 4-d$, our remaining expression is

$$-\frac{i\lambda^3}{2(4\pi)^d}\left(\frac{2}{\epsilon}\right)\int_0^1 dy\,\frac{\Gamma(\epsilon)}{\big[m^2-y(1-y)p^2\big]^{\epsilon}}$$
$$\underset{d\to 4}{\longrightarrow}\; -\frac{i\lambda^3}{2(4\pi)^4}\left(\frac{2}{\epsilon}\right)\int_0^1 dy\left(\frac{1}{\epsilon}-\gamma+\log(4\pi)-\log\big[m^2-y(1-y)p^2\big]\right), \tag{10.62}$$

where we have kept only the divergent terms in the second line. The logarithm, multiplied by the pole $2/\epsilon$, is the nonlocal divergence that we worried about in Section 10.4.

Fortunately, we must still add to this the "$t+u$" counterterm diagram of Group II. The computation of that diagram is by now a straightforward process:



$$= (-i\lambda)^3 \cdot -2iV(0) \cdot iV(p^2)$$

$$= \frac{i\lambda^3}{2(4\pi)^d} \int_0^1 dy \, \frac{\Gamma(2-\frac{d}{2})}{\left[m^2\right]^{2-d/2}} \frac{\Gamma(2-\frac{d}{2})}{\left[m^2 - y(1-y)p^2\right]^{2-d/2}}$$

$$\underset{d\to 4}{\longrightarrow} \frac{i\lambda^3}{2(4\pi)^4} \int_0^1 dy \left(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log m^2\right)$$

$$\times \left(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log\left[m^2 - y(1-y)p^2\right]\right). \quad (10.63)$$

(Again we have dropped finite terms from the last line.) This expression also contains a nonlocal divergence, given by the first pole times the second logarithm. It exactly cancels the nonlocal divergence in (10.62). The remaining terms are all either finite, or divergent but independent of momentum. This completes the proof that the two-loop contribution to the four-point function is finite.

The two features of the Group I diagrams appear here in Group II as well. The divergent pieces of (10.62) and (10.63) contain double poles that do not cancel, so we again find that the second-order vertex counterterm must contain a double pole. The finite pieces of (10.62) and (10.63) contain double logarithms, so we again find that the two-loop amplitude behaves as $\lambda^3 \log^2 p^2$ as $p \to \infty$.

Problems

10.1 One-loop structure of QED. In Section 10.1 we argued from general principles that the photon one-point and three-point functions vanish, while the four-point function is finite.

(a) Verify directly that the one-loop diagram contributing to the one-point function vanishes. There are two Feynman diagrams contributing to the three-point function at one-loop order. Show that these cancel. Show that the diagrams contributing to any n-point photon amplitude, for n odd, cancel in pairs.

(b) The photon four-point amplitude is a sum of six diagrams. Show explicitly that the potential logarithmic divergences of these diagrams cancel.

10.2 Renormalization of Yukawa theory. Consider the pseudoscalar Yukawa Lagrangian,

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \phi)^2 - \tfrac{1}{2}m^2\phi^2 + \bar{\psi}(i\partial\!\!\!/ - M)\psi - ig\bar{\psi}\gamma^5\psi\phi,$$

where ϕ is a real scalar field and ψ is a Dirac fermion. Notice that this Lagrangian is invariant under the parity transformation $\psi(t,\mathbf{x}) \to \gamma^0\psi(t,-\mathbf{x})$, $\phi(t,\mathbf{x}) \to -\phi(t,-\mathbf{x})$, in which the field ϕ carries odd parity.

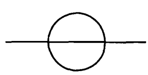
(a) Determine the superficially divergent amplitudes and work out the Feynman rules for renormalized perturbation theory for this Lagrangian. Include all necessary counterterm vertices. Show that the theory contains a superficially divergent 4ϕ amplitude. This means that the theory cannot be renormalized unless one includes a scalar self-interaction,

$$\delta\mathcal{L} = \frac{\lambda}{4!}\phi^4,$$

and a counterterm of the same form. It is of course possible to set the renormalized value of this coupling to zero, but that is not a natural choice, since the counterterm will still be nonzero. Are any further interactions required?

(b) Compute the divergent part (the pole as $d \to 4$) of each counterterm, to the one-loop order of perturbation theory, implementing a sufficient set of renormalization conditions. You need not worry about finite parts of the counterterms. Since the divergent parts must have a fixed dependence on the external momenta, you can simplify this calculation by choosing the momenta in the simplest possible way.

10.3 Field-strength renormalization in ϕ^4 theory. The two-loop contribution to the propagator in ϕ^4 theory involves the three diagrams shown in (10.31). Compute the first of these diagrams in the limit of zero mass for the scalar field, using dimensional regularization. Show that, near $d = 4$, this diagram takes the form:



$$= -ip^2 \cdot \frac{\lambda^2}{12(4\pi)^4}\left[-\frac{1}{\epsilon} + \log p^2 + \cdots\right],$$

with $\epsilon = 4 - d$. The coefficient in this equation involves a Feynman parameter integral that can be evaluated by setting $d = 4$. Verify that the second diagram in (10.31) vanishes near $d = 4$. Thus the first diagram should contain a pole only at $\epsilon = 0$, which can be canceled by a field-strength renormalization counterterm.

10.4 Asymptotic behavior of diagrams in ϕ^4 theory. Compute the leading terms in the S-matrix element for boson-boson scattering in ϕ^4 theory in the limit $s \to \infty$, t fixed. Ignore all masses on internal lines, and keep external masses nonzero only as infrared regulators where these are needed. Show that

$$i\mathcal{M}(s,t) \sim -i\lambda - i\frac{\lambda^2}{(4\pi)^2}\log s - i\frac{5\lambda^3}{2(4\pi)^4}\log^2 s + \cdots.$$

Notice that ignoring the internal masses allows some pleasing simplifications of the Feynman parameter integrals.

Chapter 11

Renormalization and Symmetry

Now that we have determined the general structure of the ultraviolet divergences of quantum field theories, it would seem natural to continue investigating the implications of these divergences in Feynman diagram calculations. However, we will now put this issue aside until Chapter 12 and set off in what may seem an unrelated direction. In Chapter 8 and in Section 9.3, we noted the formal relation between quantum field theory and statistical mechanics. The closest formal analogue of a scalar field theory was seen to be the continuum description of a ferromagnet or some other system that allows a second-order phase transition. This analogy raises the possibility that in quantum field theory as well it may be possible for the field to take on a nonzero global value. As in a magnet, this global field might have a directional character, and thus violate a symmetry of the Lagrangian. In such a case, we say that the field theory has *hidden* or *spontaneously broken* symmetry. We devote this chapter to an analysis of this mechanism of symmetry violation.

Spontaneously broken symmetry is a central concept in the study of quantum field theory, for two reasons. First, it plays a major role in the applications of quantum field theory to Nature. In this book, we will see two very different examples of such applications: Chapter 13 will apply the theory of hidden symmetry to statistical mechanics, specifically to the behavior of thermodynamic variables near second-order phase transitions. Later, in Chapter 20, we will see that hidden symmetry is an essential ingredient in the theory of the weak interactions. Spontaneous symmetry breaking also finds applications in the theory of the strong interactions, and in the search for unified models of fundamental physics.

But spontaneous symmetry breaking is also interesting from a theoretical point of view. Quantum field theories with spontaneously broken symmetry contain ultraviolet divergences. Thus, it is natural to ask whether these divergences are constrained by the underlying symmetry of the theory. The answer to this question, first presented by Benjamin Lee,* will give us further insights into the nature of ultraviolet divergences and the meaning of renormalization.

---

*A beautiful summary of Lee's analysis is given in his lecture note volume: B. Lee, *Chiral Dynamics* (Gordon and Breach, New York, 1972).

11.1 Spontaneous Symmetry Breaking

We begin with an analysis of spontaneous symmetry breaking in classical field theory. Consider first the familiar ϕ^4 theory Lagrangian,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4,$$

but with m^2 replaced by a negative parameter, $-\mu^2$:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2}\mu^2\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{11.1}$$

This Lagrangian has a discrete symmetry: It is invariant under the operation $\phi \to -\phi$. The corresponding Hamiltonian is

$$H = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}\mu^2\phi^2 + \frac{\lambda}{4!}\phi^4\right].$$

The minimum-energy classical configuration is a uniform field $\phi(x) = \phi_0$, with ϕ_0 chosen to minimize the potential

$$V(\phi) = -\frac{1}{2}\mu^2\phi^2 + \frac{\lambda}{4!}\phi^4$$

(see Fig. 11.1). This potential has two minima, given by

$$\phi_0 = \pm v = \pm\sqrt{\frac{6}{\lambda}}\,\mu. \tag{11.2}$$

The constant v is called the *vacuum expectation value* of ϕ.

To interpret this theory, suppose that the system is near one of the minima (say the positive one). Then it is convenient to define

$$\phi(x) = v + \sigma(x), \tag{11.3}$$

and rewrite $\mathcal{L}$ in terms of $\sigma(x)$. Plugging (11.3) into (11.1), we find that the term linear in σ vanishes (as it must, since the minimum of the potential is at $\sigma = 0$). Dropping the constant term as well, we obtain the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \sigma)^2 - \frac{1}{2}(2\mu^2)\sigma^2 - \sqrt{\frac{\lambda}{6}}\,\mu\sigma^3 - \frac{\lambda}{4!}\sigma^4. \tag{11.4}$$

This Lagrangian describes a simple scalar field of mass $\sqrt{2}\mu$, with σ^3 and σ^4 interactions. The symmetry $\phi \to -\phi$ is no longer apparent; its only manifestation is in the relations among the three coefficients in (11.4), which depend in a special way on only two parameters. This is the simplest example of a spontaneously broken symmetry.

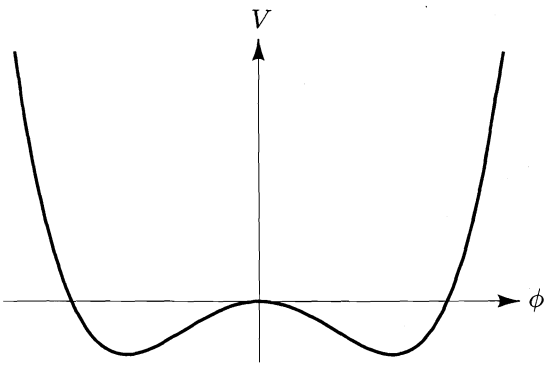


Figure 11.1. Potential for spontaneous symmetry breaking in the discrete case.

The Linear Sigma Model

A more interesting theory arises when the broken symmetry is continuous, rather than discrete. The most important example is a generalization of the preceding theory called the *linear sigma model*, which we considered briefly in Problem 4.3. We will study this model in detail throughout this chapter.

The Lagrangian of the linear sigma model involves a set of N real scalar field $\phi^i(x)$:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi^i)^2 + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}\left[(\phi^i)^2\right]^2, \tag{11.5}$$

with an implicit sum over i in each factor $(\phi^i)^2$. Note that we have rescaled the coupling λ from the ϕ^4 theory Lagrangian to remove the awkward factors of 6 in the analysis above. The Lagrangian (11.5) is invariant under the symmetry

$$\phi^i \longrightarrow R^{ij}\phi^j \tag{11.6}$$

for any $N \times N$ orthogonal matrix R. The group of transformations (11.6) is just the rotation group in N dimensions, also called the N-dimensional *orthogonal group* or simply $O(N)$.

Again the lowest-energy classical configuration is a constant field ϕ_0^i, whose value is chosen to minimize the potential

$$V(\phi^i) = -\frac{1}{2}\mu^2(\phi^i)^2 + \frac{\lambda}{4}\left[(\phi^i)^2\right]^2$$

(see Fig. 11.2). This potential is minimized for any ϕ_0^i that satisfies

$$(\phi_0^i)^2 = \frac{\mu^2}{\lambda}.$$

This condition determines only the length of the vector ϕ_0^i; its direction is arbitrary. It is conventional to choose coordinates so that ϕ_0^i points in the

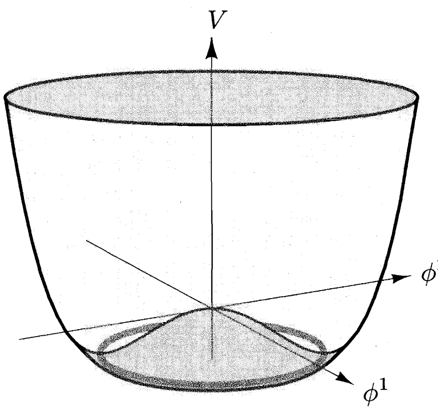


Figure 11.2. Potential for spontaneous breaking of a continuous $O(N)$ symmetry, drawn for the case $N = 2$. Oscillations along the trough in the potential correspond to the massless π fields.

Nth direction:

$$\phi_0^i = (0, 0, \ldots, 0, v), \qquad \text{where } v = \frac{\mu}{\sqrt{\lambda}}. \tag{11.7}$$

We can now define a set of shifted fields by writing

$$\phi^i(x) = \big(\pi^k(x), v + \sigma(x)\big), \qquad k = 1, \ldots, N{-}1. \tag{11.8}$$

(The notation, as in Problem 4.3, comes from the application of this formalism to pions in the case $N = 4$.)

It is now straightforward to rewrite the Lagrangian (11.5) in terms of the π and σ fields. The result is

$$\begin{aligned} \mathcal{L} = {} & \frac{1}{2}(\partial_\mu \pi^k)^2 + \frac{1}{2}(\partial_\mu \sigma)^2 - \frac{1}{2}(2\mu^2)\sigma^2 \\ & - \sqrt{\lambda}\mu\sigma^3 - \sqrt{\lambda}\mu(\pi^k)^2\sigma - \frac{\lambda}{4}\sigma^4 - \frac{\lambda}{2}(\pi^k)^2\sigma^2 - \frac{\lambda}{4}\big[(\pi^k)^2\big]^2. \end{aligned} \tag{11.9}$$

We obtain a massive σ field just as in (11.4), and also a set of $N{-}1$ massless π fields. The original $O(N)$ symmetry is hidden, leaving only the subgroup $O(N{-}1)$, which rotates the π fields among themselves. Referring to Fig. 11.2, we note that the massive σ field describes oscillations of ϕ^i in the radial direction, in which the potential has a nonvanishing second derivative. The massless π fields describe oscillations of ϕ^i in the tangential directions, along the trough of the potential. The trough is an $(N{-}1)$-dimensional surface, and all $N{-}1$ directions are equivalent, reflecting the unbroken $O(N{-}1)$ symmetry.

Goldstone's Theorem

The appearance of massless particles when a continuous symmetry is spontaneously broken is a general result, known as *Goldstone's theorem.* To state the theorem precisely, we must count the number of linearly independent continuous symmetry transformations. In the linear sigma model, there are no continuous symmetries for $N = 1$, while for $N = 2$ there is a single direction of rotation. A rotation in N dimensions can be in any one of $N(N-1)/2$ planes, so the $O(N)$-symmetric theory has $N(N-1)/2$ continuous symmetries. After spontaneous symmetry breaking there are $(N-1)(N-2)/2$ remaining symmetries, corresponding to rotations of the $(N-1)$ π fields. The number of *broken* symmetries is the difference, $N-1$.

Goldstone's theorem states that for every spontaneously broken continuous symmetry, the theory must contain a massless particle.† We have just seen that this theorem holds in the linear sigma model, at least at the classical level. The massless fields that arise through spontaneous symmetry breaking are called *Goldstone bosons.* Many light bosons seen in physics, such as the pions, may be interpreted (at least approximately) as Goldstone bosons. We conclude this section with a general proof of Goldstone's theorem for classical scalar field theories. The rest of this chapter is devoted to the quantum-mechanical analysis of theories with hidden symmetry. By the end of the chapter we will see that Goldstone bosons cannot acquire mass from any order of quantum corrections.

Consider, then, a theory involving several fields $\phi^a(x)$, with a Lagrangian of the form

$$\mathcal{L} = (\text{terms with derivatives}) - V(\phi). \tag{11.10}$$

Let ϕ_0^a be a constant field that minimizes V, so that

$$\frac{\partial}{\partial \phi^a} V\bigg|_{\phi^a(x)=\phi_0^a} = 0.$$

Expanding V about this minimum, we find

$$V(\phi) = V(\phi_0) + \frac{1}{2}(\phi - \phi_0)^a(\phi - \phi_0)^b \left(\frac{\partial^2}{\partial\phi^a \partial\phi^b} V\right)_{\phi_0} + \cdots .$$

The coefficient of the quadratic term,

$$\left(\frac{\partial^2}{\partial\phi^a \partial\phi^b} V\right)_{\phi_0} = m_{ab}^2, \tag{11.11}$$

†J. Goldstone, *Nuovo Cim.* **19**, 154 (1961). An instructive four-page paper by J. Goldstone, A. Salam, and S. Weinberg, *Phys. Rev.* **127**, 965 (1962), gives three different proofs of the theorem.

is a symmetric matrix whose eigenvalues give the masses of the fields. These eigenvalues cannot be negative, since ϕ_0 is a minimum. To prove Goldstone's theorem, we must show that every continuous symmetry of the Lagrangian (11.10) that is not a symmetry of ϕ_0 gives rise to a zero eigenvalue of this mass matrix.

A general continuous symmetry transformation has the form

$$\phi^a \longrightarrow \phi^a + \alpha \Delta^a(\phi), \tag{11.12}$$

where α is an infinitesimal parameter and Δ^a is some function of all the ϕ's. Specialize to constant fields; then the derivative terms in $\mathcal{L}$ vanish and the potential alone must be invariant under (11.12). This condition can be written

$$V(\phi^a) = V\big(\phi^a + \alpha\Delta^a(\phi)\big) \qquad \text{or} \qquad \Delta^a(\phi)\frac{\partial}{\partial\phi^a}V(\phi) = 0.$$

Now differentiate with respect to ϕ^b, and set $\phi = \phi_0$:

$$0 = \left(\frac{\partial\Delta^a}{\partial\phi^b}\right)_{\phi_0}\left(\frac{\partial V}{\partial\phi^a}\right)_{\phi_0} + \Delta^a(\phi_0)\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V\right)_{\phi_0}. \tag{11.13}$$

The first term vanishes since ϕ_0 is a minimum of V, so the second term must also vanish. If the transformation leaves ϕ_0 unchanged (i.e., if the symmetry is respected by the ground state), then $\Delta^a(\phi_0) = 0$ and this relation is trivial. A spontaneously broken symmetry is precisely one for which $\Delta^a(\phi_0) \neq 0$; in this case $\Delta^a(\phi_0)$ is our desired vector with eigenvalue zero, so Goldstone's theorem is proved.

11.2 Renormalization and Symmetry: An Explicit Example

Now let us investigate the quantum mechanics of a theory with spontaneously broken symmetry. Again we will use as our example the linear sigma model. The Lagrangian of this theory, written in terms of shifted fields, is given in Eq. (11.9). From this expression, we can read off the Feynman rules; these are shown in Fig. 11.3.

Using these Feynman rules, we can compute tree-level amplitudes without difficulty. Diagrams with loops, however, will often diverge. For the amplitude with N_e external legs, the superficial degree of divergence is

$$D = 4 - N_e,$$

just as in the discussion of ϕ^4 theory in Section 10.2. (Diagrams containing a three-point vertex will be less divergent than this expression indicates, because this vertex has a coefficient with dimensions of mass.) However, the symmetry constraints on the amplitudes are much weaker than in that earlier analysis. The linear sigma model has eight different superficially divergent amplitudes (see Fig. 11.4); several of these have $D > 0$ and therefore can contain

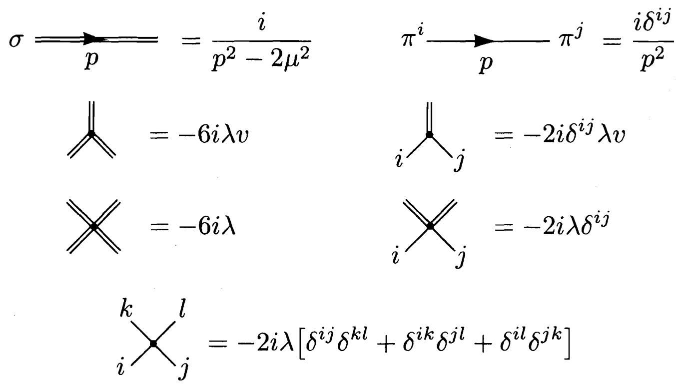


Figure 11.3. Feynman rules for the linear sigma model.

more than one infinite constant. Yet the number of bare parameters available to absorb these infinities is much smaller. If we follow the procedure of Section 10.2 to rewrite the original Lagrangian in terms of physical parameters and counterterms, we find only three counterterms:

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial_\mu\phi^i)^2 + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}\left[(\phi^i)^2\right]^2 \\ &\quad + \frac{1}{2}\delta_Z(\partial_\mu\phi^i)^2 - \frac{1}{2}\delta_\mu(\phi^i)^2 - \frac{\delta_\lambda}{4}\left[(\phi^i)^2\right]^2.\end{aligned} \tag{11.14}$$

Written in terms of π and σ fields, the second line takes the form

$$\begin{aligned}&\frac{\delta_Z}{2}(\partial_\mu\pi^k)^2 - \frac{1}{2}(\delta_\mu + \delta_\lambda v^2)(\pi^k)^2 + \frac{\delta_Z}{2}(\partial_\mu\sigma)^2 - \frac{1}{2}(\delta_\mu + 3\delta_\lambda v^2)\sigma^2 \\ &\qquad - (\delta_\mu v + \delta_\lambda v^3)\sigma - \delta_\lambda v\sigma(\pi^k)^2 - \delta_\lambda v\sigma^3 \\ &\qquad - \frac{\delta_\lambda}{4}\left[(\pi^k)^2\right]^2 - \frac{\delta_\lambda}{2}\sigma^2(\pi^k)^2 - \frac{\delta_\lambda}{4}\sigma^4.\end{aligned} \tag{11.15}$$

The Feynman rules associated with these counterterms are shown in Fig. 11.5. There are now plenty of counterterms, but they still depend on only three renormalization parameters: δ_Z, δ_μ, and δ_λ. It would be a miracle if these three parameters were able to absorb all the infinities arising in the divergent amplitudes shown in Fig. 11.4.

If this miracle did not occur, that is, if the counterterms of (11.15) did not absorb all the infinities, we could still make this theory renormalizable by introducing new, symmetry-breaking terms in the Lagrangian. These would give rise to additional counterterms, which could be adjusted to render all amplitudes finite. If desired, we could set the physical values of the symmetry-breaking coupling constants to zero. The bare values of these constants, however, would still be nonzero, so the Lagrangian itself would no longer be invariant under the $O(N)$ symmetry. We would have to conclude that the symmetry is not consistent with quantum mechanics.

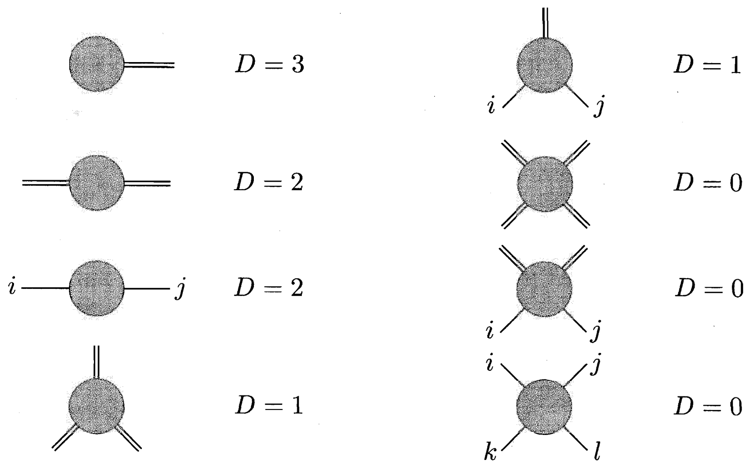


Figure 11.4. Divergent amplitudes in the linear sigma model.

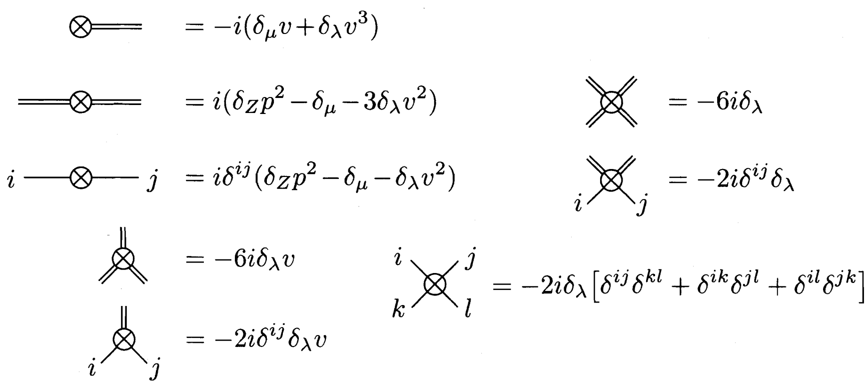
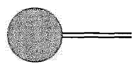


Figure 11.5. Feynman rules for counterterm vertices in the linear sigma model.

Fortunately, the miracle does occur. We will see below that the counterterms of (11.15), even though they contain only three adjustable parameters, are indeed sufficient to cancel all the infinities that occur in this theory. In this section we will demonstrate this cancellation explicitly at the one-loop level. The rest of this chapter is devoted to a more general discussion of these issues.

Renormalization Conditions

In the discussion to follow, we will keep track of only the divergent parts of Feynman diagrams. However, it will be useful to keep in mind a set of renormalization conditions that could, in principle, be used to determine also the finite parts of the counterterms. Since the counterterms contain three adjustable parameters, we need three conditions. We could take these to be the conditions (10.19) (implemented according to (10.28)), specifying the physical mass m of the σ field, its field strength, and the scattering amplitude at threshold. However, it is technically easier to replace one of these conditions with a constraint on the one-point amplitude for σ (the sum of *tadpole diagrams*):

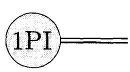
 $= 0.$

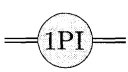
In QED the tadpole diagrams automatically vanish, as we saw in Eq. (10.5). In the linear sigma model, however, no symmetry forbids the appearance of a nonvanishing one-σ amplitude. This amplitude produces a vacuum expectation value of σ and so, since $\phi^N = v + \sigma$, shifts the vacuum expectation value of ϕ. Such a shift is quite acceptable, as long as it is finite after counterterms are properly added into the computation of the amplitude. However, it will simplify the bookkeeping to set up our conventions so that the relation

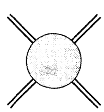
$$\langle \phi^N \rangle = \frac{\mu}{\sqrt{\lambda}} \tag{11.16}$$

is satisfied to all orders in perturbation theory. We will define λ, as in Eq. (10.19), as the scattering amplitude at threshold. Then Eq. (11.16) defines the parameter μ, so the mass m of the σ field will differ from the result of the classical equations $m^2 = 2\mu^2 = 2\lambda v^2$ by terms of order $(\lambda\mu^2)$. If indeed we can remove the divergences from the theory by adjusting three counterterms, these corrections will be finite and constitute a prediction of the quantum field theory.

To summarize, we will use the following renormalization conditions:

 $= 0;$

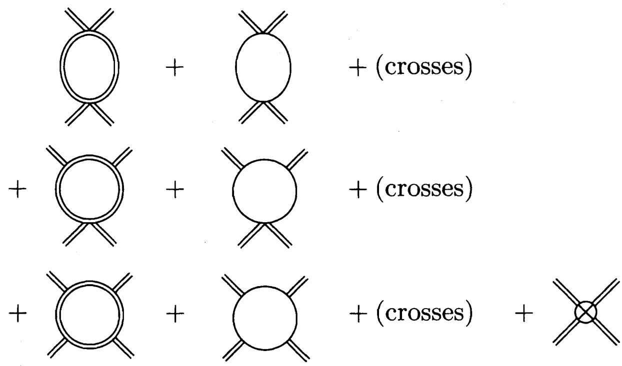
$\dfrac{d}{dp^2}\Big($  $\Big) = 0 \qquad$ at $p^2 = m^2;$ (11.17)

$\mathrm{Im}\Big($  $\Big) = -6i\lambda \qquad$ at $s = 4m^2,\ t = u = 0.$

In the last condition, the circle is the amputated four-point amplitude. Note that the last two conditions depend on the physical mass m of the σ particle. We must now show that these three conditions suffice to make all of the one-loop amplitudes of the linear sigma model finite.

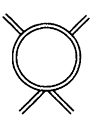
The Vertex Counterterm

We begin by determining the counterterm δ_λ by computing the 4σ amplitude. The tree-level term comes from the 4σ vertex, and is just such as to satisfy (11.17). The one-loop contribution to this amplitude is the sum of diagrams:



(11.18)

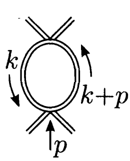
According to (11.17), we must adjust δ_λ so that this sum of diagrams vanishes at threshold. In this calculation, we will only keep track of the ultraviolet divergences. This greatly simplifies the analysis, because most of the diagrams in (11.18) are finite. All the diagrams with loops made of three or more propagators are finite, since they have at least six powers of the loop momentum in the denominator; for example,



$$\sim \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2}\frac{1}{k^2}\frac{1}{k^2}.$$

Alternatively, we can see that this diagram is finite in the following way: Each three-point vertex carries a factor of μ, which has dimensions of mass. According to the dimensional analysis argument of Section 10.1, each such factor lowers the degree of divergence of a diagram by 1. Since the 4σ amplitude already has $D = 0$, any diagram containing a three-point vertex must be finite.

We are left with the first two diagrams of (11.18) and the four diagrams related to these by crossing. Let us evaluate the first diagram using dimensional regularization:



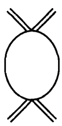
$$= \frac{1}{2}\cdot(-6i\lambda)^2 \cdot \int \frac{d^dk}{(2\pi)^d} \frac{i}{k^2 - 2\mu^2} \frac{i}{(k+p)^2 - 2\mu^2}$$

$$= 18\lambda^2 \int_0^1 dx \int \frac{d^dk}{(2\pi)^d} \frac{1}{[k^2 - \Delta]^2}$$

$$= 18\lambda^2 \int_0^1 dx \, \frac{i}{(4\pi)^{d/2}} \, \Gamma(2-\tfrac{d}{2}) \left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}}$$

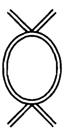
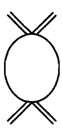
$$= 18i\lambda^2 \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} + (\text{finite terms}). \tag{11.19}$$

Here Δ is a function of p and μ, whose exact form does not concern us. Since our objective is only to demonstrate the cancellation of the divergences, we will neglect finite terms here and throughout the rest of this section. The second diagram of (11.18) (with π's instead of σ's for the internal lines) is identical, except that each vertex factor is changed from $-6i\lambda$ to $-2i\lambda\delta^{ij}$. (Roman indices $i, j, \ldots$ run from 1 to $N-1$.) We therefore have



$$= 2i\lambda^2(N-1)\,\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} + (\text{finite terms}). \tag{11.20}$$

Since the infinite part of each of these diagrams is simply a momentum-independent constant, the infinite parts of the corresponding t- and u-channel diagrams must be identical. Therefore the infinite part of the 4σ vertex is just three times the sum of (11.19) and (11.20):

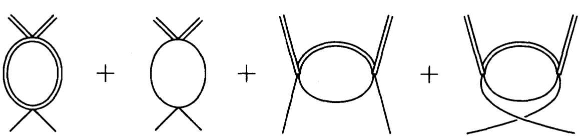
 + 

$$+ \text{ (crosses)} \quad \sim 6i\lambda^2(N+8)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \tag{11.21}$$

(In this section we use the $\sim$ symbol to indicate equality up to omitted finite corrections.) Applying the third condition of (11.17), we find that the counterterm δ_λ is given by

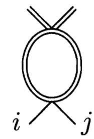
$$\delta_\lambda \sim \lambda^2(N+8)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \tag{11.22}$$

Once we have determined the value of δ_λ, we have fixed the counterterms for the two other four-point amplitudes. Are these amplitudes also made finite? Consider the amplitude with two σ's and two π's. This receives one-loop corrections from

 (11.23)

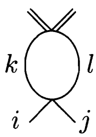
and from several diagrams with three-point vertices which, as argued earlier, are manifestly finite. Each of the diagrams in (11.23) contains a loop integral analogous to that in (11.19), whose infinite part is always $-i\Gamma(2-\frac{d}{2})/(4\pi)^2$.

The only differences are in the vertices and symmetry factors. For example, the infinite part of the first diagram of (11.23) is



$$\sim \frac{1}{2}\cdot(-6i\lambda)(-2i\lambda\delta^{ij})\cdot\frac{-i}{(4\pi)^2}\Gamma(2-\tfrac{d}{2}) = 6i\lambda^2\delta^{ij}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

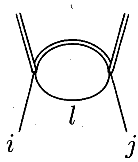
The second diagram is a bit more complicated:



$$\sim \frac{1}{2}\cdot(-2i\lambda\delta^{kl})\big(-2i\lambda(\delta^{ij}\delta^{kl}+\delta^{ik}\delta^{jl}+\delta^{il}\delta^{jk})\big)\cdot\frac{-i}{(4\pi)^2}\Gamma(2-\tfrac{d}{2})$$

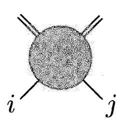
$$= 2i\lambda^2(N+1)\delta^{ij}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

In the third diagram there is no symmetry factor:



$$\sim (-2i\lambda\delta^{il})(-2i\lambda\delta^{jl})\cdot\frac{-i}{(4\pi)^2}\Gamma(2-\tfrac{d}{2}) = 4i\lambda^2\delta^{ij}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

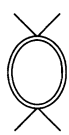
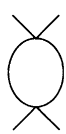
The fourth diagram of (11.23) gives an identical expression, since it is the same as the third but with i and j interchanged. The sum of the four diagrams therefore gives, for the infinite part of the $\sigma\sigma\pi\pi$ vertex,



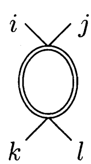
$$\sim 2i\lambda^2\,\delta^{ij}\,(N+8)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \tag{11.24}$$

This divergent term is indeed canceled by the $\sigma\sigma\pi\pi$ counterterm, with the value of δ_λ given in (11.22).

The remaining four-point amplitude has four external π fields. The divergent one-loop diagrams are:

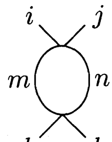
 +  + (crosses) (11.25)

These diagrams all have the same familiar form. The first is



$$\sim \frac{1}{2}\cdot(-2i\lambda\delta^{ij})(-2i\lambda\delta^{kl})\cdot\frac{-i}{(4\pi)^2}\Gamma(2-\tfrac{d}{2}) = 2i\lambda^2\delta^{ij}\delta^{kl}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

The second diagram is more complicated:

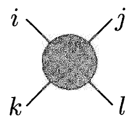


$$\sim \frac{1}{2}\cdot\big(-2i\lambda(\delta^{ij}\delta^{mn}+\delta^{im}\delta^{jn}+\delta^{in}\delta^{jm})\big)$$

$$\cdot\big(-2i\lambda(\delta^{kl}\delta^{mn}+\delta^{km}\delta^{ln}+\delta^{kn}\delta^{lm})\big)\cdot\frac{-i}{(4\pi)^2}\Gamma(2-\tfrac{d}{2})$$

$$= 2i\lambda^2\big((N+3)\delta^{ij}\delta^{kl} + 2\delta^{ik}\delta^{jl} + 2\delta^{il}\delta^{jk}\big)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

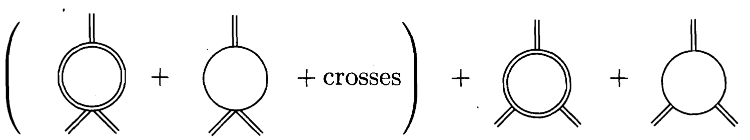
For each of these diagrams there are two corresponding cross-channel diagrams, which differ only in the ways that the external indices $ijkl$ are paired together. For instance, the t-channel diagrams are identical to the s-channel diagrams, but with j and k interchanged. Adding all six diagrams, we find for the 4π vertex



$$\sim 2i\lambda^2\,(\delta^{ij}\delta^{kl}+\delta^{ik}\delta^{jl}+\delta^{il}\delta^{jk})\,(N+8)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \qquad (11.26)$$

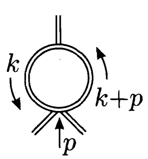
Again, the value of δ_λ given in (11.22) gives a counterterm of the correct value and index structure to cancel this divergence.

The value of δ_λ that we have determined also fixes the counterterms for the three-point amplitudes. Thus we have no further freedom in canceling the divergences in the three-point amplitudes; we can only cross our fingers and hope these also come out finite. The 3σ amplitude is given by



(11.27)

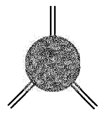
The diagrams made of three three-point vertices are finite and play no role in the cancellation of divergences. Of the divergent diagrams in (11.27), the first has the form



$$= \frac{1}{2}\cdot(-6i\lambda)(-6i\lambda v)\cdot\int\frac{d^dk}{(2\pi)^d}\,\frac{i}{k^2-2\mu^2}\,\frac{i}{(k+p)^2-2\mu^2}$$

$$\sim 18i\lambda^2 v\,\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

This is exactly the same as the corresponding diagram (11.19) for the 4σ vertex, except for the extra factor of v. The same is true of the other five divergent diagrams; thus,

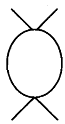


$$\sim 6i\lambda^2 v(N+8)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \tag{11.28}$$

This is precisely canceled by the 3σ counterterm vertex in Fig. 11.5, with δ_λ given by (11.22).

There is a similar correspondence between the $\sigma\pi\pi$ amplitude and the $\sigma\sigma\pi\pi$ amplitude. The four divergent diagrams in the $\sigma\pi\pi$ amplitude are identical to those in (11.23), except that each has an external σ leg replaced by a factor of v. Referring to the $\sigma\pi\pi$ counterterm vertex in Fig. 11.5, we see that the cancellation of divergences will occur here as well.

What is happening? All the divergences we have seen so far are manifestations of the basic diagram

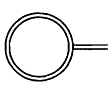
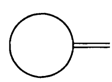
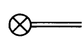


(11.29)

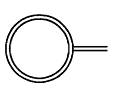
with either four external particles or with one leg set to zero momentum and associated with the vacuum expectation value of ϕ. Since the $O(N)$ symmetry is broken, this diagram manifests itself in many different ways. But apparently, the divergent part of the diagram is unaffected by the symmetry breaking.

Two-Point and One-Point Amplitudes

To complete our investigation of the one-loop structure of this theory we must evaluate the two-point and one-point amplitudes. We first determine the counterterm δ_μ by applying the first renormalization condition in (11.17). At one-loop order, this condition reads

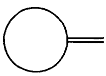
$0 =$  $+$  $+$  . (11.30)

We will later need to make use of the finite part of the counterterm, so we will pay attention to the finite terms when we evaluate (11.30). The first diagram is



$$= \frac{1}{2}(-6i\lambda v)\int\frac{d^dk}{(2\pi)^d}\frac{i}{k^2-2\mu^2} = -3i\lambda v\,\frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}}\left(\frac{1}{2\mu^2}\right)^{1-\frac{d}{2}}. \tag{11.31}$$

The second diagram involves a divergent integral over a massless propagator. To be sure that we understand how to treat this term, we will add a small mass ζ for the π field as an infrared regulator. Then the second diagram is

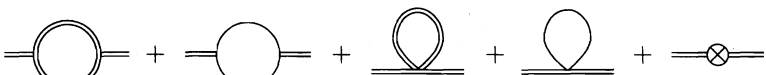


$$
\begin{aligned}
&= \frac{1}{2}(-2i\lambda v)\delta^{ij}\int\frac{d^dk}{(2\pi)^d}\frac{i\delta^{ij}}{k^2-\zeta^2} \\
&= -i(N-1)\lambda v\,\frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}}\left(\frac{1}{\zeta^2}\right)^{1-\frac{d}{2}}.
\end{aligned}
\tag{11.32}
$$

Notice that, for $d > 2$, the diagram vanishes in the limit as $\zeta \to 0$; however, it has a pole at $d = 2$. Despite these strange features, we can add (11.32) to (11.31) and impose the condition that the tadpole diagrams be canceled by the counterterm from Fig. 11.5. This condition gives

$$
(\delta_\mu + v^2\delta_\lambda) = -\lambda\,\frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}}\left(\frac{3}{(2\mu^2)^{1-d/2}} + \frac{N-1}{(\zeta^2)^{1-d/2}}\right). \tag{11.33}
$$

Now consider the 2σ amplitude. The one-particle-irreducible amplitude receives contributions from four one-loop diagrams and a counterterm:

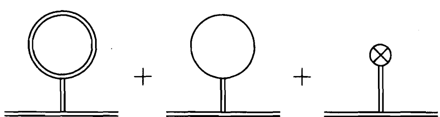


$$. \tag{11.34}$$

It is convenient to write the counterterm vertex as

$$
-i(2v^2\delta_\lambda) - i(\delta_\mu + v^2\delta_\lambda) + ip^2\delta_Z. \tag{11.35}
$$

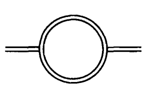
In a general renormalization scheme, the σ mass will also be shifted by the tadpole diagrams (and their counterterm):



$$. \tag{11.36}$$

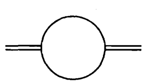
However, the first renormalization condition in (11.17) forces these diagrams to cancel precisely. This is an example of the special simplicity of this renormalization condition.

The first two diagrams are again manifestations of the generic four-point diagram (11.29), now with two external legs replaced by the vacuum expectation value of ϕ. In analogy with the preceding calculations, we find for the first diagram



$$
\sim 18i\lambda^2v^2\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2},
$$

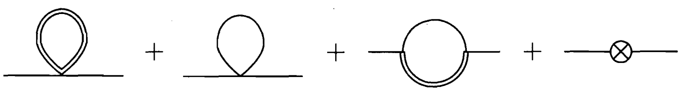
and for the second diagram



$$
\sim 2i\lambda^2v^2(N-1)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.
$$

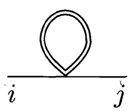
Using (11.22), we see that these two contributions are canceled by the first term of (11.35). The third and fourth diagrams of (11.34) contain precisely the same integrals as the tadpole diagrams of (11.30). Relation (11.33) implies that they are canceled by the second term in (11.35). Notice that there is no divergent term proportional to p^2 in any of the one-loop diagrams of (11.34). Thus the renormalization constant δ_Z is zero at the one-loop level, just as in ordinary ϕ^4 theory.

There remains only one potentially divergent amplitude—the $\pi\pi$ amplitude:



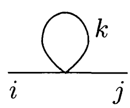
(11.37)

In analogy with (11.31), the first diagram is



$$= \frac{1}{2}(-2i\lambda\delta^{ij}) \int \frac{d^dk}{(2\pi)^d} \frac{i}{k^2 - 2\mu^2} = -i\lambda\delta^{ij} \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{1}{2\mu^2}\right)^{1-\frac{d}{2}}.$$

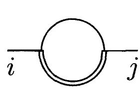
The second diagram is quite similar. As in (11.32), it is useful to introduce a small pion mass as an infrared regulator.



$$= \frac{1}{2}\left(-2i\lambda(\delta^{ij}\delta^{kk}+\delta^{ik}\delta^{jk}+\delta^{ik}\delta^{jk})\right) \int \frac{d^dk}{(2\pi)^d} \frac{1}{k^2 - \zeta^2}$$

$$= -i\lambda(N+1)\delta^{ij} \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{1}{\zeta^2}\right)^{1-\frac{d}{2}}.$$

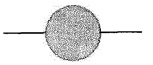
The third diagram is given by



$$= (-2i\lambda v\delta^{ik})(-2i\lambda v\delta^{kj}) \int \frac{d^dk}{(2\pi)^d} \frac{i}{k^2 - \zeta^2} \frac{i}{(k+p)^2 - 2\mu^2}$$

$$= 4i\lambda^2 v^2 \delta^{ij} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \int_0^1 dx \left(\frac{1}{2\mu^2 x + (1-x)\zeta^2 - p^2 x(1-x)}\right)^{2-\frac{d}{2}}.$$

The divergent part of this expression is independent of p, so to check the cancellation of the divergence, it suffices to set $p = 0$. It will be instructive to compute the complete amplitude at $p = 0$, including the finite terms. Adding the three loop diagrams and the counterterm, whose value is given by (11.33), we find



$$\Big|_{p=0} = (-i\lambda\delta^{ij})\left\{\frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}}\left(\frac{1}{(2\mu^2)^{1-d/2}} + \frac{N+1}{(\zeta^2)^{1-d/2}}\right)\right.$$
$$- 4\lambda v^2 \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \int_0^1 dx \left(\frac{1}{2\mu^2 x + \zeta^2(1-x)}\right)^{2-\frac{d}{2}}$$
$$\left. - \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}}\left(\frac{3}{(2\mu^2)^{1-d/2}} + \frac{N-1}{(\zeta^2)^{1-d/2}}\right)\right\}. \tag{11.38}$$

It is not hard to simplify this expression. The first and third lines can be combined to give

$$2i\lambda\delta^{ij}\frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}}\left[\frac{1}{(\zeta^2)^{1-d/2}} - \frac{1}{(2\mu^2)^{1-d/2}}\right].$$

Near $d = 2$ the quantity in brackets is proportional to $1 - d/2$, and this factor cancels the pole in the gamma function. Thus the worst divergence cancels, leaving only a pole at $d = 4$. Using the identity $\Gamma(x) = \Gamma(x+1)/x$, we can rewrite the above expression as

$$2i\lambda\delta^{ij}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}}\frac{1}{1-d/2}\left[\frac{\zeta^2}{(\zeta^2)^{2-d/2}} - \frac{2\mu^2}{(2\mu^2)^{2-d/2}}\right]. \tag{11.39}$$

The first term vanishes for $d > 2$ and $\zeta \to 0$, and can be neglected. Meanwhile, the second line of expression (11.38) involves the elementary integral

$$\int_0^1 dx\,(2\mu^2 x + (1-x)\zeta^2)^{\frac{d}{2}-2} = \frac{1}{d/2-1}\cdot\frac{(2\mu^2)^{d/2-1} - (\zeta^2)^{d/2-1}}{2\mu^2 - \zeta^2}.$$

This expression is also nonsingular at $d = 2$ and reduces to

$$\frac{1}{d/2-1}(2\mu^2)^{d/2-2}$$

for $d > 2$ and $\zeta \to 0$. Comparing this line with the remaining term from (11.39), and recalling that $\lambda v^2 = \mu^2$, we find that the $\pi\pi$ amplitude is not only finite, but *vanishes* completely at $p = 0$.

This result is very attractive. The $\pi\pi$ amplitude, at $p = 0$, is precisely the mass shift δm_π^2 of the π field. We already knew that the π particles are massless at tree level—they are the $N-1$ massless bosons required by Goldstone's theorem. We have now verified that these bosons remain massless at the one-loop level in the linear sigma model; in other words, the first quantum corrections to the linear sigma model also respect Goldstone's theorem. At the end of this chapter, we will give a general argument that Goldstone's theorem is satisfied to all orders in perturbation theory.

11.3 The Effective Action

In the first section of this chapter, we analyzed spontaneous symmetry breaking in classical field theory. That analysis was geometrical: We found the vacuum state by finding the deepest well in a potential surface, and we proved Goldstone's theorem by showing that symmetry required the presence of a line of degenerate minima at the bottom of the well. But this geometrical picture was lost, or at least disguised, in the one-loop calculations of Section 11.2. It seems worthwhile to develop a formalism that will allow us to use geometrical arguments about spontaneous symmetry breaking even at the quantum level.

To define our goal somewhat better, consider the problem of determining the vacuum expectation value of the quantum field ϕ. This expectation value should be determined as a function of the parameters of the Lagrangian. At the classical level, it is easy to compute $\langle\phi\rangle$; one minimizes the potential energy. However, as we have seen in the previous section, this classical value can be altered by perturbative loop corrections. In fact, we saw that $\langle\phi\rangle$ could be shifted by a potentially divergent quantity, which we needed to control by renormalization.

It would be wonderful if, in the full quantum field theory, there were a function whose minimum gave the exact value of $\langle\phi\rangle$. This function would agree with the classical potential energy to lowest order in perturbation theory, but it would be modified in higher orders by quantum corrections. Presumably, these corrections would need renormalization to remove infinities. Nevertheless, after renormalization, this quantity should give the same relations between $\langle\phi\rangle$ and particle masses and couplings that we would find by direct Feynman diagram calculations. In this section, we will exhibit a function with these properties, called the *effective potential.* In Section 11.4 we will explain how to compute the effective potential in perturbation theory, in terms of renormalized masses and couplings. Then we will go on to use it as a tool in analyzing the renormalizability of theories with hidden symmetry.

To identify the effective potential, consider the analogy between quantum field theory and statistical mechanics set out in Section 9.3. In that section, we derived a correspondence between the correlation functions of a quantum field and those of a related statistical system, with quantum fluctuations being replaced by thermal fluctuations. At zero temperature the thermodynamic ground state is the state of lowest energy, but at nonzero temperature we still have a geometrical picture of the preferred thermodynamic state: It is the state that minimizes the Gibbs free energy. More explicitly, taking the example of a magnetic system, one defines the Helmholtz free energy $F(H)$ by

$$Z(H) = e^{-\beta F(H)} = \int \mathcal{D}s \, \exp\Big[-\beta \int dx \, \big(\mathcal{H}[s] - Hs(x)\big)\Big], \tag{11.40}$$

where H is the external magnetic field, $\mathcal{H}[s]$ is the spin energy density, and $\beta = 1/kT$. We can find the magnetization of the system by differentiating

$F(H)$:

$$\begin{aligned} -\frac{\partial F}{\partial H}\bigg|_{\beta \text{ fixed}} &= \frac{1}{\beta}\frac{\partial}{\partial H}\log Z \\ &= \frac{1}{Z}\int dx \int \mathcal{D}s\, s(x) \exp\Big[-\beta \int dx\, \big(\mathcal{H}[s] - Hs\big)\Big] \\ &= \int dx\, \big\langle s(x)\big\rangle \equiv M. \end{aligned} \tag{11.41}$$

The Gibbs free energy G is defined by the Legendre transformation

$$G = F + MH,$$

so that it satisfies

$$\begin{aligned} \frac{\partial G}{\partial M} &= \frac{\partial F}{\partial M} + M\frac{\partial H}{\partial M} + H \\ &= \frac{\partial H}{\partial M}\frac{\partial F}{\partial H} + M\frac{\partial H}{\partial M} + H \\ &= H \end{aligned} \tag{11.42}$$

(where all partial derivatives are taken with β fixed). If $H = 0$, the Gibbs free energy reaches an extremum at the corresponding value of M. The thermodynamically most stable state is the minimum of $G(M)$. Thus the function $G(M)$ gives a picture of the preferred thermodynamic state that is geometrical and at the same time includes all effects of thermal fluctuations.

By analogy, we can construct a similar quantity in a quantum field theory. For simplicity, we will work in this section only with a theory of one scalar field. All of the results generalize straightforwardly to systems with multiple scalar, spinor, and vector fields.

Consider a quantum field theory of a scalar field ϕ, in the presence of an external source J. As in Chapter 9, it is useful to take the external source to depend on x. Thus, we define an energy functional $E[J]$ by

$$Z[J] = e^{-iE[J]} = \int \mathcal{D}\phi \exp\Big[i \int d^4x\, \big(\mathcal{L}[\phi] + J\phi\big)\Big]. \tag{11.43}$$

The right-hand side of this equation is the functional integral representation of the amplitude $\langle\Omega| e^{-iHT} |\Omega\rangle$, where T is the time extent of the functional integration, in the presence of the source J. Thus, $E[J]$ is just the vacuum energy as a function of the external source. The functional $E[J]$ is the analogue of the Helmholtz free energy, and J is the analogue of the external magnetic field.

In principle, we could now Legendre-transform $E[J]$ with respect to a constant value of the source. However, since we have already developed a formalism for functional integration and differentiation, it will not be much more difficult to work with an external source $J(x)$ that depends on x in an arbitrary way. As we will see, this generalization yields additional relations which connect this formalism to our general study of renormalization theory.‡

Consider, then, the functional derivative of $E[J]$ with respect to $J(x)$:

$$\frac{\delta}{\delta J(x)} E[J] = i \frac{\delta}{\delta J(x)} \log Z = -\frac{\int \mathcal{D}\phi\, e^{i\int(\mathcal{L}+J\phi)} \phi(x)}{\int \mathcal{D}\phi\, e^{i\int(\mathcal{L}+J\phi)}}. \tag{11.44}$$

We abbreviate this relation as

$$\frac{\delta}{\delta J(x)} E[J] = -\langle\Omega|\, \phi(x)\, |\Omega\rangle_J\, ; \tag{11.45}$$

the right-hand side is the vacuum expectation value in the presence of a nonzero source $J(x)$. This relation is a functional analogue of Eq. (11.41): The functional derivative of $E[J]$ gives the expectation value of ϕ in the presence of the spatially varying source. We should treat this expectation value as the thermodynamic variable conjugate to $J(x)$. Thus we define the quantity $\phi_{\rm cl}(x)$, called the *classical field*, by

$$\phi_{\rm cl}(x) = \langle\Omega|\, \phi(x)\, |\Omega\rangle_J\, . \tag{11.46}$$

The classical field is related to $\phi(x)$ in the same way that the magnetization M is related to the local spin field $s(x)$: It is a weighted average over all possible fluctuations. Note that $\phi_{\rm cl}(x)$ depends on the external source $J(x)$, just as M depends on H.

Now, in analogy with the construction of the Gibbs free energy, define the Legendre transform of $E[J]$:

$$\Gamma[\phi_{\rm cl}] \equiv -E[J] - \int d^4y\, J(y)\phi_{\rm cl}(y). \tag{11.47}$$

This quantity is known as the *effective action*. In analogy with Eq. (11.42), we can now compute

$$\begin{aligned} \frac{\delta}{\delta\phi_{\rm cl}(x)} \Gamma[\phi_{\rm cl}] &= -\frac{\delta}{\delta\phi_{\rm cl}(x)} E[J] - \int d^4y\, \frac{\delta J(y)}{\delta\phi_{\rm cl}(x)} \phi_{\rm cl}(y) - J(x) \\ &= -\int d^4y\, \frac{\delta J(y)}{\delta\phi_{\rm cl}(x)} \frac{\delta E[J]}{\delta J(y)} - \int d^4y\, \frac{\delta J(y)}{\delta\phi_{\rm cl}(x)} \phi_{\rm cl}(y) - J(x) \\ &= -J(x). \end{aligned} \tag{11.48}$$

In the last step we have used Eq. (11.45).

For each of the thermodynamic quantities discussed at the beginning of this section, we have now defined an analogous quantity in quantum field theory. Table 11.1 summarizes these analogies.

‡This functional generalization of thermodynamics is due to C. DeDominicis and P. Martin, *J. Math. Phys.* **5**, 14 (1964), and was formulated for relativistic field theory by G. Jona-Lasinio, *Nuovo Cim.* **34A**, 1790 (1964).

Magnetic System	Quantum Field Theory
$\mathbf{x}$	$x = (t, \mathbf{x})$
$s(\mathbf{x})$	$\phi(x)$
H	$J(x)$
$\mathcal{H}(s)$	$\mathcal{L}(\phi)$
$Z(H)$	$Z[J]$
$F(H)$	$E[J]$
M	$\phi_{\text{cl}}(x)$
$G(M)$	$-\Gamma[\phi_{\text{cl}}]$

Table 11.1. Analogous quantities in a magnetic system and a scalar quantum field theory.

Relation (11.48) implies that, if the external source is set to zero, the effective action satisfies the equation

$$\frac{\delta}{\delta\phi_{\text{cl}}(x)}\Gamma[\phi_{\text{cl}}] = 0. \tag{11.49}$$

The solutions to this equation are the values of $\langle\phi(x)\rangle$ in the stable quantum states of the theory. For a translation-invariant vacuum state, we will find a solution in which ϕ_{cl} is independent of x. Sometimes, Eq. (11.49) will have additional solutions, corresponding to localized lumps of field held together by their self-interaction. In these states, called *solitons*, the solution $\phi_{\text{cl}}(x)$ depends on x.

From here on we will assume, for the field theories we consider, that the possible vacuum states are invariant under translations and Lorentz transformations.* Then, for each possible vacuum state, the corresponding solution ϕ_{cl} will be a constant, independent of x, and the process of solving Eq. (11.49) reduces to that of solving an ordinary equation of one variable (ϕ_{cl}). Furthermore, we know that Γ is, in thermodynamic terms, an extensive quantity: It is proportional to the volume of the spacetime region over which the functional integral is taken. If T is the time extent of this region and V is its three-dimensional volume, we can write

$$\Gamma[\phi_{\text{cl}}] = -(VT)\cdot V_{\text{eff}}(\phi_{\text{cl}}). \tag{11.50}$$

The coefficient V_{eff} is called the *effective potential.* The condition that $\Gamma[\phi_{\text{cl}}]$ has an extremum then reduces to the simple equation

$$\frac{\partial}{\partial\phi_{\text{cl}}}V_{\text{eff}}(\phi_{\text{cl}}) = 0. \tag{11.51}$$

*Certain condensed matter systems have ground states with preferred orientation; see, for example P. G. de Gennes, *The Physics of Liquid Crystals* (Oxford University Press, 1974).

Each solution of Eq. (11.51) is a translation-invariant state with $J = 0$. Equation (11.47) implies that $\Gamma = -E$ in this case, and therefore that $V_{\text{eff}}(\phi_{\text{cl}})$, evaluated at a solution to (11.51), is just the energy density of the corresponding state.

Figure 11.6 illustrates one possible shape for the function $V_{\text{eff}}(\phi)$. The local maxima (or, for systems of several fields ϕ^i, possible saddle points) are unstable configurations that cannot be realized as stationary states. The figure also contains a local minimum of V_{eff} that is not the absolute minimum; this is a metastable vacuum state, which can decay to the true vacuum by quantum-mechanical tunneling. The absolute minimum of V_{eff} is the state of lowest energy in the theory, and thus the true, stable, vacuum state. A system with spontaneously broken symmetry will have several minima of V_{eff}, all with the same energy by virtue of the symmetry. The choice of one among these vacua is the spontaneous symmetry breaking.

In drawing Fig. 11.6, we have assumed that we are computing the effective potential for a fixed constant background value of ϕ. Under some circumstances, this state does not give the true minimum energy configuration for states with a given expectation value of ϕ. This mismatch can occur in the following way: In a system for which the effective potential for constant background fields is given by Fig. 11.6, consider choosing a value of ϕ_{cl} that is intermediate between the locally stable vacuum states ϕ_1 and ϕ_3:

$$\phi_{\text{cl}} = x\phi_1 + (1-x)\phi_3, \qquad 0 < x < 1. \tag{11.52}$$

The assumption of a constant background field gives a large value of the effective potential, as indicated in the figure. We can obtain a lower-energy configuration by considering states with macroscopic regions in which $\langle\phi\rangle = \phi_1$ and other regions in which $\langle\phi\rangle = \phi_3$, in such a way that the average value of $\langle\phi\rangle$ over the whole system is ϕ_{cl}. For such a configuration, the average vacuum energy is given by

$$V_{\text{eff}}(\phi_{\text{cl}}) = xV_{\text{eff}}(\phi_1) + (1-x)V_{\text{eff}}(\phi_3), \tag{11.53}$$

as shown in Fig. 11.7. We have called the left-hand side of this equation $V_{\text{eff}}(\phi_{\text{cl}})$ because the result (11.53) would be the result of an exact evaluation of the functional integral definition of V_{eff} for values of ϕ_{cl} satisfying (11.52). The interpolation (11.53) is the field theoretic analogue of the Maxwell construction for the thermodynamic free energy. In general, for any ϕ_{cl}, ϕ_1, ϕ_3 satisfying (11.52), the estimate (11.53) will be an upper bound to the effective potential; we say that the effective potential is a *convex* function of ϕ_{cl}.[†]

Just as in thermodynamics, straightforward schemes for computing the effective potential do not take account of the possibility of phase separation and so lead to a structure of unstable and metastable configurations of the

[†]The convexity of the Gibbs free energy is a well-known exact result in statistical mechanics; see, for example, D. Ruelle, *Statistical Mechanics* (W. A. Benjamin, Reading, Mass., 1969).

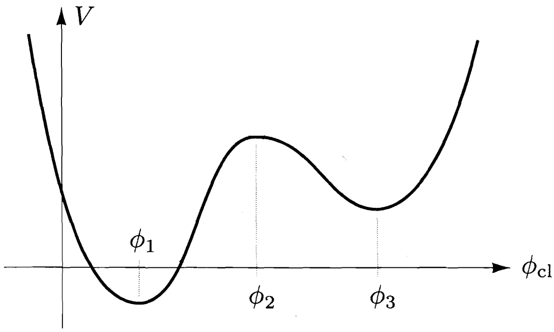


Figure 11.6. A possible form for the effective potential in a scalar field theory. The extrema of the effective potential occur at the points $\phi_{cl} = \phi_1, \phi_2, \phi_3$. The true vacuum state is the one corresponding to ϕ_1. The state ϕ_2 is unstable. The state ϕ_3 is metastable, but it can decay to ϕ_1 by quantum-mechanical tunneling.

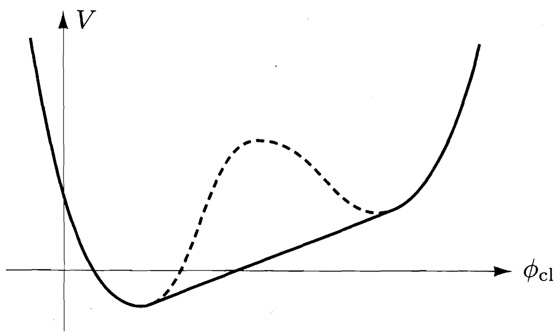


Figure 11.7. Exact convex form of the effective potential for the system of Fig. 11.6.

type shown in Fig. 11.6. The Maxwell construction must be performed by hand to yield the final form of $V_{\text{eff}}(\phi_{cl})$. Fortunately, the absolute minimum of V_{eff} is not affected by this nicety.

We have now solved the problem that we posed at the beginning of this section: The effective potential, defined by Eqs. (11.47) and (11.50), gives an easily visualized function whose minimization defines the exact vacuum state of the quantum field theory, including all effects of quantum corrections. It is not obvious from these definitions how to compute $V_{\text{eff}}(\phi_{cl})$. We will see how to do so in the next section, by direct evaluation of the functional integral.

11.4 Computation of the Effective Action

Now that we have defined the object whose minimization gives the exact vacuum state of a quantum field theory, we must learn how to compute it. This can be done in more than one way. The simplest method, which we will use here, requires that we be bold enough to evaluate the complete effective action Γ directly from its functional integral definition. After computing Γ, we can obtain V_{eff} by specializing to constant values of ϕ_{cl}.‡

Our plan is to find a perturbation expansion for the generating functional Z, starting with its functional integral definition (11.43). We will then take the logarithm to obtain the energy functional E, and finally Legendre-transform according to Eq. (11.47) to obtain Γ. We will use renormalized perturbation theory, so it is convenient to split the Lagrangian as we did in Eq. (10.18), into a piece depending on renormalized parameters and one containing the counterterms:

$$\mathcal{L} = \mathcal{L}_1 + \delta\mathcal{L}. \tag{11.54}$$

We wish to compute Γ as a function of ϕ_{cl}. But the functional $Z[J]$ depends on ϕ_{cl} through its dependence on J. Thus, we must find, at least implicitly, a relation between $J(x)$ and $\phi_{\text{cl}}(x)$. At the lowest order in perturbation theory, that relation is just the classical field equation:

$$\left.\frac{\delta\mathcal{L}}{\delta\phi}\right|_{\phi=\phi_{\text{cl}}} + J(x) = 0 \qquad \text{(to lowest order).}$$

Let us define $J_1(x)$ to be whatever function satisfies this equation exactly, when $\mathcal{L} = \mathcal{L}_1$:

$$\left.\frac{\delta\mathcal{L}_1}{\delta\phi}\right|_{\phi=\phi_{\text{cl}}} + J_1(x) = 0 \qquad \text{(exactly).} \tag{11.55}$$

We will think of the difference between J and J_1 as a counterterm, analogous to $\delta\mathcal{L}$, so we write

$$J(x) = J_1(x) + \delta J(x), \tag{11.56}$$

where δJ is determined, order by order in perturbation theory, by the original definition (11.46) of ϕ_{cl}, namely $\langle\phi(x)\rangle_J = \phi_{\text{cl}}(x)$.

Using this notation, we rewrite Eq. (11.43) as

$$e^{-iE[J]} = \int \mathcal{D}\phi\, e^{i\int d^4x(\mathcal{L}_1[\phi]+J_1\phi)} e^{i\int d^4x(\delta\mathcal{L}[\phi]+\delta J\phi)}. \tag{11.57}$$

The second exponential contains the counterterms; leave this aside for the moment. In the first exponential, expand the exponent about ϕ_{cl} by replacing

‡This method is due to R. Jackiw, *Phys. Rev.* **D9**, 1686 (1974).

$\phi(x) = \phi_{\text{cl}}(x) + \eta(x)$. This exponent takes the form

$$\int d^4x(\mathcal{L}_1 + J_1\phi) = \int d^4x(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) + \int d^4x\,\eta(x)\Big(\frac{\delta\mathcal{L}_1}{\delta\phi} + J_1\Big)$$
$$+ \frac{1}{2}\int d^4x\,d^4y\,\eta(x)\eta(y)\frac{\delta^2\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)}$$
$$+ \frac{1}{3!}\int d^4x\,d^4y\,d^4z\,\eta(x)\eta(y)\eta(z)\frac{\delta^3\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)\delta\phi(z)} + \cdots, \tag{11.58}$$

where the various functional derivatives of $\mathcal{L}_1$ are evaluated at $\phi_{\text{cl}}(x)$. Notice that the term linear in η vanishes by the use of Eq. (11.55). The integral over η is thus a Gaussian integral, with the cubic and higher terms giving perturbative corrections.

We will describe a formal evaluation of this integral, following the prescriptions of Section 9.2. The ingredients in this evaluation will be the coefficients of Eq. (11.58), that is, the successive functional derivatives of $\mathcal{L}_1$. For the moment, please accept that these give well-defined operators. After presenting a general expression for $\Gamma[\phi_{\text{cl}}]$, we will carry out this calculation explicitly in a scalar field theory example. We will see in this example that the formal operators correspond to expressions familiar from Feynman diagram perturbation theory.

Let us, then, consider performing the integral over $\eta(x)$ using the expansion (11.58). Keeping only the terms up to quadratic order in η, and still neglecting the counterterms, we have a pure Gaussian integral, which can be evaluated in terms of a functional determinant:

$$\int \mathcal{D}\eta\,\exp\Big[i\Big(\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) + \frac{1}{2}\int \eta\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\eta\Big)\Big]$$
$$= \exp\Big[i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}})\Big]\cdot\Big(\det\Big[-\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\Big]\Big)^{-1/2}. \tag{11.59}$$

This functional determinant will give us the lowest-order quantum correction to the effective action, and for many purposes it is unnecessary to go further in the expansion (11.58). Later we will see that if we do include the cubic and higher terms in η, these produce a Feynman diagram expansion of the functional integral (11.57) in which the propagator is the operator inverse

$$-i\Big(\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\Big)^{-1} \tag{11.60}$$

and the vertices are the third and higher functional derivatives of $\mathcal{L}_1$.

Finally, let us put back the effects of the second exponential in Eq. (11.57), that is, the counterterm Lagrangian. It is useful to expand this term about $\phi = \phi_{\text{cl}}$, writing it as

$$\big(\delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}}\big) + \big(\delta\mathcal{L}[\phi_{\text{cl}} + \eta] - \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\eta\big). \tag{11.61}$$

The second term of (11.61) can be expanded as a Taylor series in η; the successive terms give counterterm vertices which can be included in the aforementioned Feynman diagrams. The first term is a constant with respect to the functional integral over η, and therefore gives additional terms in the exponent of Eq. (11.59).

Combining the integral (11.59) with the contributions from higher-order vertices and counterterms, one can obtain a complete expression for the functional integral (11.57). We will see in the example below that the Feynman diagrams representing the higher-order terms can be arranged to give the exponential of the sum of connected diagrams. Thus one obtains the following expression for $E[J]$:

$$\begin{aligned} -iE[J] = i\int d^4x(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) - \frac{1}{2}\log\det\left[-\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right] \\ + (\text{connected diagrams}) + i\int d^4x(\delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}}). \end{aligned} \tag{11.62}$$

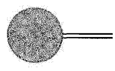
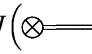
From this equation, Γ follows directly: Using $J_1 + \delta J = J$ and the Legendre transform (11.47), we find

$$\begin{aligned} \Gamma[\phi_{\text{cl}}] = \int d^4x\, \mathcal{L}_1[\phi_{\text{cl}}] + \frac{i}{2}\log\det\left[-\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right] \\ - i\cdot(\text{connected diagrams}) + \int d^4x\, \delta\mathcal{L}[\phi_{\text{cl}}]. \end{aligned} \tag{11.63}$$

Notice that there are no terms remaining that depend explicitly on J; thus, Γ is expressed as a function of ϕ_{cl}, as it should be. The Feynman diagrams contributing to $\Gamma[\phi_{\text{cl}}]$ have no external lines, and the simplest ones turn out to have two loops. The lowest-order quantum correction to Γ is given by the functional determinant, and this term is all that we will make use of in this book.

The last term of (11.63) provides a set of counterterms that can be used to satisfy the renormalization conditions on Γ and, in the process, to cancel divergences that appear in the evaluation of the functional determinant and the diagrams. We will show in the example below exactly how this cancellation works. The renormalization conditions will determine all of the counterterms in $\delta\mathcal{L}$. However, the formalism we have constructed contains a new counterterm δJ. That coefficient is determined by the following special criterion: In Eq. (11.55), we set up our analysis in such a way that, at the leading order, $\langle\phi\rangle = \phi_{\text{cl}}$. Potentially, however, this relation could break down at higher orders: The quantity $\langle\phi\rangle$ could receive additional contributions from Feynman diagrams that might shift it from the value ϕ_{cl}. This will happen if there are nonzero tadpole diagrams that contribute to $\langle\eta\rangle$. But this amplitude also receives a contribution from the counterterm $(\delta J\eta)$ in (11.61). Thus we can maintain $\langle\eta\rangle = 0$, and in the process determine δJ to any order, by adjusting

δJ to satisfy the diagrammatic equation

 $+ \ \delta J\Big($  $\Big) \ = \ 0\,.$ (11.64)

In practice, we will satisfy this condition by simply ignoring any one-particle-irreducible one-point diagram, since any such diagram will be canceled by adjustment of δJ. The removal of these tadpole diagrams, which we needed some effort to arrange in Section 11.2, is thus built in here as a natural part of the formalism.

The Effective Action in the Linear Sigma Model

In Eq. (11.63), we have given a complete, though not exactly transparent, evaluation of $\Gamma[\phi_{\rm cl}]$. Let us now clarify the meaning of this equation, and also put it to some good use, by computing $\Gamma[\phi_{\rm cl}]$ in the linear sigma model. We will see that the results that we obtained by brute-force perturbation theory in Section 11.2 emerge much more naturally from Eq. (11.63).

We begin again with the Lagrangian (11.5):

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi^i)^2 + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}\big[(\phi^i)^2\big]^2. \tag{11.65}$$

Expand about the classical field: $\phi^i = \phi^i_{\rm cl} + \eta^i$. Because we expect to find a translation-invariant vacuum state, we will specialize to the case of a *constant* classical field. This will simplify some elements of the calculation below. In particular, according to Eq. (11.50), the final result will be proportional to the four-dimensional volume (VT) of the functional intergration. When this dependence is factored out, we will obtain a well-defined intensive expression for the effective potential. In any event, after this simplification, (11.65) takes the form

$$\begin{aligned}\mathcal{L} = \frac{1}{2}\mu^2(\phi^i_{\rm cl})^2 - \frac{\lambda}{4}\big[(\phi^i_{\rm cl})^2\big]^2 + (\mu^2 - \lambda(\phi^i_{\rm cl})^2)\phi^i_{\rm cl}\eta^i \\ + \frac{1}{2}(\partial_\mu\eta^i)^2 + \frac{1}{2}\mu^2(\eta^i)^2 - \frac{\lambda}{2}\big[(\phi^i_{\rm cl})^2(\eta^i)^2 + 2(\phi^i_{\rm cl}\eta^i)^2\big] + \cdots.\end{aligned} \tag{11.66}$$

According to Eq. (11.63), we should drop the term linear in η.

From the terms quadratic in η, we can read off

$$\frac{\delta^2\mathcal{L}}{\delta\phi^i\delta\phi^j} = -\partial^2\delta^{ij} + \mu^2\delta^{ij} - \lambda\big[(\phi^k_{\rm cl})^2\delta^{ij} + 2\phi^i_{\rm cl}\phi^j_{\rm cl}\big]. \tag{11.67}$$

Notice that this object has the general form of a Klein-Gordon operator. To clarify this relation, let us orient the coordinates so that $\phi^i_{\rm cl}$ points in the Nth direction,

$$\phi^i_{\rm cl} = (0, 0, \ldots, 0, \phi_{\rm cl}), \tag{11.68}$$

as we did in Eq. (11.7). Then the operator (11.67) is just equal to the Klein-Gordon operator $(-\partial^2 - m_i^2)$, where

$$m_i^2 = \begin{cases} \lambda\phi_{\rm cl}^2 - \mu^2 & \text{acting on } \eta^1, \ldots, \eta^{N-1}; \\ 3\lambda\phi_{\rm cl}^2 - \mu^2 & \text{acting on } \eta^N. \end{cases} \tag{11.69}$$

The functional determinant in Eq. (11.63) is the product of the determinants of these Klein-Gordon operators:

$$\det \frac{\delta^2 \mathcal{L}}{\delta\phi\delta\phi} = \left[\det(\partial^2 + (\lambda\phi_{\rm cl}^2 - \mu^2))\right]^{N-1}\left[\det(\partial^2 + (3\lambda\phi_{\rm cl}^2 - \mu^2))\right]. \tag{11.70}$$

It is not difficult to obtain an explicit form for the determinant of a Klein-Gordon operator. To begin, use the trick of Eq. (9.77) to write

$$\log\det(\partial^2 + m^2) = \operatorname{Tr}\log(\partial^2 + m^2).$$

Now evaluate the trace of the operator as the sum of its eigenvalues:

$$\begin{aligned} \operatorname{Tr}\log(\partial^2 + m^2) &= \sum_k \log(-k^2 + m^2) \\ &= (VT)\int \frac{d^4k}{(2\pi)^4}\log(-k^2 + m^2). \end{aligned} \tag{11.71}$$

In the second line, we have converted the sum over momenta to an integral. The factor (VT) is the four-dimensional volume of the functional integral; we have already noted that this is expected to appear as an overall factor in $\Gamma[\phi_{\rm cl}]$. This manipulation gives an integral that can be evaluated in dimensional regularization after a Wick rotation:

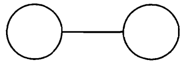
$$\begin{aligned} \int \frac{d^dk}{(2\pi)^d}\log(-k^2 + m^2) &= i\int \frac{d^dk_E}{(2\pi)^d}\log(k_E^2 + m^2) \\ &= -i\frac{\partial}{\partial\alpha}\int \frac{d^dk_E}{(2\pi)^d}\frac{1}{(k_E^2 + m^2)^\alpha}\bigg|_{\alpha=0} \\ &= -i\frac{\partial}{\partial\alpha}\left(\frac{1}{(4\pi)^{d/2}}\frac{\Gamma(\alpha - \frac{d}{2})}{\Gamma(\alpha)}\frac{1}{(m^2)^{\alpha-d/2}}\right)\bigg|_{\alpha=0} \\ &= -i\frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}}\frac{1}{(m^2)^{-d/2}}. \end{aligned} \tag{11.72}$$

In the last line, we have used $\Gamma(\alpha) \to 1/\alpha$ as $\alpha \to 0$. Thus,

$$\frac{1}{(VT)}\log\det(\partial^2 + m^2) = -i\frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}}(m^2)^{d/2}. \tag{11.73}$$

Using this result to evaluate the determinant in Eq. (11.63), and choosing

(a)



(b)

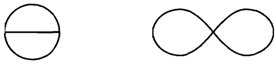


Figure 11.8. Feynman diagrams contributing to the evaluation of the effective potential of the $O(N)$ linear sigma model: (a) a diagram that is removed by (11.64); (b) the first nonzero diagrammatic corrections.

the counterterm Lagrangian as in Eq. (11.14), we find

$$\begin{aligned} V_{\text{eff}}(\phi_{\text{cl}}) &= -\frac{1}{(VT)}\Gamma[\phi_{\text{cl}}] \\ &= -\frac{1}{2}\mu^2\phi_{\text{cl}}^2 + \frac{\lambda}{4}\phi_{\text{cl}}^4 \\ &\quad - \frac{1}{2}\frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}}\left[(N-1)(\lambda\phi_{\text{cl}}^2 - \mu^2)^{d/2} + (3\lambda\phi_{\text{cl}}^2 - \mu^2)^{d/2}\right] \\ &\quad + \frac{1}{2}\delta_\mu\phi_{\text{cl}}^2 + \frac{1}{4}\delta_\lambda\phi_{\text{cl}}^4. \end{aligned} \tag{11.74}$$

Here we have written ϕ_{cl}^2 as a shorthand for $(\phi_{\text{cl}}^i)^2$. Since the second line of this result is the leading radiative correction, we might expect that the result has the structure of a one-loop Feynman diagram. Indeed, we see that this expression contains Gamma functions and ultraviolet divergences similar to those that we found in the one-loop computations of Section 11.2. We will show below that this term in fact has exactly the same ultraviolet divergences that we found in Section 11.2. These divergences will be subtracted by the counterterms in the last line of Eq. (11.74).

Since the computation of the determinant in Eq. (11.63) gives the effect of one-loop corrections, we might expect the Feynman diagrams that contribute to Eq. (11.63) to begin in two-loop order. We can see this explicitly for the case of the $O(N)$ sigma model. The perturbation expansion described below Eq. (11.60) involves the propagator that is the inverse of Eq. (11.67):

$$\langle \eta^i(k)\eta^j(-k)\rangle = \frac{i}{k^2 - m_i^2}\delta^{ij}, \tag{11.75}$$

where m_i^2 is given by (11.69). The vertices are given by the terms of order η^3 and η^4 in the expansion of the Lagrangian. Combining these ingredients, we find that the leading Feynman diagrams contributing to the vacuum energy have the forms shown in Fig. 11.8. The diagram of Fig. 11.8(a) is actually canceled by the effects of the counterterm δJ, as shown in Eq. (11.64). Thus the leading diagrammatic contribution to the effective potential comes from the two-loop diagrams of Fig. 11.8(b).

The result (11.74) is manifestly $O(N)$-symmetric. From the question that we posed at the beginning of Section 11.2, we might have feared that this property would be destroyed when we compute radiative corrections about a state with spontaneously broken symmetry. But $V_{\text{eff}}(\phi_{\text{cl}})$ is the function that we minimize to find the vacuum state, and so it should properly be symmetric, even if the lowest-energy vacuum is asymmetric. In the formalism we have constructed here, there is no need to worry. Formula (11.63) is manifestly invariant, term by term, under the original $O(N)$ symmetry of the Lagrangian. Thus we must necessarily have arrived at an $O(N)$-symmetric result for $V_{\text{eff}}(\phi_{\text{cl}})$.

Before going on to determine δ_μ and δ_λ precisely, we might first check that the counterterms in Eq. (11.74) are sufficient to make the expression for $\Gamma[\phi_{\text{cl}}]$ finite. The factor $\Gamma(-d/2)$ has poles at $d = 0, 2, 4$. The pole at $d = 0$ is a constant, independent of ϕ_{cl}, and therefore without physical significance. The pole at $d = 2$ is an even quadratic polynomial in ϕ_{cl}. The pole at $d = 4$ is an even quartic polynomial in ϕ_{cl}. Thus Eq. (11.74) becomes a finite expression in the limit $d \to 2$ if we set

$$\delta_\mu = -\lambda(N+2)\frac{\Gamma(1-\frac{d}{2})}{(4\pi)} + \text{ finite.}$$

The expression is finite as $d \to 4$ if we set

$$\begin{aligned}\delta_\mu &= -\lambda\mu^2(N+2)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} + \text{ finite;}\\ \delta_\lambda &= \lambda^2(N+8)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} + \text{ finite.}\end{aligned} \tag{11.76}$$

These expressions agree with our earlier results from Section 11.2, Eqs. (11.33) and (11.22), in the limits $d \to 2$ and $d \to 4$ respectively.

The finite parts of δ_λ and δ_μ depend on the exact form of the renormalization conditions that are imposed. For example, in Section 11.2, we imposed the condition (11.16) that the vacuum expectation value of ϕ equals $\mu/\sqrt{\lambda}$ and the additional conditions in (11.17) on the scattering amplitude and field strength of the σ. Condition (11.16) is readily expressed in terms of the effective potential as

$$\frac{\partial V_{\text{eff}}}{\partial \phi_{\text{cl}}}(\phi_{\text{cl}} = \mu/\sqrt{\lambda}) = 0.$$

Using the connection between derivatives of Γ and one-particle-irreducible amplitudes, we could write the other two conditions as Fourier transforms to momentum space of functional derivatives of $\Gamma[\phi_{\text{cl}}]$. In this way, it is possible in principle to reconstruct the particular renormalization scheme used in Section 11.2.

However, if we want to visualize the modification of the lowest-order results that is induced by the quantum corrections, we can apply a renormalization scheme that can be implemented more easily. One such scheme, known as *minimal subtraction* (MS), is simply to remove the $(1/\epsilon)$ poles (for $\epsilon = 4-d$) in potentially divergent quantities. Normally, though, these $(1/\epsilon)$ poles are accompanied by terms involving γ and $\log(4\pi)$. It is convenient, and no more arbitrary, to subtract these terms as well. In this prescription, known as *modified minimal subtraction* or $\overline{MS}$ ("em-ess-bar"), one replaces

$$\begin{aligned}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}(m^2)^{2-d/2}} &= \frac{1}{(4\pi)^2}\Big(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log(m^2)\Big) \\ &\longrightarrow \frac{1}{(4\pi)^2}\big(-\log(m^2/M^2)\big), \qquad (11.77)\end{aligned}$$

where M is an arbitrary mass parameter that we have introduced to make the final equation dimensionally correct. You should think of M as parametrizing a sequence of possible renormalization conditions. The $\overline{MS}$ renormalization scheme usually puts one-loop corrections in an especially simple form. The price of this simplicity is that it normally takes some effort to express physically measurable quantities in terms of the parameters of the $\overline{MS}$ expression.

To apply the $\overline{MS}$ renormalization prescription to (11.74), we need to expand the divergent terms in this equation in powers of ϵ. As an example, consider the $\overline{MS}$ regularization of expression (11.73):

$$\begin{aligned}\frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}}(m^2)^{d/2} &= \frac{1}{\frac{d}{2}(\frac{d}{2}-1)}\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}}(m^2)^{d/2} \\ &= \frac{m^4}{2(4\pi)^2}\Big(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log(m^2) + \frac{3}{2}\Big) \\ &\longrightarrow \frac{m^4}{2(4\pi)^2}\Big(-\log(m^2/M^2) + \frac{3}{2}\Big). \qquad (11.78)\end{aligned}$$

Modifying our result (11.74) in this way, we find

$$\begin{aligned}V_{\text{eff}} = &-\frac{1}{2}\mu^2\phi_{\text{cl}}^2 + \frac{\lambda}{4}\phi_{\text{cl}}^4 \\ &+ \frac{1}{4}\frac{1}{(4\pi)^2}\Big((N-1)(\lambda\phi_{\text{cl}}^2-\mu^2)^2\big(\log\big[(\lambda\phi_{\text{cl}}^2-\mu^2)/M^2\big] - \tfrac{3}{2}\big) \\ &\qquad + (3\lambda\phi_{\text{cl}}^2-\mu^2)^2\big(\log\big[(3\lambda\phi_{\text{cl}}^2-\mu^2)/M^2\big] - \tfrac{3}{2}\big)\Big). \qquad (11.79)\end{aligned}$$

The effective potential is thus modified to be slightly steeper at large values of ϕ_{cl} and more negative at smaller values, as shown in Fig. 11.9. For each set of values of μ, λ, and M, we can determine the preferred vacuum state by minimizing $V_{\text{eff}}(\phi)$ with respect to ϕ_{cl}. The correction to V_{eff} is undefined when the arguments of the logarithms become negative, but fortunately the minima of V_{eff} occur outside of this region, as is illustrated in the figure.

Before going on, we would like to raise two questions about this expression for the effective potential. The problems that we will raise occur generically in quantum field theory calculations, but expression (11.79) provides a concrete

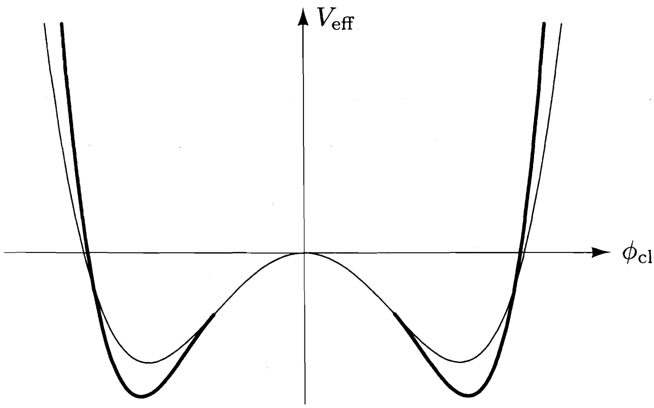


Figure 11.9. The effective potential for ϕ^4 theory ($N = 1$), with quantum corrections included as in Eq. (11.79). The lighter-weight curve shows the classical potential energy, for comparison.

illustration of these difficulties. Most of our discussion in the next two chapters will be devoted to building a formalism within which these questions can be answered.

First, it is troubling that, while our classical Lagrangian contained only two parameters, μ and λ, the result (11.79) depends on three parameters, of which one is the arbitrary mass scale M. A superficial reply to this complaint can be given as follows: Consider the change in $V_{\text{eff}}(\phi_{\text{cl}})$ that results from changing the value of M^2 to $M^2 + \delta M^2$. From the explicit form of (11.79), we can see that this change is compensated completely by shifting the values of μ and λ, according to

$$\begin{aligned} \lambda &\to \lambda + \frac{\lambda^2}{(4\pi)^2}(N+8) \cdot \frac{\delta M^2}{M^2}, \\ \mu^2 &\to \mu^2 + \frac{\lambda\mu^2}{(4\pi)^2}(N+2) \cdot \frac{\delta M^2}{M^2}. \end{aligned} \tag{11.80}$$

Thus, a change in M^2 is completely equivalent to changes in the parameters μ and λ. It is not clear, however, why this should be true or how this fact helps us understand the dependence of our formulae on M^2.

The second problem arises from the fact that the one-loop correction in Eq. (11.79) includes a logarithm that can become large enough to compensate the small coupling constant λ. The problem is particularly clear in the limit $\mu^2 \to 0$; then Eq. (11.79) takes the form

$$\begin{aligned} V_{\text{eff}} &= \frac{\lambda}{4}\phi_{\text{cl}}^4 + \frac{1}{4}\frac{\lambda^2}{(4\pi)^2}\phi_{\text{cl}}^4\Big((N+8)\big(\log(\lambda\phi_{\text{cl}}^2/M^2) - \tfrac{3}{2}\big) + 9\log 3\Big) \\ &= \frac{1}{4}\phi_{\text{cl}}^4\Big[\lambda + \frac{\lambda^2}{(4\pi)^2}\Big((N+8)\big(\log(\lambda\phi_{\text{cl}}^2/M^2) - \tfrac{3}{2}\big) + 9\log 3\Big)\Big]. \end{aligned} \tag{11.81}$$

Where is the minimum of this potential? If we take this expression at face value, we find that $V_{\text{eff}}(\phi_{\text{cl}})$ passes through zero when ϕ_{cl} reaches a very small value of order

$$\phi_{\text{cl}}^2 \sim \frac{M^2}{\lambda} \cdot \exp\left[-\frac{(4\pi)^2}{(N+8)\lambda}\right],$$

and, near this point, attains a minimum with a nonzero value of ϕ_{cl}. But the zero occurs by the cancellation of the leading term against the quantum correction. In other words, perturbation theory breaks down completely before we can address the question of whether $V_{\text{eff}}(\phi_{\text{cl}})$, for $\mu^2 = 0$, has a symmetry-breaking minimum. It seems that our present tools are quite inadequate to resolve this case.

Although it is far from obvious, these two problems turn out to be related to each other. One of our major results in Chapter 12 will be an explanation of the interrelation of M^2, λ, and μ^2 displayed in Eq. (11.80). Then, in Chapter 13, we will use the insight we have gained from this analysis to solve completely the second problem of the appearance of large logarithms. Before beginning that study, however, there are a few issues we have yet to discuss in the more formal aspects of the renormalization of theories with spontaneously broken symmetry.

11.5 The Effective Action as a Generating Functional

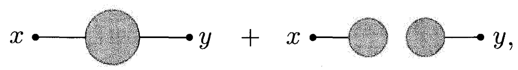
Now that we have defined the effective action and computed it for one particular theory, let us return to our goal of understanding the renormalization of theories with hidden symmetry. In Section 11.6 we will use the effective action as a tool in achieving this goal. First, however, we must investigate in more detail the relation between the effective action and Feynman diagrams.

We saw in Section 9.2 that the functional derivatives of $Z[J]$ with respect to $J(x)$ produce the correlation functions of the scalar field (see, for example, Eq. (9.35)). In other words, $Z[J]$ is the *generating functional* of correlation functions. Our goal now is to show that $\Gamma[\phi_{\text{cl}}]$ is also such a generating functional; specifically, it is the generating functional of one-particle-irreducible (1PI) correlation functions. Since the 1PI correlation functions figure prominently in the theory of renormalization, this result will be central in the discussion of renormalization in the following section.

To begin, let us consider the functional derivatives not of $\Gamma[\phi_{\text{cl}}]$, but of $E[J] = i \log Z[J]$. The first derivative, given in Eq. (11.44), is precisely $-\langle\phi(x)\rangle$. The second derivative is

$$\begin{aligned}\frac{\delta^2 E[J]}{\delta J(x)\delta J(y)} &= -\frac{i}{Z}\int \mathcal{D}\phi\, e^{i\int(\mathcal{L}+J\phi)}\phi(x)\phi(y) \\ &\qquad + \frac{i}{Z^2}\int \mathcal{D}\phi\, e^{i\int(\mathcal{L}+J\phi)}\phi(x)\cdot\int\mathcal{D}\phi\, e^{i\int(\mathcal{L}+J\phi)}\phi(y) \\ &= -i\Big[\langle\phi(x)\phi(y)\rangle - \langle\phi(x)\rangle\langle\phi(y)\rangle\Big]. \end{aligned} \tag{11.82}$$

If we were to compute the term $\langle \phi(x)\phi(y) \rangle$ from Feynman diagrams, there would be two types of contributions:



(11.83)

where each circle corresponds to a sum of *connected* diagrams. The second term in the last line of Eq. (11.82) cancels the second, disconnected, term of (11.83). Thus the second derivative of $E[J]$ contains only those contributions to $\langle \phi(x)\phi(y) \rangle$ that come from connected Feynman diagrams. Let us call this object the *connected correlator*:

$$\frac{\delta^2 E[J]}{\delta J(x)\delta J(y)} = -i \left\langle \phi(x)\phi(y) \right\rangle_{\text{conn}}. \tag{11.84}$$

Similarly, the third functional derivative of $E[J]$ is

$$\begin{aligned}\frac{\delta^3 E[J]}{\delta J(x)\delta J(y)\delta J(z)} &= \Big[\left\langle \phi(x)\phi(y)\phi(z) \right\rangle - \left\langle \phi(x)\phi(y) \right\rangle \left\langle \phi(z) \right\rangle - \left\langle \phi(x)\phi(z) \right\rangle \left\langle \phi(y) \right\rangle \\ &\qquad - \left\langle \phi(y)\phi(z) \right\rangle \left\langle \phi(x) \right\rangle + 2 \left\langle \phi(x) \right\rangle \left\langle \phi(y) \right\rangle \left\langle \phi(z) \right\rangle\Big] \\ &= \left\langle \phi(x)\phi(y)\phi(z) \right\rangle_{\text{conn}}. \end{aligned} \tag{11.85}$$

In each successive derivative of $E[J]$ all contributions cancel except for those from fully connected diagrams. The general formula for n derivatives is

$$\frac{\delta^n E[J]}{\delta J(x_1) \cdots \delta J(x_n)} = (i)^{n+1} \left\langle \phi(x_1) \cdots \phi(x_n) \right\rangle_{\text{conn}}. \tag{11.86}$$

We therefore refer to $E[J]$ as the *generating functional of connected correlation functions.*

So much for $E[J]$. Now what about the functional derivatives of the effective action? Consider first the derivative of Eq. (11.48) with respect to $J(y)$:

$$\frac{\delta}{\delta J(y)} \frac{\delta \Gamma}{\delta \phi_{\text{cl}}(x)} = -\delta(x-y).$$

We can rewrite the left-hand side of this equation using the chain rule, to obtain

$$\begin{aligned}\delta(x-y) &= -\int d^4z \, \frac{\delta \phi_{\text{cl}}(z)}{\delta J(y)} \frac{\delta^2 \Gamma}{\delta \phi_{\text{cl}}(z) \delta \phi_{\text{cl}}(x)} \\ &= \int d^4z \, \frac{\delta^2 E}{\delta J(y)\delta J(z)} \frac{\delta^2 \Gamma}{\delta \phi_{\text{cl}}(z) \delta \phi_{\text{cl}}(x)} \\ &= \left(\frac{\delta^2 E}{\delta J \delta J}\right)_{yz} \left(\frac{\delta^2 \Gamma}{\delta \phi_{\text{cl}} \delta \phi_{\text{cl}}}\right)_{zx}. \end{aligned} \tag{11.87}$$

In the second line we have used Eq. (11.45). The last line is an abstract representation of the second line, where we think of each of the second derivatives as an infinite-dimensional matrix, with the integral over z represented by matrix multiplication. What we have shown is that these two matrices are inverses of each other:

$$\left(\frac{\delta^2 E}{\delta J \delta J}\right) = \left(\frac{\delta^2 \Gamma}{\delta\phi_{\rm cl}\delta\phi_{\rm cl}}\right)^{-1}. \tag{11.88}$$

Now according to Eq. (11.84), the first of these matrices is $-i$ times the connected two-point function, that is, the exact propagator of the field ϕ. Let us call this propagator $D(x, y)$:

$$\left(\frac{\delta^2 E}{\delta J(x)\delta J(y)}\right) = -i\,\langle\phi(x)\phi(y)\rangle_{\rm conn} \equiv -iD(x, y). \tag{11.89}$$

We will therefore refer to the other matrix (times $-i$) as the *inverse propagator*:

$$\left(\frac{\delta^2 \Gamma}{\delta\phi_{\rm cl}(x)\delta\phi_{\rm cl}(y)}\right) = iD^{-1}(x, y). \tag{11.90}$$

This provides an interpretation, of sorts, for the second functional derivative of the effective action. This interpretation becomes more concrete if we go to momentum space. On a translation-invariant vacuum state (one with $\phi_{\rm cl}$ constant), the matrix $D(x, y)$ must be diagonal in momentum:

$$D(x, y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \widetilde{D}(p). \tag{11.91}$$

We showed in Eq. (7.43) that the momentum-space propagator $\widetilde{D}(p)$ is a geometric series in one-particle-irreducible Feynman diagrams. The Fourier transform of $D^{-1}(x, y)$ then gives the inverse propagator:

$$\widetilde{D}^{-1}(p) = -i(p^2 - m^2 - M^2(p^2)), \tag{11.92}$$

where $M^2(p^2)$ is the sum of one-particle-irreducible two-point diagrams.

To evaluate higher derivatives of the effective action we again use the chain rule,

$$\frac{\delta}{\delta J(z)} = \int d^4w\, \frac{\delta\phi_{\rm cl}(w)}{\delta J(z)} \frac{\delta}{\delta\phi_{\rm cl}(w)} = i\int d^4w\, D(z, w)\frac{\delta}{\delta\phi_{\rm cl}(w)}, \tag{11.93}$$

together with the standard rule for differentiating matrix inverses:

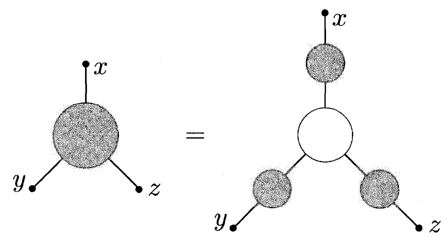
$$\frac{\partial}{\partial\alpha} M^{-1}(\alpha) = -M^{-1}\frac{\partial M}{\partial\alpha}M^{-1}. \tag{11.94}$$

Applying these identities to Eq. (11.88), we find (with some abbreviated notation)

$$\begin{aligned}\frac{\delta^3 E[J]}{\delta J_x \delta J_y \delta J_z} &= i\int d^4w\, D(z, w)\, \frac{\delta}{\delta\phi^{\rm cl}_w}\left(\frac{\delta^2\Gamma}{\delta\phi^{\rm cl}_x\delta\phi^{\rm cl}_y}\right)^{-1} \\ &= i\int d^4w\, D_{zw}\,(-1)\int d^4u\int d^4v (-iD_{xu})\frac{\delta^3\Gamma}{\delta\phi^{\rm cl}_u\delta\phi^{\rm cl}_v\delta\phi^{\rm cl}_w}(-iD_{vy})\end{aligned}$$

$$= i\int d^4u\, d^4v\, d^4w\, D_{xu}D_{yv}D_{zw}\,\frac{\delta^3\Gamma}{\delta\phi^{\rm cl}_u\delta\phi^{\rm cl}_v\delta\phi^{\rm cl}_w}. \tag{11.95}$$

This relation is more clearly expressed diagrammatically. The left-hand side is the connected three-point function. If we extract exact propagators as indicated in (11.95), this decomposes as follows:



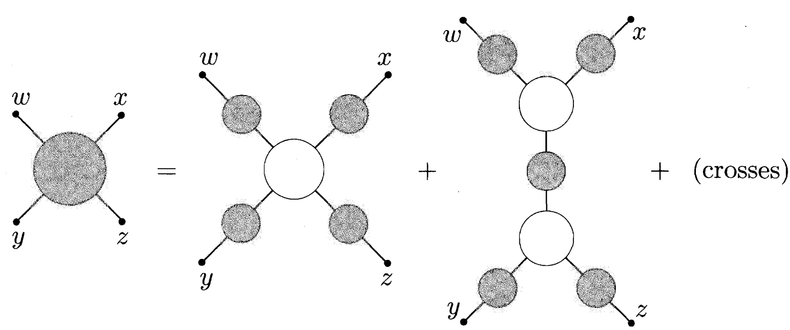
In this picture, each dark gray circle represents the sum of connected diagrams, while the light gray circle on the right-hand side represents the third derivative of $i\Gamma[\phi_{\rm cl}]$. We see that the third derivative of $i\Gamma[\phi_{\rm cl}]$ is just the connected correlation function with all three full propagators removed, that is, the *one-particle-irreducible* three-point function:

$$\frac{i\delta^3\Gamma}{\delta\phi_{\rm cl}(x)\phi_{\rm cl}(y)\phi_{\rm cl}(z)} = \langle\phi(x)\phi(y)\phi(z)\rangle_{\rm 1PI}\,.$$

By similar, if increasingly complicated, manipulations, one can derive the same relation for each successive derivative of Γ. For example, differentiating Eq. (11.95), we eventually find (using matrix notation with repeated indices implicitly integrated over)

$$\frac{-i\delta^4 E}{\delta J_w\delta J_x\delta J_y\delta J_z} = D_{sw}D_{xt}D_{yu}D_{zv}\left[\frac{i\delta^4\Gamma}{\delta\phi^{\rm cl}_s\delta\phi^{\rm cl}_t\delta\phi^{\rm cl}_u\delta\phi^{\rm cl}_v}\right.$$
$$\left.+\frac{i\delta^3\Gamma}{\delta\phi^{\rm cl}_s\delta\phi^{\rm cl}_t\delta\phi^{\rm cl}_r}D_{qr}\frac{i\delta^3\Gamma}{\delta\phi^{\rm cl}_q\delta\phi^{\rm cl}_u\delta\phi^{\rm cl}_v} + (t\leftrightarrow u) + (t\leftrightarrow v)\right].$$

Since the left-hand side of this equation is the connected four-point function, we can rewrite it diagrammatically as



As above, the dark gray circles represent the sum of connected diagrams, while the light gray circles represent i times various derivatives of Γ. Subtracting the last three terms from each side removes all one-particle reducible pieces from the connected four-point function and so identifies the fourth derivative of $i\Gamma$ as the one-particle-irreducible four-point function. The general relation (for $n \geq 3$) is

$$\frac{\delta^n \Gamma[\phi_{\rm cl}]}{\delta\phi_{\rm cl}(x_1)\cdots\delta\phi_{\rm cl}(x_n)} = -i\left\langle \phi(x_1)\cdots\phi(x_n)\right\rangle_{\rm 1PI}. \tag{11.96}$$

In other words, the effective action is the generating functional of one-particle-irreducible correlation functions.

This conclusion implies that Γ contains the complete set of physical predictions of the quantum field theory. Let us review how this information is encoded. The vacuum state of the field theory is identified as the minimum of the effective potential. The location of the minimum determines whether the symmetries of the Lagrangian are preserved or spontaneously broken. The second derivative of Γ is the inverse propagator. The poles of the propagator, or the zeros of the inverse propagator, give the values of the particle masses. Thus the particle masses m^2 are determined as the values of p^2 that solve the equation

$$i\widetilde{D}^{-1}(p^2) = \int d^4x\, e^{ip\cdot(x-y)}\frac{\delta^2\Gamma}{\delta\phi\delta\phi}(x,y) = 0. \tag{11.97}$$

The higher derivatives of Γ are the one-particle-irreducible amplitudes. These can be connected by full propagators and joined together to construct four- and higher-point connected amplitudes, which give the S-matrix elements. Thus, from the knowledge of Γ, we can reconstruct the qualitative behavior of the quantum field theory, its pattern of symmetry-breaking, and then the quantitative details of its particles and their interactions.

11.6 Renormalization and Symmetry: General Analysis

In our analysis of the divergences of quantum field theories (especially in the paragraph below Eq. (10.4)), we noted that the basic divergences of Feynman integrals are associated with one-particle-irreducible diagrams. Thus we might expect that the effective action will be a useful object in discussing the renormalizability of quantum field theories, especially those with spontaneously broken symmetry. In this section we will make use of the effective action in precisely this way.

In Section 11.4, we saw in a particular example that the formalism for calculating the effective action provides the counterterms needed to remove the ultraviolet divergences, at least at the one-loop level. These counterterms were exactly those of the original Lagrangian. We will now argue that this set of counterterms is always sufficient—to all orders and for any renormalizable field theory—by applying the power-counting arguments of Section 10.1

directly to the computation of the effective action. We will use the language of scalar field theories, but the arguments can be generalized to theories of spinor and vector fields.

Consider first the computation of the effective potential for constant (x-independent) classical fields, in a field theory with an arbitrary number of fields ϕ^i. The effective potential has mass dimension 4, so we expect that $V_{\text{eff}}(\phi_{\text{cl}})$ will have divergent terms up to Λ^4. To understand these divergences, expand $V_{\text{eff}}(\phi_{\text{cl}})$ in a Taylor series:

$$V_{\text{eff}}(\phi_{\text{cl}}) = A_0 + A_2^{ij}\phi_{\text{cl}}^i\phi_{\text{cl}}^j + A_4^{ijk\ell}\phi_{\text{cl}}^i\phi_{\text{cl}}^j\phi_{\text{cl}}^k\phi_{\text{cl}}^\ell + \cdots.$$

In theories without a symmetry $\phi^i \to -\phi^i$, there might also be terms linear and cubic in ϕ^i; we omit these for simplicity. The coefficients A_0, A_2, A_4 have mass dimension, respectively, 4, 2, and 0; thus we expect them to contain Λ^4, Λ^2, and $\log\Lambda$ divergences, respectively. The power-counting analysis predicts that all higher terms in the Taylor series expansion should be finite. The constant term A_0 is independent of ϕ_{cl}; it has no physical significance. However, the divergences in A_2 and A_4 appear in physical quantities, since these coefficients enter the inverse propagator (11.90) and the irreducible four-point function (11.96) and therefore appear in the computation of S-matrix elements. There is one further coefficient in the effective action that has non-negative mass dimension by power counting; this is the coefficient of the term quadratic in $\partial_\mu\phi_{\text{cl}}^i$, which appears when the effective action is evaluated for a nonconstant background field:

$$\Delta\Gamma[\phi_{\text{cl}}] = \int d^4x\, B_2^{ij}\,\partial_\mu\phi_{\text{cl}}^i\partial^\mu\phi_{\text{cl}}^j. \tag{11.98}$$

All other coefficients in the Taylor expansion of the effective action in powers of ϕ_{cl}^i are finite by power counting.

We can now argue that the counterterms of the original Lagrangian suffice to remove the divergences that might appear in the computation of $\Gamma[\phi_{\text{cl}}]$. The argument proceeds in two steps. We first use the BPHZ theorem to argue that the divergences of Green's functions can be removed by adjusting a set of counterterms corresponding to the possible operators that can be added to the Lagrangian with coefficients of mass dimension greater than or equal to zero. The coefficients of these counterterms are in 1-to-1 correspondence with the coefficients A_2, A_4, and B_2 of the effective action. Next, we use the fact that the effective action is manifestly invariant to the original symmetry group of the model. This is true even if the vacuum state of the model has spontaneous symmetry breaking. This symmetry of the effective action follows from the analysis of Section 11.4, since the method we presented there for computing the effective action is manifestly invariant to the original symmetry of the Lagrangian. Combining these two results, we conclude that the effective action can always be made finite by adjusting the set of counterterms that are invariant to the original symmetry of the theory, even if this symmetry is spontaneously broken. By using the results of Section 11.5, which explain how to construct the Green's functions of the theory from the functional derivatives of the effective action, this conclusion of renormalizability extends to all the Green's functions of the theory.

To make this abstract argument more concrete, we will demonstrate in a simple example how the functional derivatives of the effective action yield a set of Feynman diagrams whose divergences correspond to symmetric counterterms. Let us, then, return once again to the $O(N)$-invariant linear sigma model and compute the second functional derivative of $\Gamma[\phi_{\rm cl}]$. If the whole formalism we have constructed hangs together, we should be able to recognize the result as the Feynman diagram expansion of the inverse propagator, with divergences corresponding to the counterterms of $O(N)$-symmetric scalar field theory.

To begin, we write out expression (11.63) explicitly for this model:

$$\Gamma[\phi_{\rm cl}] = \int d^4x\Big(\frac{1}{2}(\partial_\mu^2\phi_{\rm cl}^i)^2 + \frac{1}{2}\mu^2(\phi_{\rm cl}^i)^2 - \frac{\lambda}{4}((\phi_{\rm cl}^i)^2)^2 + \frac{i}{2}\log\det[-i\mathcal{D}^{ij}] + \cdots\Big), \tag{11.99}$$

where

$$-i\mathcal{D}^{ij} = -\frac{\delta^2\mathcal{L}}{\delta\phi^i\delta\phi^j} = \partial^2\delta^{ij} + \big(\lambda(\phi_{\rm cl}^k(x))^2 - \mu^2\big)\delta^{ij} + 2\lambda\phi_{\rm cl}^i(x)\phi_{\rm cl}^j(x). \tag{11.100}$$

For constant $\phi_{\rm cl}^i$, $\mathcal{D}^{ij}$ is the operator that, acting on a given component of the scalar field, equals the Klein-Gordon operator with mass squared given by Eq. (11.69). This is the leading-order approximation to the inverse propagator of the linear sigma model.

To find the higher-order corrections to the inverse propagator, we must compute the second functional derivative of the quantum correction terms in $\Gamma[\phi_{\rm cl}]$. From (11.99), we find

$$\frac{\delta^2\Gamma}{\delta\phi_{\rm cl}^i(x)\delta\phi_{\rm cl}^j(y)} = \frac{\delta^2\mathcal{L}}{\delta\phi_{\rm cl}^i(x)\delta\phi_{\rm cl}^j(y)} + \frac{i}{2}\frac{\delta^2}{\delta\phi_{\rm cl}^i(x)\delta\phi_{\rm cl}^j(y)}\log\det[-i\mathcal{D}] + \cdots.$$

The first term is just the Klein-Gordon operator $i\mathcal{D}^{ij}\delta(x-y)$. To compute the second term, use identity (9.77) for determinants of matrices:

$$\frac{\partial}{\partial\alpha}\log\det M(\alpha) = \frac{\partial}{\partial\alpha}\operatorname{tr}\log M(\alpha) = \operatorname{tr} M^{-1}\frac{\partial M}{\partial\alpha}. \tag{11.101}$$

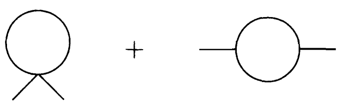
Using this identity, we find

$$\begin{aligned}\frac{i}{2}\frac{\delta}{\delta\phi_{\rm cl}^k(z)}&\log\det[-i\mathcal{D}]\\ &= i\Big[\lambda\Big(\phi_{\rm cl}^k(z)\delta^{ij} + \phi_{\rm cl}^i(z)\delta^{jk} + \phi_{\rm cl}^j(z)\delta^{ik}\Big)(i\mathcal{D}^{-1})^{ij}(z,z)\Big]\\ &= -\lambda\Big(\phi_{\rm cl}^k(z)\delta^{ij} + \phi_{\rm cl}^i(z)\delta^{jk} + \phi_{\rm cl}^j(z)\delta^{ik}\Big)(\mathcal{D}^{-1})^{ij}(z,z).\end{aligned} \tag{11.102}$$

The quantity $(\mathcal{D}^{-1})^{ij}(x,y)$ is the Klein-Gordon propagator. To differentiate a second time, we can use the identity (11.94); this yields

$$\begin{aligned}\frac{i}{2}\frac{\delta^2}{\delta\phi^k_{\rm cl}(z)\delta\phi^\ell_{\rm cl}(w)}&\log\det[-i\mathcal{D}]\\ &= -\lambda(\delta^{k\ell}\delta^{ij}+\delta^{ik}\delta^{j\ell}+\delta^{i\ell}\delta^{jk})(\mathcal{D}^{-1})^{ij}(z,z)\delta(z-w)\\ &\quad + 2i\lambda^2(\phi^k_{\rm cl}(z)\delta^{ij}+\phi^i_{\rm cl}(z)\delta^{jk}+\phi^j_{\rm cl}(z)\delta^{ik})(\mathcal{D}^{-1})^{im}(z,w)\\ &\qquad\cdot(\phi^\ell_{\rm cl}(w)\delta^{mn}+\phi^m_{\rm cl}(w)\delta^{n\ell}+\phi^n_{\rm cl}(w)\delta^{m\ell})(\mathcal{D}^{-1})^{nj}(w,z). \end{aligned} \tag{11.103}$$

This is expected to be the formal correction to the inverse propagator at one-loop order, and indeed we can recognize in (11.103) the values of the one-loop diagrams



Notice how, in this derivation, every functional derivative on $\mathcal{D}^{-1}$ adds another propagator to the diagram and thus lowers the degree of divergence, in conformity with our general arguments in Section 10.1.

This example illustrates that the successive functional derivatives of $\Gamma[\phi_{\rm cl}]$ are computed by a Feynman diagram expansion, with propagators and vertices that depend on the classical field. When the classical field is a constant, the propagators reduce to ordinary Klein-Gordon propagators and so the BPHZ theorem applies. All ultraviolet divergences can be removed from all of the amplitudes obtained by differentiating $\Gamma[\phi_{\rm cl}]$ by the use of the most general set of mass, vertex, and field-strength renormalizations. At the same time, the perturbation theory is manifestly invariant to the symmetry of the original Lagrangian, and so the only divergences that appear—and thus the only counterterms required—are those that respect this symmetry. In general, then, all amplitudes of a renormalizable theory of scalar fields invariant under a symmetry group can be made finite using only the set of counterterms invariant to the symmetry. This gives a complete and quite satisfactory answer to the question posed at the beginning of Section 11.2.

The computation of the effective action in spatially varying background fields has not been analyzed at the level of rigor involved in the proof of the BPHZ theorem. However, it is expected that in this situation also, the standard set of counterterms for the symmetric theory should suffice. We can argue this intuitively by using the fact that the ultraviolet divergences of Feynman diagrams are local in spacetime. Thus, to understand the divergences of a computation in a background $\phi_{\rm cl}(x)$ that is smoothly varying, we can divide spacetime into small boxes, in each of which $\phi_{\rm cl}(x)$ is approximately constant, and expand in the derivatives $\partial_\mu\phi_{\rm cl}(x)$. In this expansion in powers of $\partial_\mu\phi_{\rm cl}(x)$, the Taylor series coefficients are functional derivatives of Γ in a constant background, which we know can be renormalized. The conclusion of this intuitive argument has been checked at the two-loop level for several nontrivial background field configurations.

Our general result on the renormalization of theories with spontaneously broken symmetry has an important implication for the physical predictions of these theories. In a renormalizable field theory, the most basic quantities of the theory cannot be predicted, because they are the quantities that must be specified as part of the definition of the theory. For example, in QED, the mass and charge of the electron must be adjusted from outside in order to define the theory. The predictions of QED are quantities that do not appear in the basic Lagrangian, for example, the anomalous magnetic moment of the electron. In renormalizable theories with spontaneously broken symmetry, however, the symmetry-breaking produces a large number of distinct masses and couplings, which depend on the relatively small number of parameters of the original symmetric theory. After the original parameters of the theory are fixed, any additional observable of the theory can be predicted unambiguously. For example, in the linear sigma model studied in this chapter, we took the values of the four-point coupling λ and the vacuum expectation value $\langle\phi\rangle$ as input parameters; we then calculated the mass of the σ particle in terms of these parameters in an unambiguous way.

There is a general argument that implies that, once we fix the parameters of the Lagrangian, we must find an unambiguous, finite formula for the σ mass in ϕ^4 theory, or, more generally, for any additional parameter of a renormalizable quantum field theory. In general, this parameter will be determined at the classical level in terms of the couplings in the Lagrangian. For the example of the σ mass in the linear sigma model, this classical relation is

$$m - \sqrt{2\lambda}\,\langle\phi\rangle = 0, \tag{11.104}$$

where m is the mass of the σ and λ gives the four-ϕ scattering amplitude at threshold. In general, loop corrections will modify this relation, contributing some nonzero expression to the right-hand side of this equation. However, since Eq. (11.104) is valid at the classical level however the parameters of the Lagrangian are modified, it holds equally well when we add counterterms to the Lagrangian and then adjust these counterterms order by order. Thus, the counterterms must give zero contributions to the right-hand side of Eq. (11.104). Therefore, the perturbative corrections to Eq. (11.104) must be automatically ultraviolet-finite. A relation of this type, true at the classical level for all values of the couplings in the Lagrangian, but corrected by loop effects, is called a *zeroth-order natural relation.* The argument we have given implies that, for any such relation, the loop corrections are finite and constitute predictions of the quantum field theory. We will see another example of such a relation in Problem 11.2.

Goldstone's Theorem Revisited

As a final application of the effective action formalism, let us return to the question of whether Goldstone's theorem is valid in the presence of quantum corrections. Recall that we proved this theorem at the classical level at the end of Section 11.1: We showed in (11.13) that, if the Lagrangian has a continuous symmetry that is spontaneously broken, the matrix of second derivatives of the classical potential $V(\phi)$ has a corresponding zero eigenvalue. According to Eq. (11.11), this implies that the classical theory contains a massless scalar particle, associated with the spontaneously broken symmetry.

Using the effective action formalism, this argument can be repeated almost verbatim in the full quantum field theory. The effective potential $V_{\text{eff}}(\phi_{\text{cl}})$ encapsulates the full solution to the theory, including all orders of quantum corrections. At the same time, it satisfies the general properties of the classical potential: It is invariant to the symmetries of the theory, and its minimum gives the vacuum expectation value of ϕ. This means that the argument we gave in (11.13) works in exactly the same way for V_{eff} as it does for V: If a continuous symmetry of the original Lagrangian is spontaneously broken by $\langle\phi\rangle$, the matrix of second derivatives of $V_{\text{eff}}(\phi_{\text{cl}})$ has a zero eigenvalue along the symmetry direction.

We now argue that, just as at the classical level, the presence of such a zero eigenvalue implies the existence of a massless scalar particle. In our discussion of the general properties of the effective action, we showed that its second functional derivative is the inverse propagator, and that, through Eq. (11.97), this derivative yields the spectrum of masses in the quantum theory. Let us rewrite Eq. (11.97) for a theory that contains several scalar fields:

$$\int d^4x\, e^{ip\cdot(x-y)} \frac{\delta^2\Gamma}{\delta\phi^i\delta\phi^j}(x,y) = 0. \tag{11.105}$$

A particle of mass m corresponds to a zero eigenvalue of this matrix equation at $p^2 = m^2$. Now set $p = 0$. This implies that we differentiate $\Gamma[\phi_{\text{cl}}]$ with respect to constant fields. Thus, we can replace $\Gamma[\phi_{\text{cl}}]$ by its value with constant classical fields, which is just the effective potential. We find that the quantum field theory contains a scalar particle of zero mass when the matrix of second derivatives,

$$\frac{\partial^2 V_{\text{eff}}}{\partial\phi^i_{\text{cl}}\partial\phi^j_{\text{cl}}},$$

has a zero eigenvalue. This completes the proof of Goldstone's theorem.

This argument for Goldstone's theorem illustrates the power of the effective action formalism. The formalism gives a geometrical picture of spontaneous symmetry breaking that is valid to any order in quantum corrections. As a bonus, it is built up from objects that are renormalized in a simple way. This formalism will prove useful in understanding the applications of spontaneously broken symmetry that occur, in several different contexts, throughout the rest of this book.

Problems

11.1 Spin-wave theory.

(a) Prove the following wonderful formula: Let $\phi(x)$ be a free scalar field with propagator $\langle T\phi(x)\phi(0)\rangle = D(x)$. Then

$$\left\langle T e^{i\phi(x)} e^{-i\phi(0)} \right\rangle = e^{[D(x)-D(0)]}.$$

(The factor $D(0)$ gives a formally divergent adjustment of the overall normalization.)

(b) We can use this formula in Euclidean field theory to discuss correlation functions in a theory with spontaneously broken symmetry for $T < T_C$. Let us consider only the simplest case of a broken $O(2)$ or $U(1)$ symmetry. We can write the local spin density as a complex variable

$$s(x) = s^1(x) + is^2(x).$$

The global symmetry is the transformation

$$s(x) \to e^{-i\alpha} s(x).$$

If we assume that the physics freezes the modulus of $s(x)$, we can parametrize

$$s(x) = Ae^{i\phi(x)}$$

and write an effective Lagrangian for the field $\phi(x)$. The symmetry of the theory becomes the translation symmetry

$$\phi(x) \to \phi(x) - \alpha.$$

Show that (for $d > 0$) the most general renormalizable Lagrangian consistent with this symmetry is the free field theory

$$\mathcal{L} = \tfrac{1}{2}\rho(\vec{\nabla}\phi)^2.$$

In statistical mechanics, the constant ρ is called the *spin wave modulus.* A reasonable hypothesis for ρ is that it is finite for $T < T_C$ and tends to 0 as $T \to T_C$ from below.

(c) Compute the correlation function $\langle s(x)s^*(0)\rangle$. Adjust A to give a physically sensible normalization (assuming that the system has a physical cutoff at the scale of one atomic spacing) and display the dependence of this correlation function on x for $d = 1, 2, 3, 4$. Explain the significance of your results.

11.2 A zeroth-order natural relation. This problem studies an $N = 2$ linear sigma model coupled to fermions:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi^i)^2 + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}((\phi^i)^2)^2 + \bar{\psi}(i\partial\!\!\!/)\psi - g\bar{\psi}(\phi^1 + i\gamma^5\phi^2)\psi \tag{1}$$

where ϕ^i is a two-component field, $i = 1, 2$.

(a) Show that this theory has the following global symmetry:

$$\begin{aligned}\phi^1 &\to \cos\alpha\,\phi^1 - \sin\alpha\,\phi^2,\\ \phi^2 &\to \sin\alpha\,\phi^1 + \cos\alpha\,\phi^2,\\ \psi &\to e^{-i\alpha\gamma^5/2}\,\psi.\end{aligned} \tag{2}$$

Show also that the solution to the classical equations of motion with the minimum energy breaks this symmetry spontaneously.

(b) Denote the vacuum expectation value of the field ϕ^i by v and make the change of variables

$$\phi^i(x) = (v + \sigma(x), \pi(x)). \tag{3}$$

Write out the Lagrangian in these new variables, and show that the fermion acquires a mass given by

$$m_f = g \cdot v. \tag{4}$$

(c) Compute the one-loop radiative correction to m_f, choosing renormalization conditions so that v and g (defined as the $\psi\psi\pi$ vertex at zero momentum transfer) receive no radiative corrections. Show that relation (4) receives nonzero corrections but that these corrections are *finite*. This is in accord with our general discussion in Section 11.6.

11.3 The Gross-Neveu model. The Gross-Neveu model is a model in two spacetime dimensions of fermions with a discrete chiral symmetry:

$$\mathcal{L} = \bar\psi_i i\partial\!\!\!/\psi_i + \tfrac{1}{2}g^2(\bar\psi_i\psi_i)^2$$

with $i = 1, \ldots, N$. The kinetic term of two-dimensional fermions is built from matrices γ^μ that satisfy the two-dimensional Dirac algebra. These matrices can be 2×2:

$$\gamma^0 = \sigma^2, \qquad \gamma^1 = i\sigma^1,$$

where σ^i are Pauli sigma matrices. Define

$$\gamma^5 = \gamma^0\gamma^1 = \sigma^3;$$

this matrix anticommutes with the γ^μ.

(a) Show that this theory is invariant with respect to

$$\psi_i \to \gamma^5\psi_i,$$

and that this symmetry forbids the appearance of a fermion mass.

(b) Show that this theory is renormalizable in 2 dimensions (at the level of dimensional analysis).

(c) Show that the functional integral for this theory can be represented in the following form:

$$\int \mathcal{D}\bar\psi\mathcal{D}\psi\; e^{i\int d^2x\,\mathcal{L}} = \int \mathcal{D}\bar\psi\mathcal{D}\psi\mathcal{D}\sigma\;\exp\left[i\int d^2x\left\{\bar\psi_i i\partial\!\!\!/\psi_i - \sigma\bar\psi_i\psi_i - \frac{1}{2g^2}\sigma^2\right\}\right],$$

where $\sigma(x)$ (not to be confused with a Pauli matrix) is a new scalar field with no kinetic energy terms.

(d) Compute the leading correction to the effective potential for σ by integrating over the fermion fields ψ_i. You will encounter the determinant of a Dirac operator; to evaluate this determinant, diagonalize the operator by first going to Fourier components and then diagonalizing the 2×2 Pauli matrix associated with each Fourier mode. (Alternatively, you might just take the determinant of this 2×2 matrix.) This 1-loop contribution requires a renormalization proportional to σ^2 (that is, a renormalization of g^2). Renormalize by minimal subtraction.

(e) Ignoring two-loop and higher-order contributions, minimize this potential. Show that the σ field acquires a vacuum expectation value which breaks the symmetry of part (a). Convince yourself that this result does not depend on the particular renormalization condition chosen.

(f) Note that the effective potential derived in part (e) depends on g and N according to the form

$$V_{\text{eff}}(\sigma_{\text{cl}}) = N \cdot f(g^2 N).$$

(The overall factor of N is expected in a theory with N fields.) Construct a few of the higher-order contributions to the effective potential and show that they contain additional factors of N^{-1} which suppress them if we take the limit $N \to \infty$, $(g^2 N)$ fixed. In this limit, the result of part (e) is unambiguous.

Chapter 12

The Renormalization Group

In the past two chapters, our main goal has been to determine when, and how, the cancellation of ultraviolet divergences in quantum field theory takes place. We have seen that, in a large class of field theories, the divergences appear only in the values of a few parameters: the bare masses and coupling constants, or, in renormalized perturbation theory, the counterterms. Aside from the shift in these parameters, virtual particles with very large momenta have no effect on computations in these theories.

The cancellation of ultraviolet divergences is essential if a theory is to yield quantitative physical predictions. But, at a deep level, the fact that high-momentum virtual quanta can have so little effect on a theory is quite surprising. One of the essential features of quantum field theory is locality, that is, the fact that fields at different spacetime points are independent degrees of freedom with independent quantum fluctuations. The quantum fluctuations at arbitrarily short distances appear in Feynman diagram computations as virtual quanta with arbitrarily high momenta. In a renormalizable theory, the loop integrals over virtual-particle momenta are always dominated by values comparable to the finite external particle momenta. But why? It is not easy to understand how the quantum fluctuations associated with extremely short distances can be so innocuous as to affect a theory only through the values of a few of its parameters.

This chapter begins with a physical picture, due to Kenneth Wilson, that explains this unusual and counterintuitive simplification. This picture generalizes the idea of the distance- or scale-dependent electric charge, introduced at the end of Chapter 7, and suggests that all of the parameters of a renormalizable field theory can usefully be thought of as scale-dependent entities. We will see that this scale dependence is described by simple differential equations, called *renormalization group* equations. The solutions of these equations will lead to physical predictions of a completely new type: predictions that, under certain circumstances, the correlation functions of a quantum field exhibit unusual but computable scaling laws as a function of their coordinates.

12.1 Wilson's Approach to Renormalization Theory

Wilson's method is based on the functional integral approach to field theory, in which the degrees of freedom of a quantum field are variables of integration. In this approach, one can study the origin of ultraviolet divergences by isolating the dependence of the functional integral on the short-distance degrees of freedom of the field.* In this section, we will illustrate this idea in the simplest example of ϕ^4 theory.

To make our analysis more concrete, we will drop the elegant but somewhat mysterious method of dimensional regularization in this section and instead use a sharp momentum cutoff. Since we will be working here only in ϕ^4 theory, we will not be concerned that this cutoff makes it difficult to satisfy Ward identities. Wilson's analysis can be adapted to QED and other situations where this subtlety is important, but the case of ϕ^4 theory is sufficient to give us the basic qualitative results of this approach.

In Section 9.2, we constructed the Green's functions of ϕ^4 theory in terms of a functional integral representation of the generating functional $Z[J]$. The basic integration variables are the Fourier components of the field $\phi(k)$, so $Z[J]$ is given concretely by the expression

$$Z[J] = \int \mathcal{D}\phi \, e^{i\int[\mathcal{L}+J\phi]} = \left(\prod_k \int d\phi(k)\right) e^{i\int[\mathcal{L}+J\phi]}. \tag{12.1}$$

To impose a sharp ultraviolet cutoff Λ, we restrict the number of the integration variables displayed in (12.1). That is, we integrate only over $\phi(k)$ with $|k| \leq \Lambda$, and set $\phi(k) = 0$ for $|k| > \Lambda$.

This modification of the functional integral suggests a method for assessing the influence of the quantum fluctuations at very short distances or very large momenta. In the functional integral representation, these fluctuations are represented by the integrals over the Fourier components of ϕ with momenta near the cutoff. Why not explicitly perform the integrals over these variables? Then we can compare the result to the original functional integral, and determine precisely the influence of these high-momentum modes on the physical predictions of the theory.

Before beginning this analysis, though, we must introduce one modification. At first sight, it seems most natural to define the ultraviolet cutoff in Minkowski space. However, a cutoff $k^2 \leq \Lambda^2$ is not completely effective in controlling large momenta, since in lightlike directions the components of k can be very large while k^2 remains small. We will therefore consider the cutoff to be imposed on the Euclidean momenta obtained after Wick rotation. Equivalently, we consider the Euclidean form of the functional integral, presented in Section 9.3, and restrict its variables $\phi(k)$, with k Euclidean, to $|k| \leq \Lambda$.

---

*Wilson's ideas are reviewed in K. G. Wilson and J. Kogut, *Phys. Repts.* **12C,** 75 (1974).

The transition to Euclidean space also brings us closer to the connection between renormalization theory and statistical mechanics advertised in Chapter 8. As we saw in Section 9.3, the Euclidean functional integral for ϕ^4 theory has precisely the same form as the continuum description of the statistical mechanics of a magnet. The field $\phi(x)$ is interpreted as the fluctuating spin field $s(x)$. A real magnet is built of atoms, and the atomic spacing provides a physical cutoff, a shortest distance over which fluctuations can take place. The cut-off functional integral models the effects of this atomic size in a crude way.

By pursuing this analogy, we can derive some physical intuition about the effects of the ultraviolet cutoff in a field theory. In a magnet, it is quite easy to visualize statistical fluctuations of the spins at the atomic scale. In fact, for values of the temperature away from any critical points, the statistical fluctuations are restricted to this scale; over distances of tens of atomic spacings, the magnet already shows its homogeneous macroscopic behavior. We have seen in Chapter 8 that we can approximate the correlation function of the spin field by the propagator of a Euclidean ϕ^4 theory. In this approximation,

$$\langle s(x)s(0)\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{e^{ik\cdot x}}{k^2+m^2} \underset{|x|\to\infty}{\sim} e^{-m|x|}. \tag{12.2}$$

As long as the temperature is far from the critical temperature, the size of the "mass" m is determined by the one natural scale in the problem, the atomic spacing. Thus, we expect $m \approx \Lambda$. In our field theory calculations, we were specifically interested in the situation where $m \ll \Lambda$, and we adjusted the parameters of the theory to satisfy this condition. In describing a magnet, it appears that no such adjustment is called for.

However, we saw in Chapter 8 that there is one circumstance in which the correlations of the spin field are much longer than the atomic spacing, so that, indeed, $m \ll \Lambda$. When the spin system begins to magnetize, just in the vicinity of the critical point, the spins become correlated over arbitrarily long distances as the fluctuating spins attempt to choose their eventual direction of magnetization. To study these long-range correlations in a magnet, one must carefully adjust the temperature to bring the system into the vicinity of the phase transition. In the same way, we can imagine making a fine adjustment of the parameter m of ϕ^4 theory to bring the quantum field theory into a region of parameters where we do find correlations of the field $\phi(x)$ over distances much larger than $1/\Lambda$.

Integrating Over a Single Momentum Shell

With this introduction, we will now carry out the integration over the high-momentum degrees of freedom of ϕ. We begin by writing the functional integral (12.1) more explicitly for the case of ϕ^4 theory. We apply the cutoff prescription described earlier, and set $J = 0$ for simplicity. Then

$$Z = \int [\mathcal{D}\phi]_\Lambda \exp\Big(-\int d^dx \Big[\frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4\Big]\Big), \tag{12.3}$$

where

$$[\mathcal{D}\phi]_\Lambda = \prod_{|k|<\Lambda} d\phi(k). \tag{12.4}$$

In the Lagrangian of Eq. (12.3), m and λ are the bare parameters, and so there are no counterterms. As in our study of the superficial degree of divergence, it will be useful to carry out this analysis in an arbitrary spacetime dimension d.

We now divide the integration variables $\phi(k)$ into two groups. Choose a fraction $b < 1$. The variables $\phi(k)$ with $b\Lambda \le |k| < \Lambda$ are the high-momentum degrees of freedom that we will integrate over. To label these degrees of freedom, let us define

$$\hat{\phi}(k) = \begin{cases} \phi(k) & \text{for } b\Lambda \le |k| < \Lambda; \\ 0 & \text{otherwise.} \end{cases}$$

Next, let us define a new $\phi(k)$, which is identical to the old for $|k| < b\Lambda$ and zero for $|k| > b\Lambda$. Then we can replace the old ϕ in the Lagrangian with $\phi + \hat{\phi}$, and rewrite Eq. (12.3) as

$$\begin{aligned} Z &= \int \mathcal{D}\phi \int \mathcal{D}\hat{\phi} \exp\Big(-\int d^dx \Big[\frac{1}{2}(\partial_\mu\phi + \partial_\mu\hat{\phi})^2 + \frac{1}{2}m^2(\phi+\hat{\phi})^2 + \frac{\lambda}{4!}(\phi+\hat{\phi})^4\Big]\Big) \\ &= \int \mathcal{D}\phi\, e^{-\int \mathcal{L}(\phi)} \int \mathcal{D}\hat{\phi} \exp\Big(-\int d^dx \Big[\frac{1}{2}(\partial_\mu\hat{\phi})^2 + \frac{1}{2}m^2\hat{\phi}^2 \\ &\qquad + \lambda\Big(\frac{1}{6}\phi^3\hat{\phi} + \frac{1}{4}\phi^2\hat{\phi}^2 + \frac{1}{6}\phi\hat{\phi}^3 + \frac{1}{4!}\hat{\phi}^4\Big)\Big]\Big). \end{aligned} \tag{12.5}$$

In the final expression we have gathered all terms independent of $\hat{\phi}$ into $\mathcal{L}(\phi)$. Note that quadratic terms of the form $\phi\hat{\phi}$ vanish, since Fourier components of different wavelengths are orthogonal.

The next few paragraphs will explain how to perform the integral over $\hat{\phi}$. This integration will transform (12.5) into an expression of the form

$$Z = \int [\mathcal{D}\phi]_{b\Lambda} \exp\Big(-\int d^dx\, \mathcal{L}_{\text{eff}}\Big), \tag{12.6}$$

where $\mathcal{L}_{\text{eff}}(\phi)$ involves only the Fourier components $\phi(k)$ with $|k| < b\Lambda$. We will see that $\mathcal{L}_{\text{eff}}(\phi) = \mathcal{L}(\phi)$ plus corrections proportional to powers of λ. These correction terms compensate for the removal of the large-k Fourier components $\hat{\phi}$, by supplying the interactions among the remaining $\phi(k)$ that were previously mediated by fluctuations of the $\hat{\phi}$.

To carry out the integrals over the $\hat{\phi}(k)$, we use the same method that we applied in Section 9.2 to derive Feynman rules. In fact, we will see below that the new terms in $\mathcal{L}_{\text{eff}}$ can be written in a diagrammatic form. In this analysis, we treat the quartic terms in (12.5), all proportional to λ, as perturbations.

Since we are mainly interested in the situation $m^2 \ll \Lambda^2$, we will also treat the mass term $\frac{1}{2}m^2\hat{\phi}^2$ as a perturbation. Then the leading-order term in the portion of the Lagrangian involving $\hat{\phi}$ is

$$\int \mathcal{L}_0 = \frac{1}{2} \int\limits_{b\Lambda \leq |k| < \Lambda} \frac{d^dk}{(2\pi)^d} \hat{\phi}^*(k) k^2 \hat{\phi}(k). \tag{12.7}$$

This term leads to a propagator

$$\overbracket{\hat{\phi}(k)\hat{\phi}(p)} = \frac{\int \mathcal{D}\hat{\phi}\, e^{-\int \mathcal{L}_0} \hat{\phi}(k)\hat{\phi}(p)}{\int \mathcal{D}\hat{\phi}\, e^{-\int \mathcal{L}_0}} = \frac{1}{k^2}(2\pi)^d \delta^{(d)}(k+p)\Theta(k), \tag{12.8}$$

where

$$\Theta(k) = \begin{cases} 1 & \text{if } b\Lambda \leq |k| < \Lambda; \\ 0 & \text{otherwise.} \end{cases} \tag{12.9}$$

We will regard the remaining $\hat{\phi}$ terms in Eq. (12.5) as perturbations, and expand the exponential. The various contributions from these perturbations can be evaluated by using Wick's theorem with (12.8) as the propagator.

First consider the term that results from expanding to one power of the $\phi^2\hat{\phi}^2$ term in the exponent of (12.5). We find

$$-\int d^dx \, \frac{\lambda}{4} \phi^2 \overbracket{\hat{\phi}\hat{\phi}} = -\frac{1}{2} \int \frac{d^dk_1}{(2\pi)^d} \, \mu \, \phi(k_1)\phi(-k_1), \tag{12.10}$$

where the coefficient μ is the result of contracting the two $\hat{\phi}$ fields:

$$\mu = \frac{\lambda}{2} \int\limits_{b\Lambda \leq |k| < \Lambda} \frac{d^dk}{(2\pi)^d} \frac{1}{k^2} = \frac{\lambda}{(4\pi)^{d/2}\Gamma(\frac{d}{2})} \frac{1 - b^{d-2}}{d-2} \Lambda^{d-2}. \tag{12.11}$$

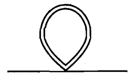
The term (12.10) could just as well have arisen from an expansion of the exponential

$$\exp\Bigl(-\int d^dx \, \frac{1}{2}\mu\phi^2 + \cdots\Bigr). \tag{12.12}$$

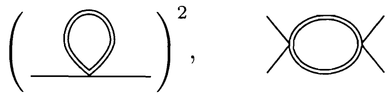
We will soon see that the rest of the perturbation series also organizes itself into this form. The coefficient μ therefore gives a positive correction to the m^2 term in $\mathcal{L}$.

The higher orders of the perturbation theory in the correction terms can be worked out in a similar way. As in our derivation of the standard perturbation theory for ϕ^4 theory, it is useful to adopt a diagrammatic notation. Represent the propagator (12.8) by a double line. This propagator will connect pairs of fields $\hat{\phi}$ from the various quartic interactions. Represent the fields ϕ in these interactions, which are not integrated over, as single external lines.

Then, for example, the contribution of (12.10) corresponds to the following diagram:



At order λ^2, we will have, among other contributions, terms involving the contractions of two interaction terms $\lambda\phi^2\hat{\phi}^2$. Each term corresponds to a vertex connecting two single lines and two double lines. There are two possible contractions:



(12.13)

Of these, the first, which is a disconnected diagram, supplies the order-λ^2 term in the exponential (12.12). The second is a new contribution, which will become a correction to the ϕ^4 interaction in $\mathcal{L}(\phi)$.

Let us now evaluate this second contribution. For simplicity, we consider the limit in which the external momenta carried by the factors ϕ are very small compared to $b\Lambda$, so we can ignore them. Then this diagram has the value

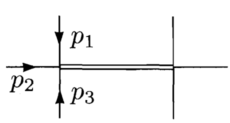
$$-\frac{1}{4!}\int d^dx\,\zeta\,\phi^4, \tag{12.14}$$

where

$$\zeta = -4!\frac{2}{2!}\left(\frac{\lambda}{4}\right)^2 \int\limits_{b\Lambda\leq|k|<\Lambda} \frac{d^dk}{(2\pi)^d}\left(\frac{1}{k^2}\right)^2 = \frac{-3\lambda^2}{(4\pi)^{d/2}\Gamma(\frac{d}{2})}\frac{(1-b^{d-4})}{d-4}\Lambda^{d-4}$$
$$\underset{d\to 4}{\longrightarrow} -\frac{3\lambda^2}{16\pi^2}\log\frac{1}{b}. \tag{12.15}$$

The 2 in the numerator counts the two possible contractions; there are no additional combinatoric factors from counting external legs or vertices. In the analysis of ϕ^4 theory in Section 10.2, we encountered a similar diagram, integrated over a range of momenta from 0 to Λ, producing a logarithmic ultraviolet divergence. In Wilson's treatment this divergence is not a pathology but simply a sign that the diagram is receiving contributions from all momentum scales. Indeed, it receives an equal contribution from each logarithmic interval between the momentum scales m and Λ. We will see below that the (finite) contribution to this diagram from each momentum interval has a natural physical importance.

The diagrammatic perturbation theory we have described not only generates contributions proportional to ϕ^2 and ϕ^4 but also to higher powers of ϕ. For example, the following diagram generates a contribution to a ϕ^6 interaction:



$$\propto \frac{\lambda^2}{(p_1+p_2+p_3)^2}\Theta(p_1+p_2+p_3). \tag{12.16}$$

There are also derivative interactions, which arise when we no longer neglect the external momenta of the diagrams. A more exact treatment would Taylor-expand in these momenta; for instance, in addition to expression (12.14), we would obtain terms with two powers of external momenta, which we could rewrite as

$$-\frac{1}{4}\int d^dx\, \eta\, \phi^2(\partial_\mu\phi)^2. \tag{12.17}$$

We would also find terms with four, six, and more powers of the momenta carried by the ϕ. In general, the procedure of integrating out the $\hat{\phi}$ generates all possible interactions of the fields ϕ and their derivatives.

The diagrammatic corrections can be simplified slightly by resumming them as an exponential. We have seen already in (12.13) that our diagrammatic expansion generates disconnected diagrams. By the same combinatoric argument that we used in Eq. (4.52), we can rewrite the sum of the series as the exponential of the sum of the connected diagrams. This leads precisely to expression (12.6), with

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{1}{4!}\lambda\phi^4 + (\text{sum of connected diagrams}). \tag{12.18}$$

The diagrammatic contributions include corrections to m^2 and λ, as well as all possible higher-dimension operators. We can now use the new Lagrangian $\mathcal{L}_{\text{eff}}(\phi)$ to compute correlation functions of the $\phi(k)$, or to compute S-matrix elements. Since the $\phi(k)$ include only momenta up to $b\Lambda$, the loop diagrams in such a calculation would be integrated only up to that lowered cutoff. The correction terms in (12.18) precisely compensate for this change.

One might well be puzzled by the appearance of higher-dimension operators in Eq. (12.18). We chose the original Lagrangian of ϕ^4 theory to contain only renormalizable interactions. At first sight, it is disturbing that all possible nonrenormalizable interactions appear when we integrate out the variables $\hat{\phi}$. However, we will see below that our procedure actually keeps the contributions of these nonrenormalizable interactions under control. In fact, our analysis will imply that the presence of nonrenormalizable interactions in the original Lagrangian, defined to be used with very large cutoff Λ, has negligible effect on physics at scales much less than Λ.

Renormalization Group Flows

Let us now make a more careful comparison of the new functional integral (12.6) and the one we started with (12.3). The most convenient way to do this is to rescale distances and momenta in (12.6) according to

$$k' = k/b, \qquad x' = xb, \tag{12.19}$$

so that the variable k' is integrated over $|k'| < \Lambda$. Let us express the explicit form of (12.18) schematically as

$$\int d^dx\,\mathcal{L}_{\text{eff}} = \int d^dx\left[\frac{1}{2}(1+\Delta Z)(\partial_\mu\phi)^2 + \frac{1}{2}(m^2+\Delta m^2)\phi^2 + \frac{1}{4!}(\lambda+\Delta\lambda)\phi^4 + \Delta C(\partial_\mu\phi)^4 + \Delta D\phi^6 + \cdots\right]. \tag{12.20}$$

In terms of the rescaled variable x', this becomes

$$\int d^dx\,\mathcal{L}_{\text{eff}} = \int d^dx'\,b^{-d}\left[\frac{1}{2}(1+\Delta Z)b^2(\partial'_\mu\phi)^2 + \frac{1}{2}(m^2+\Delta m^2)\phi^2 + \frac{1}{4!}(\lambda+\Delta\lambda)\phi^4 + \Delta C b^4(\partial'_\mu\phi)^4 + \Delta D\phi^6 + \cdots\right]. \tag{12.21}$$

Throughout this analysis, we have treated all terms beyond the first as small perturbations. As long as the original couplings are small, this is still a valid approximation in treating (12.21).

The original functional integral led to the propagator (12.8). The new action (12.21) will give rise to exactly the same propagator, if we rescale the field ϕ according to

$$\phi' = \left[b^{2-d}(1+\Delta Z)\right]^{1/2}\phi. \tag{12.22}$$

After this rescaling, the unperturbed action returns to its initial form, while the various perturbations undergo a transformation:

$$\int d^dx\,\mathcal{L}_{\text{eff}} = \int d^dx'\left[\frac{1}{2}(\partial'_\mu\phi')^2 + \frac{1}{2}m'^2\phi'^2 + \frac{1}{4!}\lambda'\phi'^4 + C'(\partial'_\mu\phi')^4 + D'\phi'^6 + \cdots\right]. \tag{12.23}$$

The new parameters of the Lagrangian are

$$\begin{aligned} m'^2 &= (m^2+\Delta m^2)(1+\Delta Z)^{-1}b^{-2},\\ \lambda' &= (\lambda+\Delta\lambda)(1+\Delta Z)^{-2}b^{d-4},\\ C' &= (C+\Delta C)(1+\Delta Z)^{-2}b^{d},\\ D' &= (D+\Delta D)(1+\Delta Z)^{-3}b^{2d-6}, \end{aligned} \tag{12.24}$$

and so on. (The original Lagrangian had $C = D = 0$, but the same equations would apply if the initial values of C and D were nonzero.) All of the corrections, Δm^2, $\Delta\lambda$, and so on, arise from diagrams and thus are small compared to the leading terms if perturbation theory is justified.

By combining the operation of integrating out high-momentum degrees of freedom with the rescaling (12.19), we have rewritten this operation as a transformation of the Lagrangian. Continuing this procedure, we could integrate over another shell of momentum space and transform the Lagrangian further. Successive integrations produce further iterations of the transformation (12.24). If we take the parameter b to be close to 1, so that the shells of momentum space are infinitesimally thin, the transformation becomes a continuous one. We can then describe the result of integrating over the high-momentum degrees of freedom of a field theory as a trajectory or a flow in the space of all possible Lagrangians.

For historical reasons, these continuously generated transformations of Lagrangians are referred to as the *renormalization group.* They do not form a group in the formal sense, because the operation of integrating out degrees of freedom is not invertible. On the other hand, they are most certainly connected to renormalization, as we will now see.

Imagine that we wish to compute a correlation function of fields whose momenta p_i are all much less than Λ. We could compute this correlation function perturbatively using either the original Lagrangian $\mathcal{L}$, or the effective Lagrangian $\mathcal{L}_{\text{eff}}$ obtained after integrating over all momentum shells down to the scale of the external momenta p_i. Both procedures must ultimately yield the same result. But in the first case, the effects of high-momentum fluctuations of the field do not show up until we compute loop diagrams. In the second case, these effects have already been absorbed into the new coupling constants (m', λ', etc.), so their influence can be seen directly from the Lagrangian. In the first procedure, the large shifts from the original (bare) parameters to the values appropriate to low-momentum processes appear suddenly in one-loop diagrams, and seem to invalidate the use of perturbation theory. In the second approach, these corrections are introduced slowly and systematically. A perturbative treatment is valid at every step as long as the effective coupling constants such as λ' remain small.

However, the parameters of the effective Lagrangian may be very different from those of the original Lagrangian, since we must iterate the transformation (12.24) many times to get from the large momentum Λ down to the momentum scale of typical experiments. Let us therefore look more closely at how the Lagrangian tends to vary under the renormalization group transformations.

The simplest case to consider is a Lagrangian in the vicinity of the point $m^2 = \lambda = C = D = \cdots = 0$, where all the perturbations vanish. We have defined our transformation so that this point is left unchanged; we say that the free-field Lagrangian

$$\mathcal{L}_0 = \tfrac{1}{2}(\partial_\mu \phi)^2 \tag{12.25}$$

is a *fixed point* of the renormalization group transformation.

In the vicinity of $\mathcal{L}_0$, we can ignore the terms Δm^2, $\Delta\lambda$, etc., in the iteration equations (12.24) and keep only those terms that are linear in the perturbations. This gives an especially simple transformation law:

$$m'^2 = m^2 b^{-2}, \quad \lambda' = \lambda b^{d-4}, \quad C' = C b^d, \quad D' = D b^{2d-6}, \quad \text{etc.} \tag{12.26}$$

Since $b < 1$, those parameters that are multiplied by negative powers of b

grow, while those that are multiplied by positive powers of b decay. If the Lagrangian contains growing coefficients, these will eventually carry it away from $\mathcal{L}_0$.

It is conventional to speak of the various terms in the effective Lagrangian as a set of local operators that can be added as perturbations to $\mathcal{L}_0$. We call the operators whose coefficients grow during the recursion procedure *relevant* operators. The coefficients that die away are associated with *irrelevant* operators. For example, the scalar field mass operator ϕ^2 is always relevant, while the ϕ^4 operator is relevant if $d < 4$. If the coefficient of some operator is multiplied by b^0 (for example, the operator ϕ^4 in $d = 4$), we call this operator *marginal*; to find out whether its coefficient grows or decays, we must include the effect of higher-order corrections.

In general, an operator with N powers of ϕ and M derivatives has a coefficient that transforms as

$$C'_{N,M} = b^{N(d/2-1)+M-d} C_{N,M}. \tag{12.27}$$

Notice that the coefficient is just $(d_{N,M} - d)$, where $d_{N,M}$ is the mass dimension of the operator as computed at the end of Section 10.1. In other words, relevant and marginal operators about the free theory $\mathcal{L}_0$ correspond precisely to super-renormalizable and renormalizable interaction terms in the power-counting analysis of Section 10.1.

We can also understand the evolution of coefficients near the free-field fixed point using straightforward dimensional analysis. An operator with mass dimension d_i has a coefficient with dimension $(\text{mass})^{d-d_i}$. The natural order of magnitude for this mass is the cutoff Λ. Thus, if $d_i < d$, the perturbation is increasingly important at low momenta. On the other hand, if $d_i > d$, the relative size of this term decreases as $(p/\Lambda)^{d_i - d}$ as the momentum $p \to 0$; thus the term is truly irrelevant.

We have now shown that, at least in the vicinity of the zero-coupling fixed point, an arbitrarily complicated Lagrangian at the scale of the cutoff degenerates to a Lagrangian containing only a finite number of renormalizable interactions. It is instructive to compare this result with the conclusions of Chapter 10. There we took the philosophy that the cutoff Λ should be disposed of by taking the limit $\Lambda \to \infty$ as quickly as possible. We found that this limit gives well-defined predictions only if the Lagrangian contains no parameters with negative mass dimension. From this viewpoint, it seemed exceedingly fortunate that QED, for example, contained no such parameters, since otherwise this theory would not yield well-defined predictions.

Wilson's analysis takes just the opposite point of view, that any quantum field theory is defined fundamentally with a cutoff Λ that has some physical significance. In statistical mechanical applications, this momentum scale is the inverse atomic spacing. In QED and other quantum field theories appropriate to elementary particle physics, the cutoff would have to be associated with some fundamental graininess of spacetime, perhaps a result of quantum fluctuations in gravity. We discuss some speculations on the nature of this

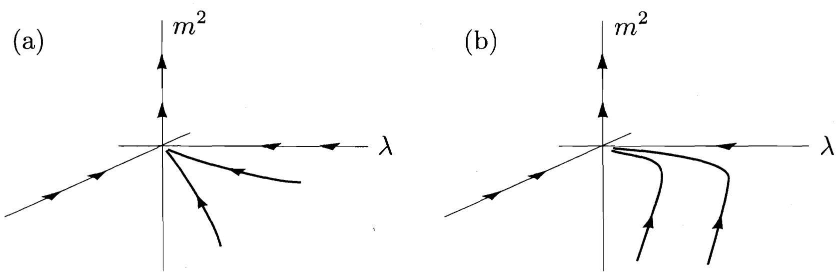


Figure 12.1. Renormalization group flows near the free-field fixed point in scalar field theory: (a) $d > 4$; (b) $d = 4$.

cutoff in the Epilogue. But whatever this scale is, it lies far beyond the reach of present-day experiments. The argument we have just given shows that this circumstance *explains* the renormalizability of QED and other quantum field theories of particle interactions. Whatever the Lagrangian of QED was at its fundamental scale, as long as its couplings are sufficiently weak, it must be described at the energies of our experiments by a renormalizable effective Lagrangian.

On the other hand, we should emphasize that these simple conclusions can be altered by sufficiently strong field theory interactions. Away from the free-field fixed point, the simple transformation laws (12.26) receive corrections proportional to higher powers of the coupling constants. If these corrections are large enough, they can halt or reverse the renormalization group flow. They could even create new fixed points, which would give new types of $\Lambda \to \infty$ limits.

To illustrate the possible influences of interactions in a relatively simple context, let us discuss the renormalization group flows near $\mathcal{L}_0$ for the specific case of ϕ^4 theory. It is instructive to consider the three cases $d > 4$, $d = 4$, and $d < 4$ in turn. When $d > 4$, the only relevant operator is the scalar field mass term. Then the renormalization group flows near $\mathcal{L}_0$ have the form shown in Fig. 12.1(a). The ϕ^4 interaction and possible higher-order interactions die away, while the mass term increases in importance.

In previous chapters, we have always discussed ϕ^4 theory in the limit in which the mass is small compared to the cutoff. Let us take a moment to rewrite this condition in the language of renormalization group flows. In the course of the flow, the effective mass term m'^2 becomes large and eventually comes to equal the current cutoff. For example, near the free-field fixed point, after n iterations, $m'^2 = m^2 b^{-2n}$, and eventually there is an n such that $m'^2 \sim \Lambda^2$. At this point, we have integrated out the entire momentum region between the original Λ and the effective mass of the scalar field. The mass term then suppresses the remaining quantum fluctuations. In general, the criterion that the scalar field mass is small compared to the cutoff is equivalent to the statement that $m'^2 \sim \Lambda^2$ only after a large number of iterations of the renormalization group transformation.

This criterion is met whenever the initial conditions for the renormalization group flow are adjusted so that the trajectory passes very close to a fixed point. In principle, the flow could begin far away, along the direction of an irrelevant operator. The original value of m^2 need not be particularly small, as long as this original value is canceled by corrections arising from the diagrammatic contributions to $\mathcal{L}_{\text{eff}}$. Thus we could imagine constructing a scalar field theory in $d > 4$ by writing a complicated nonlinear Lagrangian, but adjusting the original m^2 so the trajectory that begins at this Lagrangian eventually passes close to the free-field fixed point $\mathcal{L}_0$. In this case, the effective theory at momenta small compared to the cutoff should be extremely simple: It will be a free field theory with negligible nonlinear interaction. As will be discussed in the next chapter, this remarkable prediction has been verified in mathematical models of magnetic systems in more than four dimensions: Even though the original model is highly nonlinear, the correlation function of spins near the phase transition has the free-field form given by the higher-dimensional analogue of Eq. (12.2).

Next consider the case $d = 4$. For this case, Eq. (12.26) does not give enough information to tell us whether the ϕ^4 interaction is important or unimportant at large distances. So we must go back to the complete transformation law (12.24). The leading contribution to $\Delta\lambda$ is given by Eq. (12.15). The leading contribution to ΔZ is of order λ^2 and can be neglected. (This is just what happened with the first correction to δ_Z in Section 10.2.) Thus we find the transformation

$$\lambda' = \lambda - \frac{3\lambda^2}{16\pi^2}\log(1/b). \tag{12.28}$$

This says that λ slowly decreases as we integrate out high-momentum degrees of freedom.

The diagram contributing to the correction $\Delta\lambda$ has the same structure as the one-loop diagrams computed in Section 10.2. In fact, these are essentially the same diagrams, and differ only in whether the integrals are carried out iteratively or all at once. However, whereas the diagrams in Section 10.2 had ultraviolet divergences, the corresponding diagram in Wilson's approach is well defined and gives the coefficient of a simple evolution equation of the coupling constant. This transformation gives a first example of the reinterpretation of ultraviolet divergences that we will make in this chapter.

The transformation law (12.28) implies that the renormalization group flows near $\mathcal{L}_0$ have the form shown in Fig. 12.1(b), with one slowly decaying direction. If we follow the flows far enough, the behavior should again be that of a free field. This picture has the puzzling implication that four-dimensional interacting ϕ^4 theory does not exist in the limit in which the cutoff goes to infinity. We will discuss this result further—and explain why it nevertheless makes sense to use ϕ^4 theory as a model field theory—in Section 12.3.

Finally consider the case $d < 4$. Now λ becomes a relevant parameter.

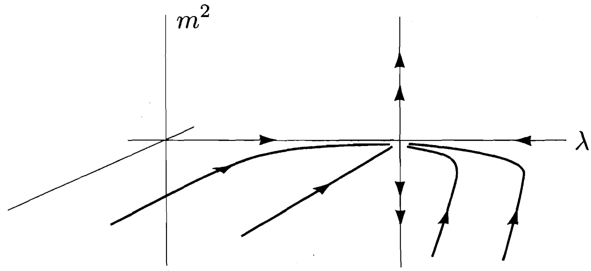


Figure 12.2. Renormalization group flows near the free-field fixed point in scalar field theory: $d < 4$.

The theory thus flows away from the free theory $\mathcal{L}_0$ as we integrate out degrees of freedom; at large distances, the ϕ^4 interaction becomes increasingly important. However, when λ becomes large, the nonlinear corrections such as that displayed in Eq. (12.28) must also be considered. If we include this specific effect in $d < 4$, we find the recursion formula

$$\lambda' = \left[\lambda - \frac{3\lambda^2}{(4\pi)^{d/2}\Gamma(\frac{d}{2})} \frac{b^{d-4} - 1}{4 - d} \Lambda^{d-4}\right] b^{d-4}. \tag{12.29}$$

This equation implies that there is a value of λ at which the increase due to rescaling is compensated by the decrease caused by the nonlinear effect. At this value, λ is unchanged when we integrate out degrees of freedom. The corresponding Lagrangian is a second fixed point of the renormalization group flow. In the limit $d \to 4$, the flow (12.29) tends to (12.28) and so the new fixed point merges with the free field fixed point. For d sufficiently close to 4, the new fixed point will share with $\mathcal{L}_0$ the property that the mass parameter m^2 is increased by the iteration. Then the mass operator will be a relevant operator near the new fixed point, so that the renormalization group flows will have the form shown in Fig. 12.2.

In this example, the new fixed point of the renormalization group had a Lagrangian with couplings weak enough that the transformation equations could be computed in perturbation theory. In principle, one could also find fixed points whose Lagrangians are strongly coupled, so that the renormalization group transformations cannot be understood by Feynman diagram analysis. Many examples of such fixed points are known in exactly solvable model field theories in two dimensions.† However, up to the present, all of the examples of quantum field theories that are important for physical applications have been found to be controlled either by the free field fixed point or by fixed points, like the one described in the previous paragraph, that approach the free-field fixed point in a specific limit. No one understands why this should be. This observation implies that Feynman diagram analysis has

†We mention some of these examples, and discuss other nonperturbative approaches to quantum field theory, in the Epilogue.

unexpected power in evaluating the physical consequences of quantum field theories.

One more aspect of ϕ^4 theory deserves comment. Since the mass term, $m^2\phi^2$, is a relevant operator, its coefficient diverges rapidly under the renormalization group flow. We have seen above that, in order to end up at the desired value of m^2 at low momentum, we must imagine that the value of m^2 in the original Lagrangian has been adjusted very delicately. This adjustment has a natural interpretation in a magnetic system as the need to sensitively adjust the temperature to be very close to the critical point. However, it seems quite artificial when applied to the quantum field theory of elementary particles, which purports to be a fundamental theory of Nature. This problem appears only for scalar fields, since for fermions the renormalization of the mass is proportional to the bare mass rather than being an arbitrary additive constant. Perhaps this is the reason why there seem to be no elementary scalar fields in Nature. We will return to this question in the Epilogue.

12.2 The Callan-Symanzik Equation

Wilson's picture of renormalization, as a flow in the space of possible Lagrangians, is beautifully intuitive, and gives us a deep understanding of why Nature should be describable in terms of renormalizable quantum field theories. In addition, however, this idea can be applied to extract further quantitative predictions from these theories. In the remainder of this chapter we will develop a formalism for extracting these predictions. Specifically, we will see that Wilson's picture leads to predictions for the form of the high- and low-momentum behavior of correlation functions. In the simplest cases, the correlation functions turn out to scale as powers of their external momenta, with power laws that do not appear at any fixed order of perturbation theory.

It is possible to derive these predictions directly from Wilson's procedure of integrating out slices in momentum space, as Wilson originally did. However, now that we understand the basic idea of renormalization group flows, it will be technically easier to work in the more familiar context of ordinary renormalized perturbation theory. The discussion of the previous section was physically motivated but technically complex. It involved awkward integrals over finite domains, and used the artificial parameter b, which must cancel out in any final results. Furthermore, we know from Section 7.5 that a cutoff regulator leads to even more trouble in QED, since it conflicts with the Ward identity. The discussion of the present section will be much more abstract and formal, but it will remove these technical problems. In this section and the next we will derive a flow equation for the coupling constant, similar to the one we derived in Section 12.1. To obtain the flows of the most general Lagrangians, we will need some additional tools, to be developed in Sections 12.4 and 12.5.

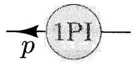
How can we hope to obtain information on renormalization group flows from the expressions for renormalized Green's functions, in which the cutoff has already been taken to infinity? We must first realize that renormalized quantum field theories correspond to a restricted class of the full set of possible Lagrangians that we considered in the previous section. In Wilson's language, a renormalized field theory with the cutoff taken arbitrarily large corresponds to a trajectory that takes an arbitrarily long time to evolve to a large value of the mass parameter. Such a trajectory must, then, pass arbitrarily close to a fixed point, which we will assume to be the weak-coupling fixed point. In the slow evolution past this fixed point, the irrelevant operators in the original Lagrangian die away, and we are left only with the relevant and marginal operators. The coefficients of these operators are in one-to-one correspondence with the parameters of the renormalizable field theory. Thus, in working with a renormalized field theory, we are throwing away information on the evolution of irrelevant perturbations, but keeping information on the flows of relevant and marginal perturbations.

The flows of these parameters cannot be determined from the cutoff dependence, because, in this framework, the cutoff has already been sent to infinity. However, we have an alternative, though more abstract, tool at our disposal. The parameters of a renormalized field theory are determined by a set of renormalization conditions, which are applied at a certain momentum scale (called the *renormalization scale*). By looking at how the parameters of the theory depend on the renormalization scale, we can recover the information contained in the renormalization group flows of the previous section.

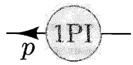
We consider first the specific case of ϕ^4 theory in four dimensions, where the coupling constant λ is dimensionless and the corresponding operator is marginal. For simplicity, we will also assume that the mass term m^2 has been adjusted to zero, so that the theory sits just at its critical point. We will perform this analysis in Minkowski space, using spacelike reference momenta. However, the analysis would be essentially identical if carried out in Euclidean space. If we wish to consider renormalization group predictions at timelike momenta, we must consider the possibilities of new singularities which make the analysis more complicated. These include both physical thresholds and the Sudakov double logarithms discussed in Section 6.4. We postpone discussion of these complications until Chapters 17 and 18.

Renormalization Conditions

To define the theory properly, we must specify the renormalization conditions. In Chapter 10 we used a natural set of renormalization conditions (10.19) for ϕ^4 theory, defined in terms of the physical mass m. However, in a theory where $m = 0$, these conditions cannot be used because they lead to singularities in the counterterms. (Consider, for example, the limit $m^2 \to 0$ of Eq. (10.24).) To avoid such singularities, we choose an arbitrary momentum scale M and impose the renormalization conditions at a spacelike momentum p with $p^2 = -M^2$:

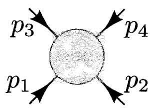


$= 0 \quad \text{at } p^2 = -M^2;$

$\frac{d}{dp^2}\Big($ 

$\Big) = 0 \quad \text{at } p^2 = -M^2;$

(12.30)



$= -i\lambda$

at $(p_1 + p_2)^2 = (p_1 + p_3)^2 = (p_1 + p_4)^2 = -M^2.$

The parameter M is called the *renormalization scale.* These conditions define the values of the two- and four-point Green's functions at a certain point and, in the process, remove all ultraviolet divergences. Speaking loosely, we say that we are "defining the theory at the scale M".

These new renormalization conditions take some getting used to. The second condition, in particular, implies that the two-point Green's function has a coefficient of 1 at the unphysical momentum $p^2 = -M^2$, rather than on shell (at $p^2 = 0$):

$$\langle\Omega|\,\phi(p)\phi(-p)\,|\Omega\rangle = \frac{i}{p^2} \qquad \text{at } p^2 = -M^2.$$

Here ϕ is the renormalized field, related to the bare field ϕ_0 by a scale factor that we again call Z:

$$\phi = Z^{-1/2}\phi_0. \tag{12.31}$$

This Z, however, is not the residue of the physical pole in the two-point Green's function of bare fields, as it was in Chapters 7 and 10. Instead, we now have

$$\langle\Omega|\,\phi_0(p)\phi_0(-p)\,|\Omega\rangle = \frac{iZ}{p^2} \qquad \text{at } p^2 = -M^2.$$

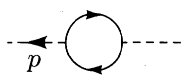
The Feynman rules for renormalized perturbation theory are the same as in Chapter 10, with the same relation between Z and the counterterm δ_Z,

$$\delta_Z = Z - 1.$$

Now, however, the counterterms δ_Z and δ_λ must be adjusted to maintain the new conditions (12.30).

The first renormalization condition in (12.30) holds the physical mass of the scalar field fixed at zero. We saw in Chapter 10 that, in ϕ^4 theory, the one-loop propagator correction is momentum-independent and is completely canceled by the mass renormalization counterterm. At two-loop order, however, the situation becomes more complicated, and the propagator corrections require both mass and field strength renormalizations. In more general scalar field theories, such as the Yukawa theory example considered at the end of Section 10.2, this complication arises already at one-loop order. Since the field strength renormalization counterterm will play an important role in the discussion below, it will be helpful to discuss briefly how we will treat this double subtraction.

The evaluation of propagator corrections has some special simplifications for the case of a massless scalar field, which we consider here, and specifically with the use of dimensional regularization. Consider, for example, the one-loop propagator correction in Yukawa theory. In Section 10.2 we found an expression of the form



$$\sim \frac{\Gamma(1-\frac{d}{2})}{\Delta^{1-d/2}}, \tag{12.32}$$

where Δ is a linear combination of the fermion mass m_f and p^2. If we compute the diagram using massless propagators only, Δ is proportional to p^2. Expression (12.32) has a pole at $d = 2$, corresponding to the quadratically divergent mass renormalization. However, the residue of this pole is independent of p^2, so we can completely cancel the pole with the mass counterterm δ_m. This allows us to analytically continue (12.32) to $d = 4$. Then this expression takes the form

$$-p^2\Big(\frac{1}{2-d/2} + \log\frac{1}{-p^2} + C\Big), \tag{12.33}$$

and gives no additional mass shift but only a field strength renormalization. The remaining divergence is canceled by the counterterm δ_Z. If we adopt the rule that we should simply continue expressions of the form (12.32) to $d = 4$, we can forget about the counterterm δ_m altogether.

In a regularization scheme with a momentum cutoff, the contributions to δ_m and δ_Z become tangled up with one another. Then it is more awkward to define the massless limit. In the following discussion, we will assume the use of dimensional regularization. However, to emphasize the physical role of the cutoff, we will write expressions of the form (12.33) as

$$-p^2\Big(\log\frac{\Lambda^2}{-p^2} + C\Big). \tag{12.34}$$

The logarithmically divergent terms proportional to p^2 will agree with the divergences obtained with a momentum cutoff; the constant terms will not agree, but these will drop out of our final results.

In ϕ^4 theory, where the one-loop propagator correction is momentum-independent, the one-loop diagram is simply set to zero by this prescription. Then the preceding analysis applies to the two-loop and higher correction terms.

The generalization of the analysis of this section to massive scalar field theory requires some additional formalism, which we postpone to Section 12.5.

The Callan-Symanzik Equation

In the renormalization conditions (12.30), the renormalization scale M is arbitrary. We could just as well have defined the same theory at a different scale M'. By "the same theory", we mean a theory whose bare Green's functions,

$$\langle\Omega|\, T\phi_0(x_1)\phi_0(x_2)\cdots\phi_0(x_n)\,|\Omega\rangle\,,$$

are given by the same functions of the bare coupling constant λ_0 and the cutoff Λ. These functions make no reference to M. The dependence on M enters only when we remove the cutoff dependence by rescaling the fields and eliminating λ_0 in favor of the renormalized coupling λ. The renormalized Green's functions are numerically equal to the bare Green's functions, up to a rescaling by powers of the field strength renormalization Z:

$$\langle\Omega|\, T\phi(x_1)\phi(x_2)\cdots\phi(x_n)\,|\Omega\rangle = Z^{-n/2}\,\langle\Omega|\, T\phi_0(x_1)\phi_0(x_2)\cdots\phi_0(x_n)\,|\Omega\rangle\,. \tag{12.35}$$

The renormalized Green's functions could be defined equally well at another scale M', using a new renormalized coupling λ' and a new rescaling factor Z'.

Let us write more explicitly the effect of an infinitesimal shift of M. Let $G^{(n)}(x_1,\cdots,x_n)$ be the connected n-point function, computed in renormalized perturbation theory:

$$G^{(n)}(x_1,\cdots,x_n) = \langle\Omega|\, T\phi(x_1)\cdots\phi(x_n)\,|\Omega\rangle_{\text{connected}}\,. \tag{12.36}$$

Now suppose that we shift M by δM. There is a corresponding shift in the coupling constant and the field strength such that the bare Green's functions remain fixed:

$$\begin{aligned} M &\to M + \delta M,\\ \lambda &\to \lambda + \delta\lambda,\\ \phi &\to (1+\delta\eta)\phi. \end{aligned} \tag{12.37}$$

Then the shift in any renormalized Green's function is simply that induced by the field rescaling,

$$G^{(n)} \to (1+n\delta\eta)G^{(n)}.$$

If we think of $G^{(n)}$ as a function of M and λ, we can write this transformation as

$$dG^{(n)} = \frac{\partial G^{(n)}}{\partial M}\delta M + \frac{\partial G^{(n)}}{\partial \lambda}\delta\lambda = n\delta\eta G^{(n)}. \tag{12.38}$$

Rather than writing this relation in terms of $\delta\lambda$ and $\delta\eta$, it is conventional to define the dimensionless parameters

$$\beta \equiv \frac{M}{\delta M}\delta\lambda; \qquad \gamma \equiv -\frac{M}{\delta M}\delta\eta. \tag{12.39}$$

Making these substitutions in Eq. (12.38) and multiplying through by $M/\delta M$, we obtain

$$\left[M\frac{\partial}{\partial M} + \beta\frac{\partial}{\partial\lambda} + n\gamma\right] G^{(n)}(x_1,\cdots,x_n;M,\lambda) = 0. \tag{12.40}$$

The parameters β and γ are the same for every n, and must be independent of the x_i. Since the Green's function $G^{(n)}$ is renormalized, β and γ cannot depend on the cutoff, and hence, by dimensional analysis, these functions cannot depend on M. Therefore they are functions only of the dimensionless variable λ. We conclude that any Green's function of massless ϕ^4 theory must satisfy

$$\left[M\frac{\partial}{\partial M} + \beta(\lambda)\frac{\partial}{\partial\lambda} + n\gamma(\lambda)\right]G^{(n)}(\{x_i\}; M, \lambda) = 0. \tag{12.41}$$

This relation is called the Callan-Symanzik equation.‡ It asserts that there exist two universal functions $\beta(\lambda)$ and $\gamma(\lambda)$, related to the shifts in the coupling constant and field strength, that compensate for the shift in the renormalization scale M.

The preceding argument generalizes without difficulty to other massless theories with dimensionless couplings. In theories with multiple fields and couplings, there is a γ term for each field and a β term for each coupling. For example, we can define QED at zero electron mass by introducing a renormalization scale as in Eqs. (12.30). The renormalization conditions for the propagators are applied at $p^2 = -M^2$, and those for the vertex at a point where all three invariants are of order $-M^2$. Then the renormalized Green's functions of this theory satisfy the Callan-Symanzik equation

$$\left[M\frac{\partial}{\partial M} + \beta(e)\frac{\partial}{\partial e} + n\gamma_2(e) + m\gamma_3(e)\right]G^{(n,m)}(\{x_i\}; M, e) = 0, \tag{12.42}$$

where n and m are, respectively, the number of electron and photon fields in the Green's function $G^{(n,m)}$ and γ_2 and γ_3 are the rescaling functions of the electron and photon fields.

Computation of β and γ

Before we work out the implications of the Callan-Symanzik equation, let us look more closely at the functions β and γ that appear in it. From their definitions (12.39), we see that they are proportional to the shift in the coupling constant and the shift in the field normalization, respectively, when the renormalization scale M is increased. The behavior of the coupling constant as a function of M is of particular interest, since it determines the strength of the interaction and the conditions under which perturbation theory is valid. We will see in the next section that the shift in the field strength is also reflected directly in the values of Green's functions.

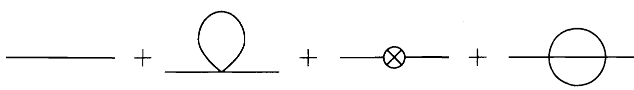
The easiest way to compute the Callan-Symanzik functions is to begin with explicit perturbative expressions for some conveniently chosen Green's functions. If we insist that these expressions satisfy the Callan-Symanzik equation, we will obtain equations that can be solved for β and γ. Because the

‡C. G. Callan, *Phys. Rev.* **D2**, 1541 (1970), K. Symanzik, *Comm. Math. Phys.* **18**, 227 (1970).

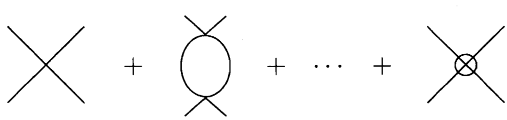
M dependence of a renormalized Green's function originates in the counterterms that cancel its logarithmic divergences, we will find that the β and γ functions are simply related to these counterterms, or equivalently, to the coefficients of the divergent logarithms. The precise formulae that relate β and γ to the counterterms will depend on the specific renormalization prescription and other details of the calculational scheme. At one-loop order, however, the expressions for β and γ are simple and unambiguous.

As a first example, let us calculate the one-loop contributions to $\beta(\lambda)$ and $\gamma(\lambda)$ in massless ϕ^4 theory. We can simplify the analysis by working in momentum space rather than coordinate space. Our strategy will be to apply the Callan-Symanzik equation to the diagrammatic expressions for the two- and four-point Green's functions.

The two-point function is given by

$G^{(2)}(p) =$  $+ \cdots .$

In massless ϕ^4 theory, the one-loop propagator correction is completely canceled by the mass counterterm. Then the first nontrivial correction to the propagator comes from the two-loop diagram and its counterterm, and is of order λ^2. Meanwhile, the four-point function is given by

$G^{(4)} =$  $+ \quad \mathcal{O}(\lambda^3)\,,$

where we have omitted the canceled one-loop propagator corrections to the external legs. The diagrams of order λ^3 include nonvanishing two-loop propagator corrections to the external legs.

To calculate β, we apply the Callan-Symanzik equation to the four-point function:

$$\Big[M\frac{\partial}{\partial M} + \beta(\lambda)\frac{\partial}{\partial\lambda} + 4\gamma(\lambda)\Big]G^{(4)}(p_1,\ldots,p_4) = 0. \tag{12.43}$$

Borrowing our result (10.21) from Section 10.2, we can write $G^{(4)}$ as

$$G^{(4)} = \big[-i\lambda + (-i\lambda)^2\big[iV(s) + iV(t) + iV(u)\big] - i\delta_\lambda\big]\cdot \prod_{i=1,\ldots,4}\frac{i}{p_i^2},$$

where $V(s)$ represents the loop integral in (10.20). Our renormalization condition (12.30) requires that the correction terms cancel at $s = t = u = -M^2$. The order-λ^2 vertex counterterm is therefore

$$\delta_\lambda = (-i\lambda)^2\cdot 3V(-M^2) = \frac{3\lambda^2}{2(4\pi)^{d/2}}\int_0^1 dx\,\frac{\Gamma(2-\frac{d}{2})}{(x(1-x)M^2)^{2-d/2}}. \tag{12.44}$$

The last expression follows from setting $m = 0$ and $p^2 = -M^2$ in Eq. (10.23) for $V(p^2)$. In the limit as $d \to 4$, Eq. (12.44) becomes

$$\delta_\lambda = \frac{3\lambda^2}{2(4\pi)^2}\left[\frac{1}{2-d/2} - \log M^2 + \text{finite}\right], \tag{12.45}$$

where the finite terms are independent of M. This counterterm gives $G^{(4)}$ its M dependence:

$$M\frac{\partial}{\partial M}G^{(4)} = \frac{3i\lambda^2}{(4\pi)^2}\prod_i \frac{i}{p_i^2}.$$

Let us assume for the moment that $\gamma(\lambda)$ has no term of order λ; we will justify this in the next paragraph. Then the Callan-Symanzik equation (12.43) can be satisfied to order λ^2 only if the β function of ϕ^4 theory is given by

$$\beta(\lambda) = \frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3). \tag{12.46}$$

Next, consider the Callan-Symanzik equation for the two-point function:

$$\left[M\frac{\partial}{\partial M} + \beta(\lambda)\frac{\partial}{\partial\lambda} + 2\gamma(\lambda)\right]G^{(2)}(p) = 0. \tag{12.47}$$

Since, to one-loop order, there are no propagator corrections to $G^{(2)}$, no dependence on M or λ is introduced to order λ. Thus the γ function is zero to this order:

$$\gamma = 0 + \mathcal{O}(\lambda^2). \tag{12.48}$$

This justifies the assumption made in the previous paragraph. The two-loop propagator correction is divergent, and its counterterm contains a term of order λ^2 which depends on M. This contributes to the first term in Eq. (12.47). Since β is of order λ^2 and the corrections to $G^{(2)}$ are of order λ^2, the leading contributions to the second term in (12.47) are of order λ^3. Thus γ acquires a nonzero contribution in order λ^2. This leading contribution to γ is computed in Problem 13.2.

The preceding example illustrates how β and γ can be calculated in more general theories with dimensionless couplings. In such theories, the M dependence of Green's functions enters through the field-strength and vertex counterterms, which are used to subtract the divergent logarithms. The lowest-order expressions for β and γ can be computed directly from these counterterms, or from the coefficients of the divergent logarithms.

In any renormalizable massless scalar field theory, the two-point Green's function has the generic form

$$\begin{aligned} G^{(2)}(p) &= \text{———} \quad + \quad \text{(loop diagrams)} \quad + \quad \text{—}\otimes\text{—} \quad + \cdots \\ &= \frac{i}{p^2} + \frac{i}{p^2}\left(A\log\frac{\Lambda^2}{-p^2} + \text{finite}\right) + \frac{i}{p^2}(ip^2\delta_Z)\frac{i}{p^2} + \cdots. \end{aligned} \tag{12.49}$$

The M dependence of this expression, to lowest order, comes entirely from the counterterm δ_Z. Applying the Callan-Symanzik equation to $G^{(2)}(p)$, and neglecting the β term (which is always smaller by at least one power of the coupling constant), we find

$$-\frac{i}{p^2} M \frac{\partial}{\partial M} \delta_Z + 2\gamma \frac{i}{p^2} = 0,$$

or

$$\gamma = \frac{1}{2} M \frac{\partial}{\partial M} \delta_Z \qquad \text{(to lowest order)}. \tag{12.50}$$

To make this result more explicit, note that the counterterm must be

$$\delta_Z = A \log \frac{\Lambda^2}{M^2} + \text{finite}$$

in order to cancel the divergent logarithm in $G^{(2)}$. Thus γ is simply the coefficient of the logarithm:

$$\gamma = -A \qquad \text{(to lowest order)}. \tag{12.51}$$

In most theories (e.g., Yukawa theory or QED), the first logarithmic divergence in δ_Z occurs at the one-loop level. However, even in ϕ^4 theory, formulae (12.50) and (12.51) are true for the first nonvanishing term in δ_Z, in this case the two-loop contribution.* By replacing the scalar field propagator (i/p^2) with a fermion propagator $(i/\not{p})$, we could repeat this argument line for line to compute the γ function for a fermion field in terms of its field strength counterterm δ_Z.

We can derive similar expressions for the β function of a generic dimensionless coupling constant g, associated with an n-point vertex. Taking propagator corrections into account, the full connected Green's function, to one-loop order, has the general form

$$\begin{aligned} G^{(n)} &= \begin{pmatrix} \text{tree-level} \\ \text{diagram} \end{pmatrix} + \begin{pmatrix} \text{1PI loop} \\ \text{diagrams} \end{pmatrix} + \begin{pmatrix} \text{vertex} \\ \text{counterterm} \end{pmatrix} + \begin{pmatrix} \text{external leg} \\ \text{corrections} \end{pmatrix} \\ &= \Big(\prod_i \frac{i}{p_i^2}\Big)\Big[-ig - iB \log \frac{\Lambda^2}{-p^2} - i\delta_g + (-ig) \sum_i \Big(A_i \log \frac{\Lambda^2}{-p_i^2} - \delta_{Zi}\Big)\Big] \\ &\qquad + \text{finite terms}. \end{aligned} \tag{12.52}$$

In this expression, p_i are the momenta on the external legs, and p^2 represents a typical invariant built from these momenta. We assume that renormalization conditions are applied at a point where all such invariants are spacelike and of order $-M^2$. The M dependence of this expression comes from the counterterms δ_g and δ_{Zi}. Applying the Callan-Symanzik equation, we obtain

$$M \frac{\partial}{\partial M}\Big(\delta_g - g \sum_i \delta_{Zi}\Big) + \beta(g) + g \sum_i \frac{1}{2} M \frac{\partial}{\partial M} \delta_{Zi} = 0,$$

*At one loop, formula (12.33) implies that we can also identify A as the coefficient of $2/(4-d)$ in the 1PI self-energy, in the limit $d \to 4$. This relation changes in higher loops. However, Eq. (12.50) remains correct.

or

$$\beta(g) = M\frac{\partial}{\partial M}\Big(-\delta_g + \frac{1}{2}g\sum_i \delta_{Zi}\Big) \qquad \text{(to lowest order).} \tag{12.53}$$

To be more explicit, we note that

$$\delta_g = -B\log\frac{\Lambda^2}{M^2} + \text{finite.}$$

Thus the β function is just a combination of the coefficients of the divergent logarithms:

$$\beta(g) = -2B - g\sum_i A_i \qquad \text{(to lowest order).} \tag{12.54}$$

Notice that the finite parts of counterterms are independent of M and therefore never contribute to β or γ. This means that, to compute the leading terms in the Callan-Symanzik functions, we needn't be too precise in specifying renormalization conditions: Any momentum scale of order M^2 will yield the same results. The divergent parts of the counterterms can be estimated simply by setting all invariants inside of logarithms equal to M^2, as we did above in our expression for the n-point Green's function.

As in the computation of γ, this argument can be applied almost without change to coupling constants for fields with spin. In Yukawa theory, for example, we consider the three-point function with one incoming fermion, one outgoing fermion, and one scalar, with momenta p_1, p_2, and p_3, respectively. Then the tree-level expression for the three-point function is

$$\frac{i}{\not p_1}\frac{i}{\not p_2}\frac{1}{p_3^2}(-ig). \tag{12.55}$$

The one-loop corrections replace the quantity $(-ig)$ by the expression in brackets in Eq. (12.52). Then formulae (12.53) and (12.54) hold also for the β function of this theory.

Similar expressions also apply in QED, though there are a number of small complications. The first comes in computing the γ function for the photon propagator. In Eq. (7.74), we saw that the general form of the photon propagator in Feynman gauge is

$$D^{\mu\nu}(q) = D(q)\Big(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\Big) + \frac{-i}{q^2}\frac{q^\mu q^\nu}{q^2}. \tag{12.56}$$

The coefficient of the last term in (12.56) depends on the gauge. Fortunately, this term drops out of all gauge-invariant observables. Thus it makes sense to concentrate on the first term, projecting all external photons onto their transverse components. Projecting the photon propagator, we see that $D(q)$ satisfies the Callan-Symanzik equation. Since the corrections to this function have the form (12.49), the arguments following that formula are valid for photons as well as for electrons and scalars. Thus, to leading order,

$$\gamma_2 = \frac{1}{2}M\frac{\partial}{\partial M}\delta_2, \qquad \gamma_3 = \frac{1}{2}M\frac{\partial}{\partial M}\delta_3, \tag{12.57}$$

where δ_2 and δ_3 are the counterterms defined in Section 10.3.

Similarly, we may consider the three-point connected Green's function $\langle\bar{\psi}(p_1)\psi(p_2)A_\mu(q)\rangle$, projected onto transverse components of the photon. At leading order, this function equals

$$\frac{i}{\not{p}_1}(-ie\gamma^\mu)\frac{i}{\not{p}_2}\frac{-i}{q^2}\Big(g^{\mu\nu}-\frac{q^\mu q^\nu}{q^2}\Big).$$

The divergent one-loop corrections have the same form, with $(-ie)$ replaced by logarithmically divergent terms. Thus, Eq. (12.53) gives the lowest-order expression for the β function:

$$\beta(e)=M\frac{\partial}{\partial M}\big(-e\delta_1+e\delta_2+\frac{e}{2}\delta_3\big). \tag{12.58}$$

To find explicit expressions for the Callan-Symanzik functions of QED, we must write expressions for the counterterms δ_1, δ_2, δ_3. In Section 10.3, we evaluated these counterterms using on-shell renormalization conditions with massive fermions. We must now re-evaluate these terms for massless fermions and renormalization at $-M^2$. Fortunately, we need only evaluate the logarithmically divergent pieces of these counterterms, which are identical in the two cases. Reading from Eqs. (10.43) and (10.44), we find

$$\begin{aligned}\delta_1=\delta_2&=-\frac{e^2}{(4\pi)^2}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}+\text{ finite},\\ \delta_3&=-\frac{e^2}{(4\pi)^2}\frac{4}{3}\frac{\Gamma(2-\frac{d}{2})}{(M^2)^{2-d/2}}+\text{ finite}.\end{aligned} \tag{12.59}$$

Using formulae (12.57) and (12.59), we obtain at leading order

$$\gamma_2(e)=\frac{e^2}{16\pi^2},\qquad \gamma_3(e)=\frac{e^2}{12\pi^2}. \tag{12.60}$$

And from Eq. (12.58), we find

$$\beta(e)=\frac{e^3}{12\pi^2}. \tag{12.61}$$

It is important to remember that the expression we have used for δ_2 explicitly assumes the use of Feynman gauge. In fact, γ_2 depends on the gauge parameter, and this makes sense, because Green's functions of individual ψ and $\bar{\psi}$ fields are not gauge invariant. On the other hand, the QED vacuum polarization, and therefore γ_3 and β, are gauge invariant.

The Meaning of β and γ

We can obtain a deeper insight into the nature of β and γ by expressing them in terms of the parameters of bare perturbation theory: Z, λ_0, and Λ for the case of ϕ^4 theory.

First recall that the bare and renormalized field are related by

$$\phi(p) = Z(M)^{-1/2}\phi_0(p). \tag{12.62}$$

This equation expresses the dependence of the field rescaling on M. If M is increased by δM, the renormalized field is shifted by

$$\delta\eta = \frac{Z(M+\delta M)^{-1/2}}{Z(M)^{-1/2}} - 1.$$

Hence our original definition (12.39) of γ gives us immediately

$$\gamma(\lambda) = \frac{1}{2}\frac{M}{Z}\frac{\partial}{\partial M}Z. \tag{12.63}$$

Since $\delta_Z = Z - 1$ (Eq. (10.17)), this formula is in agreement with (12.50) to leading order. Formula (12.63), however, is an exact relation. This expression clarifies the relation of γ to the field strength rescaling. However, it obscures the fact that γ is independent of the cutoff Λ. To understand this aspect of γ, we have to go back to the original definition of this function in terms of renormalized Green's functions, whose cutoff independence follows from the renormalizability of the theory.

Similarly, we can find an instructive expression for β in terms of the parameters of bare perturbation theory. Our original definition of β in Eq. (12.39) made use of a quantity $\delta\lambda$, defined to be the shift of the renormalized coupling λ needed to preserve the values of the bare Green's functions when the renormalization point is shifted infinitesimally. Since the bare Green's functions depend on the bare coupling λ_0 and the cutoff, this definition can be rewritten as

$$\beta(\lambda) = M\frac{\partial}{\partial M}\lambda\bigg|_{\lambda_0,\Lambda}. \tag{12.64}$$

Thus the β function is the rate of change of the renormalized coupling at the scale M corresponding to a fixed bare coupling. Recalling our analysis in Section 12.1, it is tempting to associate $\lambda(M)$ with the coupling constant λ' obtained by integrating out degrees of freedom down to the scale M. With this correspondence, the β function is just the rate of the renormalization group flow of the coupling constant λ. A positive sign for the β function indicates a renormalized coupling that increases at large momenta and decreases at small momenta. We can see explicitly that this relation works for ϕ^4 theory, to leading order in λ, by comparing Eqs. (12.28) and (12.46). We will justify this correspondence further in the following section.

The equality of the exact formula (12.64) with the first-order formula (12.53) again follows from the counterterm definitions (10.17). As with (12.63), it is not obvious that this formula for $\beta(\lambda)$ is independent of Λ, but that fact again follows from renormalizability. Conversely, it is possible to prove the renormalizability of ϕ^4 theory by demonstrating, order by order in perturbation theory, that expressions (12.63) and (12.64) are independent of Λ.†

12.3 Evolution of Coupling Constants

Now that we have discussed all of the ingredients of the Callan-Symanzik equation, let us investigate its implications. We begin by finding the explicit solution to the Callan-Symanzik equation for the simplest situation, the two-point Green's function of a scalar field theory. This solution will clarify the physical implications of the equation. In particular, it will cement the relation suggested at the end of the previous section, which identifies the β function with the rate of the renormalization group flow of the coupling constant. We will then use this relation to discuss the qualitative features of the renormalization group flow in renormalizable field theories.

Solution of the Callan-Symanzik Equation

We would like to solve the Callan-Symanzik equation for the two-point Green's function, $G^{(2)}(p)$, in a theory with a single scalar field. Since $G^{(2)}(p)$ has dimensions of $(\text{mass})^{-2}$, we can express its dependence on p and M as

$$G^{(2)}(p) = \frac{i}{p^2} g(-p^2/M^2). \tag{12.65}$$

This equation allows us to trade the derivative with respect to M for a derivative with respect to p^2. For the remainder of this chapter, we will use the variable p to represent the magnitude of the spacelike momentum: $p = (-p^2)^{1/2}$. Then we can rewrite the Callan-Symanzik equation as

$$\left[p\frac{\partial}{\partial p} - \beta(\lambda)\frac{\partial}{\partial \lambda} + 2 - 2\gamma(\lambda)\right] G^{(2)}(p) = 0. \tag{12.66}$$

In free field theory, β and γ vanish and we recover the trivial result

$$G^{(2)}(p) = \frac{i}{p^2}. \tag{12.67}$$

In an interacting theory, β and γ are nonzero functions of λ. However, it is still possible to write the explicit solution to the Callan-Symanzik equation, using the method of characteristics. Equivalently (for those not well versed in the theory of partial differential equations), we will apply a lovely hydrodynamic-bacteriological analogy due to Sidney Coleman.‡ Imagine a narrow pipe running in the x direction, containing a fluid whose velocity

†Callan has given a beautiful proof of the renormalizability of ϕ^4 theory, based on proving that the Callan-Symanzik equation holds order by order in λ, in his article in *Methods in Field Theory*, R. Balian and J. Zinn-Justin, eds. (North Holland, Amsterdam, 1976).

‡Coleman (1985), chap. 3.

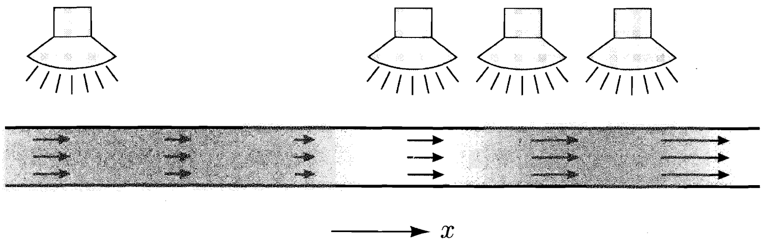


Figure 12.3. Coleman's bacteriological analogy to the Callan-Symanzik equation. The pipe is inhabited by bacteria with a given initial density $D_i(x)$. The growth rate (determined by the illumination) and flow velocity are given functions of x. The problem is to determine the density $D(t,x)$ at all subsequent times.

is $v(x)$, as shown in Fig. 12.3. The pipe is inhabited by bacteria, whose density is $D(t,x)$ and whose rate of growth is $\rho(x)$. Then the future behavior of the function $D(t,x)$ is governed by the differential equation

$$\left[\frac{\partial}{\partial t} + v(x)\frac{\partial}{\partial x} - \rho(x)\right] D(t,x) = 0. \tag{12.68}$$

The second term allows for the fact that the bacteria are swept along with the fluid, so their present density here determines their future density not here, but some distance ahead. This equation is identical to Eq. (12.66), with the replacements

$$\begin{aligned} \log(p/M) &\leftrightarrow t, \\ \lambda &\leftrightarrow x, \\ -\beta(\lambda) &\leftrightarrow v(x), \\ 2\gamma(\lambda){-}2 &\leftrightarrow \rho(x), \\ G^{(2)}(p,\lambda) &\leftrightarrow D(t,x). \end{aligned} \tag{12.69}$$

Now suppose we know the initial concentration of the bacteria: $D(t,x) = D_i(x)$ at time $t = 0$. Then we can determine the concentration of bacteria in a fluid element at the point x at any later time by computing the history of that fluid element and then integrating the rate of growth along that path. Consider the fluid element that is at x at the time t. We can find out where it was at time zero by integrating its motion backward in time. The position of this element at time $t = 0$ is given by $\bar{x}(t;x)$, which satisfies the differential equation

$$\frac{d}{dt'}\bar{x}(t';x) = -v(\bar{x}), \qquad \text{with} \qquad \bar{x}(0;x) = x. \tag{12.70}$$

Then, immediately,

$$\begin{aligned} D(t,x) &= D_i\big(\bar{x}(t;x)\big)\cdot\exp\biggl(\int_0^t dt'\,\rho\big(\bar{x}(t';x)\big)\biggr) \\ &= D_i\big(\bar{x}(t;x)\big)\cdot\exp\biggl(\int_{\bar{x}(t)}^{x} dx'\,\frac{\rho(x')}{v(x')}\biggr). \end{aligned} \tag{12.71}$$

Now bring this solution back to our field theory problem by replacing each bacteriological parameter with its corresponding field theory parameter. The time $t = 0$ corresponds to $-p^2 = M^2$, and the initial concentration $D_i(x)$ becomes an unknown function $\hat{\mathcal{G}}(\lambda)$. Then

$$G^{(2)}(p,\lambda) = \hat{\mathcal{G}}\big(\bar{\lambda}(p;\lambda)\big)\cdot\exp\biggl(-\int_{p'=M}^{p'=p} d\log(p'/M)\cdot 2\big[1-\gamma\big(\bar{\lambda}(p';\lambda)\big)\big]\biggr), \tag{12.72}$$

where $\bar{\lambda}(p;\lambda)$ solves

$$\frac{d}{d\log(p/M)}\bar{\lambda}(p;\lambda) = \beta(\bar{\lambda}), \qquad \bar{\lambda}(M;\lambda) = \lambda. \tag{12.73}$$

This differential equation describes the flow of a modified coupling constant $\bar{\lambda}(p;\lambda)$ as a function of momentum. The rate of this flow is just the β function. Thus, this flow is strongly reminiscent of the dependence of the renormalized coupling on the renormalization scale given by Eq. (12.64). We will refer to $\bar{\lambda}(p)$ as the *running coupling constant.* Its equation (12.73) is often called the *renormalization group equation.*

One can check directly that (12.72) solves the Callan-Symanzik equation by using the identity

$$\int_{\lambda}^{\bar{\lambda}} \frac{d\lambda'}{\beta(\lambda')} = \int_{p'=M}^{p'=p} d\log(p'/M), \tag{12.74}$$

from which it follows that

$$\Big(p\frac{\partial}{\partial p} - \beta(\lambda)\frac{\partial}{\partial\lambda}\Big)\bar{\lambda} = 0. \tag{12.75}$$

A convenient way of writing the solution (12.72) is

$$G^{(2)}(p,\lambda) = \frac{i}{p^2}\mathcal{G}\big(\bar{\lambda}(p;\lambda)\big)\cdot\exp\biggl(2\int_M^p d\log(p'/M)\,\gamma\big(\bar{\lambda}(p';\lambda)\big)\biggr), \tag{12.76}$$

in which $\mathcal{G}(\bar{\lambda})$ is a function that must be determined. This function cannot be determined from the general principles of renormalization theory. Instead, we must compute $G^{(2)}(p)$ as a perturbation series in λ and match terms to the expansion of (12.76) as a series in the same parameter. For the two-point function in ϕ^4 theory, this matching is rather trivial: $\mathcal{G}(\bar{\lambda}) = 1 + \mathcal{O}(\bar{\lambda}^2)$.

The preceding analysis can be applied to any family of Green's functions that are related by uniform rescaling of the momenta. Consider, for example, the connected four-point function of ϕ^4 theory evaluated at spacelike momenta p_i such that $p_i^2 = -P^2$, $p_i \cdot p_j = 0$, so that s, t, and u are of order $-P^2$. To leading order in perturbation theory, this function is given by

$$G^{(4)}(P) = \left(\frac{i}{P^2}\right)^4 (-i\lambda). \tag{12.77}$$

Using the fact that $G^{(4)}$ has dimensions of $(\text{mass})^{-8}$, we can exchange M for P in the Callan-Symanzik equation and write this equation as

$$\left[P\frac{\partial}{\partial P} - \beta(\lambda)\frac{\partial}{\partial \lambda} + 8 - 4\gamma(\lambda)\right] G^{(4)}(P;\lambda) = 0. \tag{12.78}$$

The solution to this equation is

$$G^{(4)}(P;\lambda) = \frac{1}{P^8}\mathcal{G}^{(4)}(\bar{\lambda}(p;\lambda)) \cdot \exp\left(4\int_M^p d\log(p'/M)\, \gamma(\bar{\lambda}(p';\lambda))\right). \tag{12.79}$$

This formula must agree with (12.77) to leading order in λ; this matching requires that

$$\mathcal{G}^{(4)}(\bar{\lambda}(p;\lambda)) = -i\bar{\lambda} + \mathcal{O}(\bar{\lambda}^2). \tag{12.80}$$

We can now see the physical implication of the Callan-Symanzik equation. The ordinary Feynman perturbation series for a Green's function depends both on the coupling constant λ and on the dimensionless parameter $\log(-p^2/M^2)$. The perturbation theory can be badly behaved even when λ is small if the ratio p^2/M^2 is large. The solutions (12.76) and (12.79) reorganize this dependence into a function of the running coupling constant and an exponential scale factor. We consider these two pieces in turn.

The first factor in Eqs. (12.76) and (12.79) is a function of the running coupling constant, evaluated at the momentum scale p. If p were of order M, the renormalization scale, this function would essentially be the ordinary perturbative evaluation of the Green's function. The results (12.76) and (12.79) instruct us to make use of this same expression at the scale p, but to replace λ with a new coupling constant $\bar{\lambda}$ appropriate to that scale. Thus, the running coupling constant $\bar{\lambda}(p)$ is precisely the effective coupling constant of the renormalization group flow. This interpretation is particularly clear in the solution (12.79) for $G^{(4)}(P)$, since this function directly measures the strength of the ϕ^4 coupling constant.

The exponential factor in Eqs. (12.76) and (12.79) has an equally simple interpretation: It is the accumulated field strength rescaling of the correlation function from the reference point M to the actual momentum p at which the

Green's function is evaluated. This factor receives a multiplicative contribution from each intermediate scale between M and p. Each of these contributions is, appropriately, computed using the running coupling constant at that particular scale.

As a check on these formal arguments, we can use the explicit form of the β function of ϕ^4 theory found in Eq. (12.46) and the renormalization group equation (12.73) to evaluate the running coupling constant of ϕ^4 theory. This running coupling constant satisfies the differential equation

$$\frac{d}{d\log(p/M)}\bar{\lambda} = \frac{3\bar{\lambda}^2}{16\pi^2}, \qquad \text{with} \qquad \bar{\lambda}(M;\lambda) = \lambda. \tag{12.81}$$

Integrating, we find

$$\left(\frac{3}{16\pi^2}\right)^{-1}\left[\frac{1}{\lambda} - \frac{1}{\bar{\lambda}}\right] = \log\frac{p}{M},$$

and thus,

$$\bar{\lambda}(p) = \frac{\lambda}{1 - (3\lambda/16\pi^2)\log(p/M)}. \tag{12.82}$$

Many properties of the solution to the Callan-Symanzik equation are visible in this relation. First, the expansion of this formula for $\bar{\lambda}$ to order λ^2 agrees precisely with Eq. (12.28), the rate of the renormalization group flow from Wilson's method. Second, this expression for the running coupling constant goes to zero at a logarithmic rate as $p \to 0$. This coincides with our expectation that a positive value for the β function should imply an effective coupling that becomes stronger at large momenta and weaker at small momenta.

If we expand the running coupling constant $\bar{\lambda}(p)$ in powers of λ, we find that the successive powers of the coupling constant are multiplied by powers of logarithms,

$$\lambda^{n+1}(\log p/M)^n,$$

which become large and invalidate a simple perturbation expansion for p much greater or much less than M. We have seen this problem of large logarithms arising several times in our diagram calculations, and we have remarked on it specifically as a problem in the discussion following Eq. (11.81). We now see that the renormalization group gives a partial solution to this problem. In this example, and in many others that we will study, the Callan-Symanzik equation tells us how to sum these large logarithms into the running coupling constant and multiplicative rescalings. If the running coupling constant becomes large, as happens in ϕ^4 theory for $p \to \infty$, the perturbation expansion will break down anyway, and we will need more advanced methods. However, if the running coupling constant becomes small, as for ϕ^4 theory as $p \to 0$, we will have successfully organized the powers of logarithms into a meaningful and controlled expression. The specific problem posed at the end of Section 11.4 will be solved explicitly by this method in Section 13.2.

An Application to QED

For a more concrete application of the Callan-Symanzik equation, we can look again at the electromagnetic potential between static charges, $V(\mathbf{x})$, which we studied in Section 7.5. At very short distances or at large momenta, we can ignore the electron mass in the computation of QED corrections to this potential. In this approximation, the potential should obey the Callan-Symanzik equation of massless QED. We could write this equation either for $V(\mathbf{x})$ itself or for its Fourier transform; we choose to work in Fourier space in order to make contact more easily with the results of Section 7.5.

We define the massless limit of QED by specifying a renormalization scale M at which the renormalized coupling e_r is defined. If M is taken close to the electron mass m, at the point where the massless approximation is just becoming valid, then the value of e_r will be close to the physical electron charge e. The potential between static charges is a measurable energy, so its normalization is unambiguous and is not shifted from one renormalization point to another. Thus the Callan-Symanzik equation for the Fourier transform of the potential has no γ term, being simply

$$\left[M\frac{\partial}{\partial M} + \beta(e_r)\frac{\partial}{\partial e_r}\right]V(q; M, e_r) = 0. \tag{12.83}$$

The Fourier transform of the potential has dimensions of $(\text{mass})^{-2}$, so we can trade dependence on M for dependence on q as in the scalar field theory discussion above. This gives

$$\left[q\frac{\partial}{\partial q} - \beta(e_r)\frac{\partial}{\partial e_r} + 2\right]V(q; M, e_r) = 0. \tag{12.84}$$

Equation (12.84) is almost the same as Eq. (12.66), so we can immediately write down the solution as a special case of (12.76):

$$V(q, e_r) = \frac{1}{q^2}\mathcal{V}\big(\bar{e}(q; e_r)\big), \tag{12.85}$$

where $\bar{e}(q)$ is the solution of the renormalization group equation

$$\frac{d}{d\log(q/M)}\bar{e}(q; e_r) = \beta(\bar{e}), \qquad \bar{e}(M; e_r) = e_r. \tag{12.86}$$

By comparing this formula for $V(q)$ to the leading-order result

$$V(q) \approx \frac{e^2}{q^2},$$

we can identify $\mathcal{V}(\bar{e}) = \bar{e}^2 + \mathcal{O}(\bar{e}^4)$. Then

$$V(q, e_r) = \frac{\bar{e}^2(q; e_r)}{q^2}, \tag{12.87}$$

up to corrections that are suppressed by powers of e_r^2 and contain no compensatory large logarithms of q/M.

To turn Eq. (12.87) into a completely explicit formula, we need only solve the renormalization group equation (12.86). Using the QED β function (12.61), we can integrate (12.86) to find

$$\frac{12\pi^2}{2}\left(\frac{1}{e_r^2}-\frac{1}{\bar{e}^2}\right)=\log\frac{q}{M}.$$

This simplifies to

$$\bar{e}^2(q)=\frac{e_r^2}{1-(e_r^2/6\pi^2)\log(q/M)}. \tag{12.88}$$

This result is almost identical to the formula for the effective electric charge that we found in Eq. (7.96). To cement the identification, set M to be of order the electron mass, $M^2=Am^2$, and approximate e_r at this point by e, with $\alpha=e^2/4\pi$. Then Eq. (12.88) takes the form

$$\bar{\alpha}(q)=\frac{\alpha}{1-(\alpha/3\pi)\log(-q^2/Am^2)}. \tag{12.89}$$

The particular choice $A=\exp(5/3)$ reproduces Eq. (7.96). Of course, we could not find this exact correspondence without the detailed one-loop calculation of Section 7.5. Nevertheless, our present analysis produces the correct asymptotic formula for the effective charge. Furthermore, our present formalism can be applied to any renormalizable quantum field theory; it does not rely on the special symmetries of QED that we exploited in Section 7.5.

Alternatives for the Running of Coupling Constants

Now that we have computed the behavior of the running coupling constant in two specific quantum field theories, let us consider more generally what behaviors of the running coupling constant are possible in principle. We continue to restrict our discussion to renormalizable theories in the massless limit, with a single dimensionless coupling constant λ.

By the arguments of the previous section, the Green's functions in any such theory obey a Callan-Symanzik equation. The solution of this equation depends on a running coupling constant, $\bar{\lambda}(p)$, which satisfies a differential equation

$$\frac{\partial}{\partial\log(p/M)}\bar{\lambda}=\beta(\bar{\lambda}), \tag{12.90}$$

in which the function $\beta(\lambda)$ is computable as a power series in the coupling constant. In the examples we have just discussed, the leading coefficient in this power series was positive. However, as a matter of principle, three behaviors are possible in the region of small λ:

$$\begin{aligned}&(1)\quad \beta(\lambda)>0;\\ &(2)\quad \beta(\lambda)=0;\\ &(3)\quad \beta(\lambda)<0.\end{aligned}$$

Examples of quantum fields are known that exhibit each of these behaviors.

We have already seen how, in theories of the first class, the running coupling constant goes to zero in the infrared, leading to definite predictions about the small-momentum behavior of the theory. However, the running coupling constant becomes large in the region of high momenta. Thus the short-distance behavior of the theory cannot be computed using Feynman diagram perturbation theory. In fact, in the examples studied above, the coupling constant formally goes to infinity at a large but finite value of the momentum; thus it is not even clear that these theories possess a nontrivial limit $\Lambda \to \infty$. A Feynman diagram analysis is useful in such theories if one is mainly interested in large-distance or macroscopic behavior. In Chapter 13 we will use this observation to solve problems in the statistical mechanics of systems with critical points.

In theories of the second class, the coupling constant does not flow. In these theories, the running coupling constant is independent of the momentum scale, and thus equal to the bare coupling. This means that there can be no ultraviolet divergences in the relation of coupling constants. The only possible ultraviolet divergences in such theories are those associated with field rescaling, which automatically cancel in the computation of S-matrix elements. Such theories are called *finite* quantum field theories. Before the emergence of our modern understanding of renormalization, these theories would have been embraced as the solution to the problem of ultraviolet infinities. But in fact the known finite field theories in four dimensions are very special constructions—the so-called gauge theories with extended supersymmetry—with no known physical application.

In theories of the third class, the running coupling constant becomes large in the large-distance regime and becomes small at large momenta or short distances. Imagine, for instance, that the sign of the QED β function were reversed:

$$\beta(e) = -\tfrac{1}{2}Ce^3. \tag{12.91}$$

Then, following our earlier analysis, we would have

$$\bar{e}^2(p) = \frac{e^2}{1 + Ce^2\log(p/M)}. \tag{12.92}$$

This coupling constant tends to zero at a logarithmic rate as the momentum scale increases. Such theories are called *asymptotically free.* In theories of this class, the short-distance behavior is completely solvable by Feynman diagram methods. Though ultraviolet divergences appear in every order of perturbation theory, the renormalization group tells us that the sum of these divergences is completely harmless. If we interpret these theories in terms of a bare coupling e_b and a finite cutoff Λ, the result (12.92) indicates that there is a smooth limit in which e_b tends to zero as Λ tends to infinity. Thus, asymptotically free theories give another, more sophisticated, resolution of the problem of ultraviolet divergences. In Chapter 17, we will see that asymptotic freedom plays an essential role in the formulation of a field theory that describes the strong interactions of elementary particle physics.

Now that we have enumerated the possibilities for the renormalization group flow in the region of weak coupling, let us turn our attention to the region of strong coupling. Here we will not be able to compute the β function quantitatively, but we can at least use the renormalization group equation to discuss qualitatively the possibilities for the coupling constant flow. All of our explicit solutions for running coupling constants—Eqs. (12.82), (12.88), and (12.92)—predict that the running coupling becomes infinite at a finite value of the momentum p. For example, according to Eq. (12.82), the running coupling constant of ϕ^4 theory should diverge at

$$p \sim M \exp\left(\frac{16\pi^2}{3\lambda}\right). \tag{12.93}$$

It is possible that this is the true behavior of the quantum field theory, but we have not proved this, because when the running coupling constant becomes large, the approximation we have made, ignoring the higher-order terms in the β function, is no longer valid. It is a logical possibility that the higher terms of the β function are negative, so that the β function has the form shown in Fig. 12.4(a). In this case the β function has a zero at a nonzero value λ_*. When $\bar{\lambda}$ approaches this value, the renormalization group flow slows to a halt; thus $\lambda = \lambda_*$ would be a nontrivial fixed point of the renormalization group. In this model, the running coupling constant $\bar{\lambda}$ tends to λ_* in the limit of large momentum.

For the specific case of ϕ^4 theory in four dimensions, we have strong evidence from numerical studies that there is no such nontrivial fixed point. However, we will soon demonstrate that there is a nontrivial fixed point in ϕ^4 theory in $d < 4$, and many more examples are known. It is thus worthwhile to explore the implications of a fixed point in the renormalization group flow.

For a β function of the form of Fig. 12.4(a), the β function behaves in the vicinity of the fixed point as

$$\beta \approx -B(\lambda - \lambda_*), \tag{12.94}$$

where B is a positive constant. For $\bar{\lambda}$ near λ_*,

$$\frac{d}{d\log p}\bar{\lambda} \approx -B(\bar{\lambda} - \lambda_*). \tag{12.95}$$

The solution of this equation is

$$\bar{\lambda}(p) = \lambda_* + C\left(\frac{M}{p}\right)^B. \tag{12.96}$$

Thus, $\bar{\lambda}$ indeed tends to λ_* as $p \to \infty$, and the rate of approach is governed by the slope of the β function at the fixed point.

This behavior has a dramatic consequence for the exact solution (12.72) of the Callan-Symanzik equation for $G(p)$. For p sufficiently large, the integral

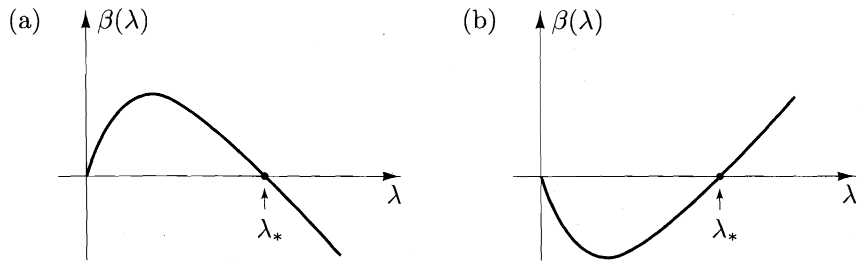


Figure 12.4. Possible forms of the β function with nontrivial zeros: (a) ultraviolet-stable fixed point; (b) infrared-stable fixed point.

in the exponential factor in this equation will be dominated by values of p for which $\bar{\lambda}(p)$ is close to λ_*. Then

$$\begin{aligned} G(p) &\approx \mathcal{G}(\lambda_*) \exp\Big[-\Big(\log\frac{p}{M}\Big)\cdot 2\big(1-\gamma(\lambda_*)\big)\Big] \\ &\approx C\cdot\Big(\frac{1}{p^2}\Big)^{1-\gamma(\lambda_*)}. \end{aligned} \tag{12.97}$$

Thus the two-point correlation function returns to the form of a simple scaling law, but with a power law different from that expected by dimensional analysis. At the fixed point we have a scale-invariant quantum field theory in which the interactions of the theory affect the law of rescaling. The shift of the exponent $\gamma(\lambda_*)$ is called the *anomalous dimension* of the scalar field. By convention, the function $\gamma(\lambda)$ is often called the anomalous dimension even if there is no fixed point in the theory.

A similar behavior is possible in an asymptotically free theory. If the β function has the form shown in Fig. 12.4(b), the running coupling constant will tend to a fixed point λ_* as $p \to 0$. The two-point correlation function of fields $G(p)$ will tend to a power law as in (12.97) for asymptotically small momenta. The two cases shown in Figs. 12.4(a) and (b) are called, respectively, *ultraviolet-stable* and *infrared-stable* fixed points.

In the previous section, we saw that the leading-order expressions for the Callan-Symanzik functions β and γ are related in a simple way to the ultraviolet divergent parts of the one-loop counterterms. However, we noted that, in higher orders of perturbation theory, β and γ depend on the specific renormalization conventions used to define the Green's functions. Still, there are some properties of these functions that are independent of any convention. The coefficient of the logarithm in the denominator of such expressions as (12.82) or (12.89) can be determined unambiguously from experiments that measure this coupling constant. This confirms the convention independence of the first β-function coefficient. Experiments sensitive to the coupling constant can also determine the existence of a zero of the β function at strong coupling, and the rate of approach to this asymptote. Thus the existence of a zero of