

In the reference frame where the particle is initially at rest, its momentum vector is $p^\mu = (p^0, \mathbf{0})$ and the vector potential reduces to

$$A^\mu(x) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \frac{e}{|\mathbf{k}|^2} \cdot (1, \mathbf{0}).$$

This is just the Coulomb potential of an unaccelerated charge. As we would expect, there is no radiation field before the particle is scattered.

After scattering ($t > 0$), we close the contour downward, picking up the three poles below the real axis. The pole at $k^0 = \mathbf{k} \cdot \mathbf{p}'/p'^0$ gives the Coulomb potential of the outgoing particle. Thus the other two poles are completely responsible for the radiation field. Their contribution gives

$$\begin{aligned} A_{\text{rad}}^\mu(x) &= \int \frac{d^3k}{(2\pi)^3} \frac{-e}{2|\mathbf{k}|} \left\{ e^{-ik\cdot x} \left(\frac{p'^\mu}{\mathbf{k} \cdot \mathbf{p}'} - \frac{p^\mu}{\mathbf{k} \cdot \mathbf{p}} \right) + \text{c.c.} \right\} \Big|_{k^0=|\mathbf{k}|} \\ &= \text{Re} \int \frac{d^3k}{(2\pi)^3} \mathcal{A}^\mu(\mathbf{k}) e^{-ik\cdot x}, \end{aligned} \quad (6.6)$$

where the momentum-space amplitude $\mathcal{A}(\mathbf{k})$ is given by

$$\mathcal{A}^\mu(\mathbf{k}) = \frac{-e}{|\mathbf{k}|} \left(\frac{p'^\mu}{\mathbf{k} \cdot \mathbf{p}'} - \frac{p^\mu}{\mathbf{k} \cdot \mathbf{p}} \right). \quad (6.7)$$

(The condition $k^0 = |\mathbf{k}|$ is implicit here and in the rest of this calculation.)

To calculate the energy radiated, we must find the electric and magnetic fields. It is easiest to write \mathbf{E} and \mathbf{B} as the real parts of complex Fourier integrals, just as we did for A^μ :

$$\begin{aligned} \mathbf{E}(x) &= \text{Re} \int \frac{d^3k}{(2\pi)^3} \mathcal{E}(\mathbf{k}) e^{-ik\cdot x}; \\ \mathbf{B}(x) &= \text{Re} \int \frac{d^3k}{(2\pi)^3} \mathcal{B}(\mathbf{k}) e^{-ik\cdot x}. \end{aligned} \quad (6.8)$$

The momentum-space amplitudes $\mathcal{E}(\mathbf{k})$ and $\mathcal{B}(\mathbf{k})$ of the radiation fields are then simply

$$\begin{aligned} \mathcal{E}(\mathbf{k}) &= -i\mathbf{k}\mathcal{A}^0(\mathbf{k}) + ik^0\mathcal{A}(\mathbf{k}); \\ \mathcal{B}(\mathbf{k}) &= i\mathbf{k} \times \mathcal{A}(\mathbf{k}) = \hat{\mathbf{k}} \times \mathcal{E}(\mathbf{k}). \end{aligned} \quad (6.9)$$

Using the explicit form (6.7) of $\mathcal{A}^\mu(\mathbf{k})$, you can easily check that the electric field is transverse: $\mathbf{k} \cdot \mathcal{E}(\mathbf{k}) = 0$.

Having expressed the fields in this way, we can compute the energy radiated:

$$\text{Energy} = \frac{1}{2} \int d^3x (|\mathbf{E}(x)|^2 + |\mathbf{B}(x)|^2). \quad (6.10)$$

The first term is

$$\begin{aligned} & \frac{1}{8} \int d^3x \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \left(\mathcal{E}(\mathbf{k}) e^{-ikx} + \mathcal{E}^*(\mathbf{k}) e^{ikx} \right) \cdot \left(\mathcal{E}(\mathbf{k}') e^{-ik'x} + \mathcal{E}^*(\mathbf{k}') e^{ik'x} \right) \\ &= \frac{1}{8} \int \frac{d^3k}{(2\pi)^3} \left(\mathcal{E}(\mathbf{k}) \cdot \mathcal{E}(-\mathbf{k}) e^{-2ik^0t} + 2\mathcal{E}(\mathbf{k}) \cdot \mathcal{E}^*(\mathbf{k}) + \mathcal{E}^*(\mathbf{k}) \cdot \mathcal{E}^*(-\mathbf{k}) e^{2ik^0t} \right). \end{aligned}$$

A similar expression involving $\mathcal{B}(\mathbf{k})$ holds for the second term. Using (6.9) and the fact that $\mathcal{E}(\mathbf{k})$ is transverse, you can show that the time-dependent terms cancel between \mathcal{E} and \mathcal{B} , while the remaining terms add to give

$$\text{Energy} = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} \mathcal{E}(\mathbf{k}) \cdot \mathcal{E}^*(\mathbf{k}). \quad (6.11)$$

Since $\mathcal{E}(\mathbf{k})$ is transverse, let us introduce two transverse unit polarization vectors $\epsilon_\lambda(\mathbf{k})$, $\lambda = 1, 2$. We can then write the integrand as

$$\mathcal{E}(\mathbf{k}) \cdot \mathcal{E}^*(\mathbf{k}) = \sum_{\lambda=1,2} |\epsilon_\lambda(\mathbf{k}) \cdot \mathcal{E}(\mathbf{k})|^2 = |\mathbf{k}|^2 \sum_{\lambda=1,2} |\epsilon_\lambda(\mathbf{k}) \cdot \mathcal{A}(\mathbf{k})|^2.$$

Using the explicit form of $\mathcal{A}(\mathbf{k})$ (6.7), we finally arrive at an expression for the energy radiated*:

$$\text{Energy} = \int \frac{d^3k}{(2\pi)^3} \sum_{\lambda=1,2} \frac{e^2}{2} \left| \epsilon_\lambda(\mathbf{k}) \cdot \left(\frac{\mathbf{p}'}{k \cdot p'} - \frac{\mathbf{p}}{k \cdot p} \right) \right|^2. \quad (6.12)$$

We can freely change ϵ , \mathbf{p}' , and \mathbf{p} into 4-vectors in this expression. Then, noting that substituting k^μ for ϵ^μ would give zero,

$$k_\mu \left(\frac{p'^\mu}{k \cdot p'} - \frac{p^\mu}{k \cdot p} \right) = 0,$$

we find that we can perform the sum over polarizations using the trick of Section 5.5, replacing $\sum \epsilon_\mu \epsilon_\nu^*$ by $-g_{\mu\nu}$. Our result then becomes

$$\begin{aligned} \text{Energy} &= \int \frac{d^3k}{(2\pi)^3} \frac{e^2}{2} (-g_{\mu\nu}) \left(\frac{p'^\mu}{k \cdot p'} - \frac{p^\mu}{k \cdot p} \right) \left(\frac{p'^\nu}{k \cdot p'} - \frac{p^\nu}{k \cdot p} \right) \\ &= \int \frac{d^3k}{(2\pi)^3} \frac{e^2}{2} \left(\frac{2p \cdot p'}{(k \cdot p')(k \cdot p)} - \frac{m^2}{(k \cdot p')^2} - \frac{m^2}{(k \cdot p)^2} \right). \end{aligned} \quad (6.13)$$

To make this formula more explicit, choose a frame in which $p^0 = p'^0 = E$. Then the momenta are

$$k^\mu = (k, \mathbf{k}), \quad p^\mu = E(1, \mathbf{v}), \quad p'^\mu = E(1, \mathbf{v}').$$

*This result is also derived in Jackson (1975), p. 703.

In such a frame our formula becomes

$$\text{Energy} = \frac{e^2}{(2\pi)^2} \int dk \mathcal{I}(\mathbf{v}, \mathbf{v}'), \quad (6.14)$$

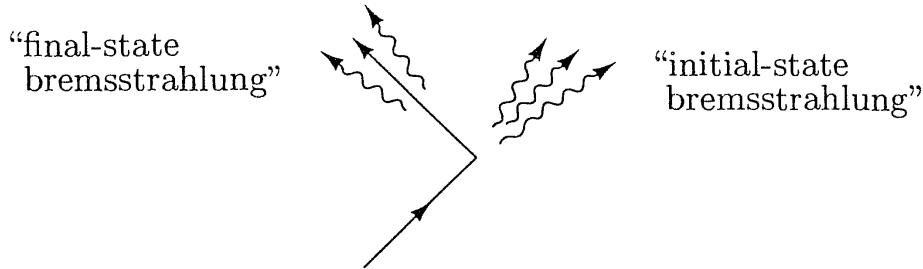
where $\mathcal{I}(\mathbf{v}, \mathbf{v}')$ (which is essentially the differential intensity $d(\text{Energy})/dk$) is given by

$$\mathcal{I}(\mathbf{v}, \mathbf{v}') = \int \frac{d\Omega_{\hat{k}}}{4\pi} \left(\frac{2(1-\mathbf{v} \cdot \mathbf{v}')}{(1-\hat{k} \cdot \mathbf{v})(1-\hat{k} \cdot \mathbf{v}')} - \frac{m^2/E^2}{(1-\hat{k} \cdot \mathbf{v}')^2} - \frac{m^2/E^2}{(1-\hat{k} \cdot \mathbf{v})^2} \right). \quad (6.15)$$

Since $\mathcal{I}(\mathbf{v}, \mathbf{v}')$ does not depend on k , we see that the integral over k in (6.14) is trivial but divergent. This divergence comes from our idealization of an infinitely sudden change in momentum. We expect our formula to be valid only for radiation whose frequency is less than the reciprocal of the scattering time. For a relativistic electron, another possible cutoff would take effect when individual photons carry away a sizable fraction of the electron's energy. In either case our formula is valid in the low-frequency limit, provided that we cut off the integral at some maximum frequency k_{\max} . We then have

$$\text{Energy} = \frac{\alpha}{\pi} \cdot k_{\max} \cdot \mathcal{I}(\mathbf{v}, \mathbf{v}'). \quad (6.16)$$

The integrand of $\mathcal{I}(\mathbf{v}, \mathbf{v}')$ peaks when \hat{k} is parallel to either \mathbf{v} or \mathbf{v}' :



In the extreme relativistic limit, most of the radiated energy comes from the two peaks in the first term of (6.15). Let us evaluate $\mathcal{I}(\mathbf{v}, \mathbf{v}')$ in this limit, by concentrating on the regions around these peaks. Break up the integral into a piece for each peak, and let $\theta = 0$ along the peak in each case. Integrate over a small region around $\theta = 0$, as follows:

$$\begin{aligned} \mathcal{I}(\mathbf{v}, \mathbf{v}') &\approx \int_{\hat{k} \cdot \mathbf{v} = \mathbf{v}' \cdot \mathbf{v}}^{\cos \theta = 1} d\cos \theta \frac{1 - \mathbf{v} \cdot \mathbf{v}'}{(1 - v \cos \theta)(1 - \mathbf{v} \cdot \mathbf{v}')} \\ &+ \int_{\hat{k} \cdot \mathbf{v} = \mathbf{v}' \cdot \mathbf{v}}^{\cos \theta = 1} d\cos \theta \frac{1 - \mathbf{v} \cdot \mathbf{v}'}{(1 - \mathbf{v} \cdot \mathbf{v}')(1 - v' \cos \theta)}. \end{aligned}$$

(The lower limits on the integrals are not critical; an equally good choice would be $\hat{k} \cdot \mathbf{v} = 1 - x(1 - \mathbf{v} \cdot \mathbf{v}')$, as long as x is neither too close to 0 nor too much bigger than 1. It is then easy to show that the leading term in the

relativistic limit does not depend on x .) The integrals are easy to perform, and we obtain

$$\begin{aligned}\mathcal{I}(\mathbf{v}, \mathbf{v}') &\approx \log\left(\frac{1 - \mathbf{v}' \cdot \mathbf{v}}{1 - |\mathbf{v}|}\right) + \log\left(\frac{1 - \mathbf{v}' \cdot \mathbf{v}}{1 - |\mathbf{v}'|}\right) = \log\left(\frac{(E^2 - \mathbf{p} \cdot \mathbf{p}')^2}{E^2(E - p)^2}\right) \\ &\approx 2 \log\left(\frac{p \cdot p'}{(E^2 - p^2)/2}\right) = 2 \log\left(\frac{-q^2}{m^2}\right),\end{aligned}\quad (6.17)$$

where $q^2 = (p' - p)^2$.

In conclusion, we have found that the radiated energy at low frequencies is given by

$$\text{Energy} = \frac{\alpha}{\pi} \int_0^{k_{\max}} dk \mathcal{I}(\mathbf{v}, \mathbf{v}') \xrightarrow{E \gg m} \frac{2\alpha}{\pi} \int_0^{k_{\max}} dk \log\left(\frac{-q^2}{m^2}\right). \quad (6.18)$$

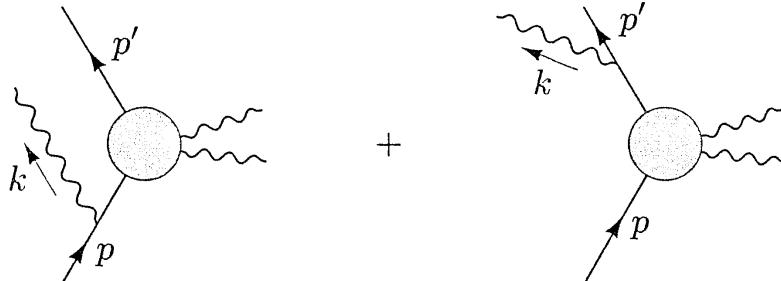
If this energy is made up of photons, each photon contributes energy k . We would then expect

$$\text{Number of photons} = \frac{\alpha}{\pi} \int_0^{k_{\max}} dk \frac{1}{k} \mathcal{I}(\mathbf{v}, \mathbf{v}'). \quad (6.19)$$

We hope that a quantum-mechanical calculation will confirm this result.

Quantum Computation

Consider now the quantum-mechanical process in which one photon is radiated during the scattering of an electron:



Let \mathcal{M}_0 denote the part of the amplitude that comes from the electron's interaction with the external field. Then the amplitude for the whole process is

$$\begin{aligned}i\mathcal{M} &= \bar{u}(p') \left(\mathcal{M}_0(p', p - k) \frac{i(p' - k + m)}{(p - k)^2 - m^2} \gamma^\mu \epsilon_\mu^*(k) \right. \\ &\quad \left. + \gamma^\mu \epsilon_\mu^*(k) \frac{i(p' + k + m)}{(p' + k)^2 - m^2} \mathcal{M}_0(p' + k, p) \right) u(p).\end{aligned}\quad (6.20)$$

Since we are interested in connecting with the classical limit, assume that the photon radiated is soft: $|k| \ll |p' - p|$. Then we can approximate

$$\mathcal{M}_0(p', p - k) \approx \mathcal{M}_0(p' + k, p) \approx \mathcal{M}_0(p', p), \quad (6.21)$$

and we can ignore \not{p} in the numerators of the propagators. The numerators can be further simplified with some Dirac algebra. In the first term we have

$$\begin{aligned} (\not{p} + m)\gamma^\mu \epsilon_\mu^* u(p) &= [2p^\mu \epsilon_\mu^* + \gamma^\mu \epsilon_\mu^* (-\not{p} + m)] u(p) \\ &= 2p^\mu \epsilon_\mu^* u(p). \end{aligned}$$

Similarly, in the second term,

$$\bar{u}(p') \gamma^\mu \epsilon_\mu^* (\not{p}' + m) = \bar{u}(p') 2p'^\mu \epsilon_\mu^*.$$

The denominators of the propagators also simplify:

$$(p - k)^2 - m^2 = -2p \cdot k; \quad (p' - k)^2 - m^2 = 2p' \cdot k.$$

So in the soft-photon approximation, the amplitude becomes

$$i\mathcal{M} = \bar{u}(p') [\mathcal{M}_0(p', p)] u(p) \cdot \left[e \left(\frac{p' \cdot \epsilon^*}{p' \cdot k} - \frac{p \cdot \epsilon^*}{p \cdot k} \right) \right]. \quad (6.22)$$

This is just the amplitude for elastic scattering (without bremsstrahlung), times a factor (in brackets) for the emission of the photon.

The cross section for our process is also easy to express in terms of the elastic cross section; just insert an additional phase-space integration for the photon variable k . Summing over the two photon polarization states, we have

$$d\sigma(p \rightarrow p' + \gamma) = d\sigma(p \rightarrow p') \cdot \int \frac{d^3 k}{(2\pi)^3} \frac{1}{2k} \sum_{\lambda=1,2} e^2 \left| \frac{p' \cdot \epsilon^{(\lambda)}}{p' \cdot k} - \frac{p \cdot \epsilon^{(\lambda)}}{p \cdot k} \right|^2. \quad (6.23)$$

Thus the differential probability of radiating a photon with momentum k , given that the electron scatters from p to p' , is

$$d(\text{prob}) = \frac{d^3 k}{(2\pi)^3} \sum_{\lambda} \frac{e^2}{2k} \left| \epsilon_{\lambda} \cdot \left(\frac{\mathbf{p}'}{p' \cdot k} - \frac{\mathbf{p}}{p \cdot k} \right) \right|^2. \quad (6.24)$$

This looks very familiar; if we multiply by the photon energy k to compute the expected energy radiated, we recover the classical expression (6.12).

But there is a problem. Equation (6.24) is an expression not for the expected number of photons radiated, but for the probability of radiating a single photon. The problem becomes worse if we integrate over the photon momentum. As in (6.16), we can integrate only up to the energy at which our soft-photon approximations break down; a reasonable estimate for this energy is $|\mathbf{q}| = |\mathbf{p} - \mathbf{p}'|$. The integral is therefore

$$\text{Total probability} \approx \frac{\alpha}{\pi} \int_0^{|\mathbf{q}|} dk \frac{1}{k} \mathcal{I}(\mathbf{v}, \mathbf{v}'). \quad (6.25)$$

Since $\mathcal{I}(\mathbf{v}, \mathbf{v}')$ is independent of k , the integral diverges at its lower limit (where all our approximations are well justified). In other words, the total

probability of radiating a very soft photon is infinite. This is the famous problem of infrared divergences in QED perturbation theory.

We can artificially make the integral in (6.25) well-defined by pretending that the photon has a very small mass μ . This mass would then provide a lower cutoff for the integral, allowing us to write the result of this section as

$$d\sigma(p \rightarrow p' + \gamma(k)) = d\sigma(p \rightarrow p') \cdot \frac{\alpha}{2\pi} \log\left(\frac{-q^2}{\mu^2}\right) \mathcal{I}(\mathbf{v}, \mathbf{v}') \quad (6.26)$$

$$\underset{-q^2 \rightarrow \infty}{\approx} d\sigma(p \rightarrow p') \cdot \frac{\alpha}{\pi} \log\left(\frac{-q^2}{\mu^2}\right) \log\left(\frac{-q^2}{m^2}\right).$$

The q^2 dependence of this result, known as the *Sudakov double logarithm*, is physical and will appear again in Section 6.4. The dependence on μ , however, presents a problem that we must solve. It is not hard to guess that the resolution of this problem will involve reinterpreting (6.24) as the expected number of radiated photons, rather than the probability of radiating a single photon. We will see in Sections 6.4 and 6.5 how this reinterpretation follows from the Feynman diagrams. To prepare for that discussion, however, we need to improve our understanding of the amplitude for scattering without radiation.

6.2 The Electron Vertex Function: Formal Structure

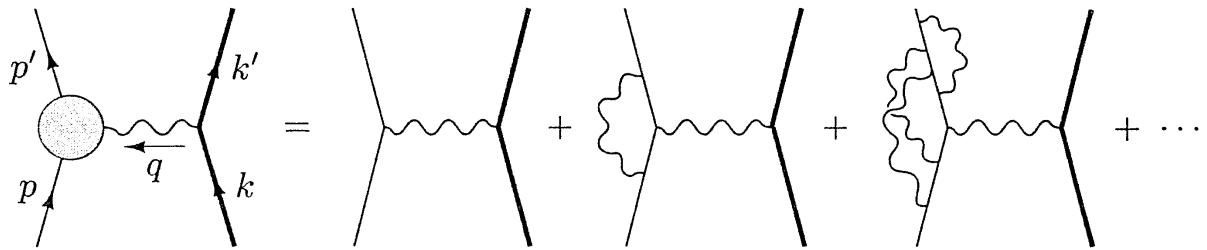
Having briefly discussed QED radiative corrections due to emission of photons (bremsstrahlung), let us now study the correction to electron scattering that comes from the presence of an additional *virtual* photon:



This will be our first experience with a Feynman diagram containing a loop. Such diagrams give rise to significant and profound complications in quantum field theory.

The result of computing this diagram will be rather complicated, so it will be useful to think ahead about what form we expect this correction to take and how to interpret its various possible terms. In this section, we will consider the general properties of vertex correction diagrams. We will see that the basic requirements of Lorentz invariance, the discrete symmetries of QED, and the Ward identity strongly constrain the form of the vertex.

Consider, then, the class of diagrams



where the gray circle indicates the sum of the lowest-order electron-photon vertex and all amputated loop corrections. We will call this sum of vertex diagrams $-ie\Gamma^\mu(p', p)$. Then, according to our master formula (4.103) for S -matrix elements, the amplitude for electron scattering from a heavy target is

$$i\mathcal{M} = ie^2 \left(\bar{u}(p') \Gamma^\mu(p', p) u(p) \right) \frac{1}{q^2} \left(\bar{u}(k') \gamma_\mu u(k) \right). \quad (6.28)$$

More generally, the function $\Gamma^\mu(p', p)$ appears in the S -matrix element for the scattering of an electron from an external electromagnetic field. As in Problem 4.4, add to the Hamiltonian of QED the interaction

$$\Delta H_{\text{int}} = \int d^3x e A_\mu^{\text{cl}} j^\mu, \quad (6.29)$$

where $j_\mu(x) = \bar{\psi}(x) \gamma^\mu \psi(x)$ is the electromagnetic current and A_μ^{cl} is a fixed classical potential. In the leading order of perturbation theory, the S -matrix element for scattering from this field is

$$i\mathcal{M} (2\pi) \delta(p^{0'} - p^0) = -ie\bar{u}(p') \gamma^\mu u(p) \cdot \tilde{A}_\mu^{\text{cl}}(p' - p),$$

where $\tilde{A}_\mu^{\text{cl}}(q)$ is the Fourier transform of $A_\mu^{\text{cl}}(x)$. The vertex corrections modify this expression to

$$i\mathcal{M} (2\pi) \delta(p^{0'} - p^0) = -ie\bar{u}(p') \Gamma^\mu(p', p) u(p) \cdot \tilde{A}_\mu^{\text{cl}}(p' - p). \quad (6.30)$$

In writing (6.28) and (6.30), we have deliberately omitted the contribution of vacuum polarization diagrams, such as the fourth diagram of (6.1). The reason for this omission is that these diagrams should be considered corrections to the electromagnetic field itself, while the diagrams included in Γ^μ represent corrections to the electron's response to a given applied field.[†]

We can use general arguments to restrict the form of $\Gamma^\mu(p', p)$. To lowest order, $\Gamma^\mu = \gamma^\mu$. In general, Γ^μ is some expression that involves p , p' , γ^μ , and constants such as m , e , and pure numbers. This list is exhaustive, since no other objects appear in the Feynman rules for evaluating the diagrams that contribute to Γ^μ . The only other object that could appear in any theory is $\epsilon^{\mu\nu\rho\sigma}$ (or equivalently, γ^5); but this is forbidden in any parity-conserving theory.

[†]To justify this statement, we must give a careful definition of an applied external field in a quantum field theory. We will do this in Chapter 11.

We can narrow down the form of Γ^μ considerably by appealing to Lorentz invariance. Since Γ^μ transforms as a vector (in the same sense that γ^μ does), it must be a linear combination of the vectors from the list above: γ^μ , p^μ , and p'^μ . Using the combinations $p' + p$ and $p' - p$ for convenience, we have

$$\Gamma^\mu = \gamma^\mu \cdot A + (p'^\mu + p^\mu) \cdot B + (p'^\mu - p^\mu) \cdot C. \quad (6.31)$$

The coefficients A , B , and C could involve Dirac matrices dotted into vectors, that is, \not{p} or \not{p}' . But since $\not{p}u(p) = m \cdot u(p)$ and $\bar{u}(p')\not{p}' = \bar{u}(p') \cdot m$, we can write the coefficients in terms of ordinary numbers without loss of generality. The only nontrivial scalar available is $q^2 = -2p' \cdot p + 2m^2$, so A , B , and C must be functions only of q^2 (and of constants such as m).

The list of allowed vectors can be further shortened by applying the Ward identity (5.79): $q_\mu \Gamma^\mu = 0$. (Note that our arguments for this identity in Section 5.4—and the proof in Section 7.4—do not require $q^2 = 0$.) Dotting q_μ into (6.31), we find that the second term vanishes, as does the first when sandwiched between $\bar{u}(p')$ and $u(p)$. The third term does not automatically vanish, so C must be zero.

We can make no further simplifications of (6.31) on general principles. It is conventional, however, to rewrite (6.31) by means of the Gordon identity (see Problem 3.2):

$$\bar{u}(p')\gamma^\mu u(p) = \bar{u}(p') \left[\frac{p'^\mu + p^\mu}{2m} + \frac{i\sigma^{\mu\nu}q_\nu}{2m} \right] u(p). \quad (6.32)$$

This identity allows us to swap the $(p' + p)$ term for one involving $\sigma^{\mu\nu}q_\nu$. We write our final result as

$$\Gamma^\mu(p', p) = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m} F_2(q^2), \quad (6.33)$$

where F_1 and F_2 are unknown functions of q^2 called *form factors*.

To lowest order, $F_1 = 1$ and $F_2 = 0$. In the next section we will compute the one-loop (order- α) corrections to the form factors, due to the vertex correction diagram (6.27). In principle, the form factors can be computed to any order in perturbation theory.

Since F_1 and F_2 contain complete information about the influence of an electromagnetic field on the electron, they should, in particular, contain the electron's gross electric and magnetic couplings. To identify the electric charge of the electron, we can use (6.30) to compute the amplitude for elastic Coulomb scattering of a nonrelativistic electron from a region of nonzero electrostatic potential. Set $A_\mu^{\text{cl}}(x) = (\phi(x), \mathbf{0})$. Then $\tilde{A}_\mu^{\text{cl}}(q) = ((2\pi)\delta(q^0)\tilde{\phi}(\mathbf{q}), \mathbf{0})$. Inserting this into (6.30), we find

$$i\mathcal{M} = -ie\bar{u}(p') \Gamma^0(p', p) u(p) \cdot \tilde{\phi}(\mathbf{q}).$$

If the electrostatic field is very slowly varying over a large (perhaps macroscopic) region, $\phi(\mathbf{q})$ will be concentrated about $\mathbf{q} = 0$; then we can take the

limit $\mathbf{q} \rightarrow 0$ in the spinor matrix element. Only the form factor F_1 contributes. Using the nonrelativistic limit of the spinors,

$$\bar{u}(p')\gamma^0 u(p) = u^\dagger(p')u(p) \approx 2m\xi'^\dagger\xi,$$

the amplitude for electron scattering from an electric field takes the form

$$i\mathcal{M} = -ieF_1(0)\tilde{\phi}(\mathbf{q}) \cdot 2m\xi'^\dagger\xi. \quad (6.34)$$

This is the Born approximation for scattering from a potential

$$V(\mathbf{x}) = eF_1(0)\phi(\mathbf{x}).$$

Thus $F_1(0)$ is the electric charge of the electron, in units of e . Since $F_1(0) = 1$ already in the leading order of perturbation theory, radiative corrections to $F_1(q^2)$ should vanish at $q^2 = 0$.

By repeating this analysis for an electron scattering from a static vector potential, we can derive a similar connection between the form factors and the electron's magnetic moment.[‡] Set $A_\mu^{\text{cl}}(x) = (0, \mathbf{A}^{\text{cl}}(\mathbf{x}))$. Then the amplitude for scattering from this field is

$$i\mathcal{M} = +ie\left[\bar{u}(p')\left(\gamma^i F_1 + \frac{i\sigma^{i\nu}q_\nu}{2m}F_2\right)u(p)\right]\tilde{A}_{\text{cl}}^i(\mathbf{q}). \quad (6.35)$$

The expression in brackets vanishes at $\mathbf{q} = 0$, so we must carefully extract from it a contribution linear in q^i . To do this, insert the nonrelativistic expansion of the spinors $u(p)$, keeping terms through first order in momenta:

$$u(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi \\ \sqrt{p \cdot \bar{\sigma}} \xi \end{pmatrix} \approx \sqrt{m} \begin{pmatrix} (1 - \mathbf{p} \cdot \sigma/2m)\xi \\ (1 + \mathbf{p} \cdot \sigma/2m)\xi \end{pmatrix}. \quad (6.36)$$

Then the F_1 term can be simplified as follows:

$$\bar{u}(p')\gamma^i u(p) = 2m\xi'^\dagger \left(\frac{\mathbf{p}' \cdot \sigma}{2m} \sigma^i + \sigma^i \frac{\mathbf{p} \cdot \sigma}{2m} \right) \xi.$$

Applying the identity $\sigma^i \sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k$, we find a spin-independent term, proportional to $(\mathbf{p}' + \mathbf{p})$, and a spin-dependent term, proportional to $(\mathbf{p}' - \mathbf{p})$. The first of these terms is the contribution of the operator $[\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}]$ in the standard kinetic energy term of nonrelativistic quantum mechanics. The second is the magnetic moment interaction we are seeking. Retaining only the latter term, we have

$$\bar{u}(p')\gamma^i u(p) = 2m\xi'^\dagger \left(\frac{-i}{2m} \epsilon^{ijk} q^j \sigma^k \right) \xi.$$

The F_2 term already contains an explicit factor of q , so we can evaluate it using the leading-order term of the expansion of the spinors. This gives

$$\bar{u}(p') \left(\frac{i}{2m} \sigma^{i\nu} q_\nu \right) u(p) = 2m\xi'^\dagger \left(\frac{-i}{2m} \epsilon^{ijk} q^j \sigma^k \right) \xi.$$

[‡]The following argument contains numerous factors of (-1) from raising and lowering spacelike indices. Be careful in verifying the algebra.

Thus, the complete term linear in q^j in the electron-photon vertex function is

$$\bar{u}(p') \left(\gamma^i F_1 + \frac{i\sigma^{i\nu} q_\nu}{2m} F_2 \right) u(p) \underset{q \rightarrow 0}{\approx} 2m \xi'^\dagger \left(\frac{-i}{2m} \epsilon^{ijk} q^j \sigma^k [F_1(0) + F_2(0)] \right) \xi.$$

Inserting this expression into (6.35), we find

$$i\mathcal{M} = -i(2m) \cdot e \xi'^\dagger \left(\frac{-1}{2m} \sigma^k [F_1(0) + F_2(0)] \right) \xi \tilde{B}^k(\mathbf{q}),$$

where

$$\tilde{B}^k(\mathbf{q}) = -i\epsilon^{ijk} q^i \tilde{A}_{\text{cl}}^j(\mathbf{q})$$

is the Fourier transform of the magnetic field produced by $\mathbf{A}^{\text{cl}}(\mathbf{x})$.

Again we can interpret \mathcal{M} as the Born approximation to the scattering of the electron from a potential well. The potential is just that of a magnetic moment interaction,

$$V(\mathbf{x}) = -\langle \boldsymbol{\mu} \rangle \cdot \mathbf{B}(\mathbf{x}),$$

where

$$\langle \boldsymbol{\mu} \rangle = \frac{e}{m} [F_1(0) + F_2(0)] \xi'^\dagger \frac{\sigma}{2} \xi.$$

This expression for the magnetic moment of the electron can be rewritten in the standard form

$$\boldsymbol{\mu} = g \left(\frac{e}{2m} \right) \mathbf{S},$$

where \mathbf{S} is the electron spin. The coefficient g , called the *Landé g-factor*, is

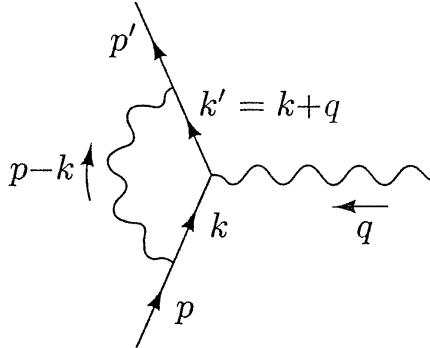
$$g = 2 [F_1(0) + F_2(0)] = 2 + 2F_2(0). \quad (6.37)$$

Since the leading order of perturbation theory gives no F_2 term, QED predicts $g = 2 + \mathcal{O}(\alpha)$. The leading term is the standard prediction of the Dirac equation. In higher orders, however, we will find a nonzero F_2 and thus a small difference between the electron's magnetic moment and the Dirac value. We will compute the order- α contribution to this *anomalous magnetic moment* in the next section.

Since our derivation of the structure (6.33) for the vertex function used only general symmetry principles, we expect this formula to apply not only to the electron but to any fermion with electromagnetic interactions. For example, the electromagnetic scattering amplitude of the proton should also be described by two invariant functions of q^2 . Since the proton is not an elementary particle, we should not expect the Dirac equation values $F_1 = 1$ and $F_2 = 0$ to be good approximations to the form factors of the proton. In fact, both proton form factors depend strongly on q^2 . However, the description of the vertex function in term of form factors provides a useful summary of data on scattering at many energies and angles. The precise transcription between form factors and cross sections is worked out in Problem 6.1. In addition, the general constraints at $q^2 = 0$ that we have just derived apply to the proton: $F_1(0) = 1$, and $2F_2(0) = (g_p - 2)$, though the g -factor of the proton differs by 40% from the Dirac value.

6.3 The Electron Vertex Function: Evaluation

Now that we know what form the answer is to take (Eq. (6.33)), we are ready to evaluate the one-loop contribution to the electron vertex function. Assign momenta on the diagram as follows:



Applying the Feynman rules, we find, to order α , that $\Gamma^\mu = \gamma^\mu + \delta\Gamma^\mu$, where

$$\begin{aligned} \delta\Gamma^\mu(p', p) &= \int \frac{d^4 k}{(2\pi)^4} \frac{-ig_{\nu\rho}}{(k-p)^2 + i\epsilon} \bar{u}(p') (-ie\gamma^\nu) \frac{i(k' + m)}{k'^2 - m^2 + i\epsilon} \gamma^\mu \frac{i(k + m)}{k^2 - m^2 + i\epsilon} (-ie\gamma^\rho) u(p) \\ &= 2ie^2 \int \frac{d^4 k}{(2\pi)^4} \frac{\bar{u}(p') [k'\gamma^\mu k' + m^2\gamma^\mu - 2m(k + k')^\mu] u(p)}{((k - p)^2 + i\epsilon)(k'^2 - m^2 + i\epsilon)(k^2 - m^2 + i\epsilon)}. \end{aligned} \quad (6.38)$$

In the second line we have used the contraction identity $\gamma^\nu\gamma^\mu\gamma_\nu = -2\gamma^\mu$. Note that the $+i\epsilon$ terms in the denominators cannot be dropped; they are necessary for proper evaluation of the loop-momentum integral.

The integral looks impossible, and in fact it will not be easy. The evaluation of such integrals requires another piece of computational technology, known as the method of *Feynman parameters* (although a very similar method was introduced earlier by Schwinger).

Feynman Parameters

The goal of this method is to squeeze the three denominator factors of (6.38) into a single quadratic polynomial in k , raised to the third power. We can then shift k by a constant to complete the square in this polynomial and evaluate the remaining spherically symmetric integral without difficulty. The price will be the introduction of auxiliary parameters to be integrated over.

It is easiest to begin with the simpler case of two factors in the denominator. We would then use the identity

$$\frac{1}{AB} = \int_0^1 dx \frac{1}{[xA + (1-x)B]^2} = \int_0^1 dx dy \delta(x+y-1) \frac{1}{[xA + yB]^2}. \quad (6.39)$$

An example of its use might look like this:

$$\begin{aligned} \frac{1}{(k-p)^2(k^2-m^2)} &= \int_0^1 dx dy \delta(x+y-1) \frac{1}{[x(k-p)^2 + y(k^2-m^2)]^2} \\ &= \int_0^1 dx dy \delta(x+y-1) \frac{1}{[k^2 - 2xk \cdot p + xp^2 - ym^2]^2}. \end{aligned}$$

If we now let $\ell \equiv k - xp$, we see that the denominator depends only on ℓ^2 . Integrating over d^4k would now be much easier, since $d^4k = d^4\ell$ and the integrand is spherically symmetric with respect to ℓ . The variables x and y that make this transformation possible are called *Feynman parameters*.

Our integral (6.38) involves a denominator with three factors, so we need a slightly better identity. By differentiating (6.39) with respect to B , it is easy to prove

$$\frac{1}{AB^n} = \int_0^1 dx dy \delta(x+y-1) \frac{ny^{n-1}}{[xA + yB]^{n+1}}. \quad (6.40)$$

But this still isn't quite good enough. The formula we need is

$$\frac{1}{A_1 A_2 \cdots A_n} = \int_0^1 dx_1 \cdots dx_n \delta(\sum x_i - 1) \frac{(n-1)!}{[x_1 A_1 + x_2 A_2 + \cdots + x_n A_n]^n}. \quad (6.41)$$

The proof of this identity is by induction. The case $n = 2$ is just Eq. (6.39); the induction step is not difficult and involves the use of (6.40).

By repeated differentiation of (6.41), you can derive the even more general identity

$$\frac{1}{A_1^{m_1} A_2^{m_2} \cdots A_n^{m_n}} = \int_0^1 dx_1 \cdots dx_n \delta(\sum x_i - 1) \frac{\prod x_i^{m_i-1}}{[\sum x_i A_i]^{\sum m_i}} \frac{\Gamma(m_1 + \cdots + m_n)}{\Gamma(m_1) \cdots \Gamma(m_n)}. \quad (6.42)$$

This formula is true even when the m_i are not integers; in Section 10.5 we will apply it in such a case.

Evaluation of the Form Factors

Now let us apply formula (6.41) to the denominator of (6.38):

$$\frac{1}{((k-p)^2+i\epsilon)(k'^2-m^2+i\epsilon)(k^2-m^2+i\epsilon)} = \int_0^1 dx dy dz \delta(x+y+z-1) \frac{2}{D^3},$$

where the new denominator D is

$$\begin{aligned} D &= x(k^2 - m^2) + y(k'^2 - m^2) + z(k - p)^2 + (x + y + z)i\epsilon \\ &= k^2 + 2k \cdot (yq - zp) + yq^2 + zp^2 - (x + y)m^2 + i\epsilon. \end{aligned} \quad (6.43)$$

In the second line we have used $x + y + z = 1$ and $k' = k + q$. Now shift k to complete the square:

$$\ell \equiv k + yq - zp.$$

After a bit of algebra we find that D simplifies to

$$D = \ell^2 - \Delta + i\epsilon,$$

where

$$\Delta \equiv -xyq^2 + (1 - z)^2m^2. \quad (6.44)$$

Since $q^2 < 0$ for a scattering process, Δ is positive; we can think of it as an effective mass term.

Next we must express the numerator of (6.38) in terms of ℓ . This task is simplified by noting that since D depends only on the magnitude of ℓ ,

$$\int \frac{d^4\ell}{(2\pi)^4} \frac{\ell^\mu}{D^3} = 0; \quad (6.45)$$

$$\int \frac{d^4\ell}{(2\pi)^4} \frac{\ell^\mu \ell^\nu}{D^3} = \int \frac{d^4\ell}{(2\pi)^4} \frac{\frac{1}{4}g^{\mu\nu}\ell^2}{D^3}. \quad (6.46)$$

The first identity follows from symmetry. To prove the second, note that the integral vanishes by symmetry unless $\mu = \nu$. Lorentz invariance therefore requires that we get something proportional to $g^{\mu\nu}$. To check the coefficient, contract each side with $g_{\mu\nu}$. Using these identities, we have

$$\begin{aligned} \text{Numerator} &= \bar{u}(p') \left[k' \gamma^\mu k'' + m^2 \gamma^\mu - 2m(k + k')^\mu \right] u(p) \\ &\rightarrow \bar{u}(p') \left[-\frac{1}{2} \gamma^\mu \ell^2 + (-y\cancel{q} + z\cancel{p}) \gamma^\mu ((1 - y)\cancel{q} + z\cancel{p}) \right. \\ &\quad \left. + m^2 \gamma^\mu - 2m((1 - 2y)q^\mu + 2zp^\mu) \right] u(p). \end{aligned}$$

(Remember that $k' = k + q$.)

Putting the numerator into a useful form is now just a matter of some tedious Dirac algebra (about a page or two). This is where our work in the last section pays off, since it tells us what kind of an answer to expect. We eventually want to group everything into two terms, proportional to γ^μ and $i\sigma^{\mu\nu}q_\nu$. The most straightforward way to accomplish this is to aim instead for an expression of the form

$$\gamma^\mu \cdot A + (p'^\mu + p^\mu) \cdot B + q^\mu \cdot C,$$

just as in (6.31). Attaining this form requires only the anticommutation relations (for example, $\cancel{p}\gamma^\mu = 2p^\mu - \gamma^\mu\cancel{p}$) and the Dirac equation ($\cancel{p}u(p) = m u(p)$)

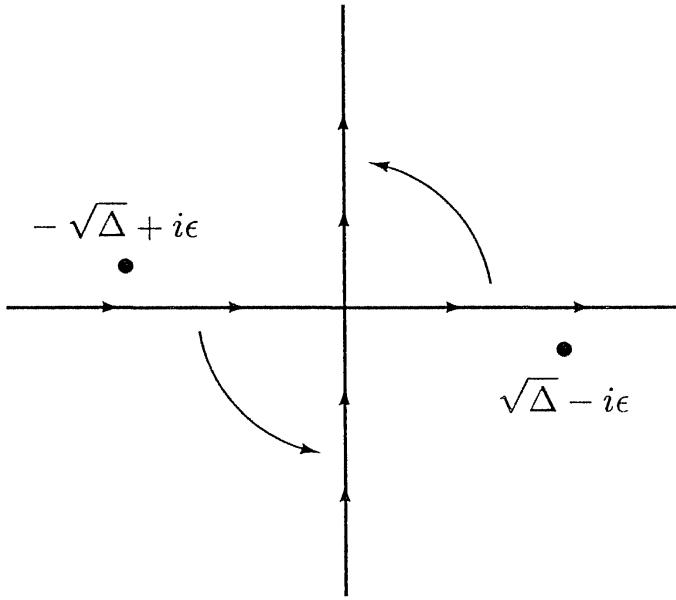


Figure 6.1. The contour of the ℓ^0 integration may be rotated as shown.

and $\bar{u}(p')\not{p}'' = \bar{u}(p')m$; note that this implies $\bar{u}(p')\not{u}(p) = 0$). It is also useful to remember that $x + y + z = 1$. When the smoke clears, we have

$$\begin{aligned} \text{Numerator} = \bar{u}(p') & \left[\gamma^\mu \cdot \left(-\frac{1}{2}\ell^2 + (1-x)(1-y)q^2 + (1-2z-z^2) \right) \right. \\ & \left. + (p'^\mu + p^\mu) \cdot mz(z-1) + q^\mu \cdot m(z-2)(x-y) \right] u(p). \end{aligned}$$

The coefficient of q^μ must vanish according to the Ward identity, as discussed after Eq. (6.31). To see that it does, note from (6.44) that the denominator is symmetric under $x \leftrightarrow y$. The coefficient of q^μ is odd under $x \leftrightarrow y$ and therefore vanishes when integrated over x and y .

Still following our work in the previous section, we now use the Gordon identity (6.32) to eliminate $(p' + p)$ in favor of $i\sigma^{\mu\nu}q_\nu$. Our entire expression for the $\mathcal{O}(\alpha)$ contribution to the electron vertex then becomes

$$\begin{aligned} \delta\Gamma^\mu(p', p) = 2ie^2 \int \frac{d^4\ell}{(2\pi)^4} \int_0^1 & dx dy dz \delta(x+y+z-1) \frac{2}{D^3} \\ & \times \bar{u}(p') \left[\gamma^\mu \cdot \left(-\frac{1}{2}\ell^2 + (1-x)(1-y)q^2 + (1-4z+z^2)m^2 \right) \right. \\ & \left. + \frac{i\sigma^{\mu\nu}q_\nu}{2m} (2m^2 z(1-z)) \right] u(p), \end{aligned} \quad (6.47)$$

where as before,

$$D = \ell^2 - \Delta + i\epsilon, \quad \Delta = -xyq^2 + (1-z)^2m^2 > 0.$$

The decomposition into form factors is now manifest.

With most of the work behind us, our main remaining task is to perform the momentum integral. It is not difficult to evaluate the ℓ^0 integral as a

contour integral, then do the spatial integrals in spherical coordinates. We will use an even easier method, making use of a trick called *Wick rotation*. Note that if it were not for the minus signs in the Minkowski metric, we could perform the entire four-dimensional integral in four-dimensional “spherical” coordinates. To remove the minus signs, consider the contour of integration in the ℓ^0 -plane (see Fig. 6.1). The locations of the poles, and the fact that the integrand falls off sufficiently rapidly at large $|\ell^0|$, allow us to rotate the contour counterclockwise by 90° . We then define a *Euclidean* 4-momentum variable ℓ_E :

$$\ell^0 \equiv i\ell_E^0; \quad \ell = \ell_E. \quad (6.48)$$

Our rotated contour goes from $\ell_E^0 = -\infty$ to ∞ . By simply changing variables to ℓ_E , we can now evaluate the integral in four-dimensional spherical coordinates.

Let us first evaluate

$$\begin{aligned} \int \frac{d^4\ell}{(2\pi)^4} \frac{1}{[\ell^2 - \Delta]^m} &= \frac{i}{(-1)^m} \frac{1}{(2\pi)^4} \int d^4\ell_E \frac{1}{[\ell_E^2 + \Delta]^m} \\ &= \frac{i(-1)^m}{(2\pi)^4} \int d\Omega_4 \int_0^\infty d\ell_E \frac{\ell_E^3}{[\ell_E^2 + \Delta]^m}. \end{aligned}$$

(Here we need only the case $m = 3$, but the more general result will be useful for other loop calculations.) The factor $\int d\Omega_4$ is the surface “area” of a four-dimensional unit sphere, which happens to equal $2\pi^2$. (One way to compute this area is to use four-dimensional spherical coordinates,

$$x = (r \sin \omega \sin \theta \cos \phi, r \sin \omega \sin \theta \sin \phi, r \sin \omega \cos \theta, r \cos \omega).$$

The integration measure is then $d^4x = r^3 \sin^2 \omega \sin \theta d\phi d\theta d\omega dr$.) The rest of the integral is straightforward, and we have

$$\int \frac{d^4\ell}{(2\pi)^4} \frac{1}{[\ell^2 - \Delta]^m} = \frac{i(-1)^m}{(4\pi)^2} \frac{1}{(m-1)(m-2)} \frac{1}{\Delta^{m-2}}. \quad (6.49)$$

Similarly,

$$\int \frac{d^4\ell}{(2\pi)^4} \frac{\ell^2}{[\ell^2 - \Delta]^m} = \frac{i(-1)^{m-1}}{(4\pi)^2} \frac{2}{(m-1)(m-2)(m-3)} \frac{1}{\Delta^{m-3}}. \quad (6.50)$$

Note that this second result is valid only when $m > 3$. When $m = 3$, the Wick rotation cannot be justified, and the integral is in any event divergent. But it is just this case that we need for (6.47).

We will eventually explore the physical meaning of this divergence, but for the moment we simply introduce an artificial prescription to make our integral finite. Go back to the original expression for the Feynman integral in

(6.38), and replace in the photon propagator

$$\frac{1}{(k-p)^2 + i\epsilon} \longrightarrow \frac{1}{(k-p)^2 + i\epsilon} - \frac{1}{(k-p)^2 - \Lambda^2 + i\epsilon}, \quad (6.51)$$

where Λ is a very large mass. The integrand is unaffected for small k (since Λ is large), but cuts off smoothly when $k \gtrsim \Lambda$. We can think of the second term as the propagator of a fictitious heavy photon, whose contribution is subtracted from that of the ordinary photon. In terms involving the heavy photon, the numerator algebra is unchanged and the denominator is altered by

$$\Delta \longrightarrow \Delta_\Lambda = -xyq^2 + (1-z)^2m^2 + z\Lambda^2. \quad (6.52)$$

The integral (6.50) is then replaced with a convergent integral, which can be Wick-rotated and evaluated:

$$\begin{aligned} \int \frac{d^4\ell}{(2\pi)^4} \left(\frac{\ell^2}{[\ell^2 - \Delta]^3} - \frac{\ell^2}{[\ell^2 - \Delta_\Lambda]^3} \right) &= \frac{i}{(4\pi)^2} \int_0^\infty d\ell_E^2 \left(\frac{\ell_E^4}{[\ell_E^2 + \Delta]^3} - \frac{\ell_E^4}{[\ell_E^2 + \Delta_\Lambda]^3} \right) \\ &= \frac{i}{(4\pi)^2} \log\left(\frac{\Delta_\Lambda}{\Delta}\right). \end{aligned} \quad (6.53)$$

The convergent terms in (6.47) are modified by terms of order Λ^{-2} , which we ignore.

This prescription for rendering Feynman integrals finite by introducing fictitious heavy particles is known as *Pauli-Villars regularization*. Please note that the fictitious photon has no physical significance, and that this method is only one of many for defining the divergent integrals. (We will discuss other methods in the next chapter; see especially Problem 7.2.) We must hope that the new parameter Λ will not appear in our final results for observable cross sections.

Using formulae (6.49) and (6.53) to evaluate the integrals in (6.47), we obtain an explicit, though complicated, expression for the one-loop vertex correction:

$$\begin{aligned} \text{Diagram} &= \frac{\alpha}{2\pi} \int_0^1 dx dy dz \delta(x+y+z-1) \\ &\quad \times \bar{u}(p') \left(\gamma^\mu \left[\log \frac{z\Lambda^2}{\Delta} + \frac{1}{\Delta} \left((1-x)(1-y)q^2 + (1-4z+z^2)m^2 \right) \right] \right. \\ &\quad \left. + \frac{i\sigma^{\mu\nu}q_\nu}{2m} \left[\frac{1}{\Delta} 2m^2 z(1-z) \right] \right) u(p). \end{aligned} \quad (6.54)$$

The bracketed expressions are our desired corrections to the form factors.

Before we try to interpret this result, let us summarize the calculational methods we used. The techniques are common to all loop calculations:

1. Draw the diagram(s) and write down the amplitude.
2. Introduce Feynman parameters to combine the denominators of the propagators.
3. Complete the square in the new denominator by shifting to a new loop momentum variable, ℓ .
4. Write the numerator in terms of ℓ . Drop odd powers of ℓ , and rewrite even powers using identities like (6.46).
5. Perform the momentum integral by means of a Wick rotation and four-dimensional spherical coordinates.

The momentum integral in the last step will often be divergent. In that case we must define (or *regularize*) the integral using the Pauli-Villars prescription or some other device.

Now that we have parametrized the ultraviolet divergence in (6.54), let us try to interpret it. Notice that the divergence appears in the worst possible place: It corrects $F_1(q^2 = 0)$, which should (according to our discussion at the end of the previous section) be fixed at the value 1. But this is the only effect of the divergent term. We will therefore adopt a simple but completely *ad hoc* fix for this difficulty: Subtract from the above expression a term proportional to the zeroth-order vertex function ($\bar{u}(p')\gamma^\mu u(p)$), in such a way as to maintain the condition $F_1(0) = 1$. In other words, make the substitution

$$\delta F_1(q^2) \rightarrow \delta F_1(q^2) - \delta F_1(0) \quad (6.55)$$

(where δF_1 denotes the first-order correction to F_1). The justification of this procedure involves the minor correction to our S -matrix formula (4.103) mentioned in Section 4.5. In brief, the term we are subtracting corrects for our omission of the external leg correction diagrams of (6.1). We postpone the justification of this statement until Section 7.2.

There is also an infrared divergence in $F_1(q^2)$, coming from the $1/\Delta$ term. For example, at $q^2 = 0$ this term is

$$\begin{aligned} \int_0^1 dx dy dz \delta(x+y+z-1) \frac{1-4z+z^2}{\Delta(q^2=0)} &= \int_0^1 dz \int_0^{1-z} dy \frac{-2 + (1-z)(3-z)}{m^2(1-z)^2} \\ &= \int_0^1 dz \frac{-2}{m^2(1-z)} + \text{finite terms.} \end{aligned}$$

We can cure this disease by pretending that the photon has a small nonzero mass μ . Then in the denominator of the photon propagator, $(k-p)^2$ would become $(k-p)^2 - \mu^2$. This denominator was multiplied by z in (6.43), so the net effect is to add a term $z\mu^2$ to Δ . We will discuss the infrared divergence further in the next two sections.

With both of these provisional modifications, the form factors are

$$F_1(q^2) = 1 + \frac{\alpha}{2\pi} \int_0^1 dx dy dz \delta(x+y+z-1) \times \left[\log\left(\frac{m^2(1-z)^2}{m^2(1-z)^2 - q^2 xy}\right) + \frac{m^2(1-4z+z^2) + q^2(1-x)(1-y)}{m^2(1-z)^2 - q^2 xy + \mu^2 z} - \frac{m^2(1-4z+z^2)}{m^2(1-z)^2 + \mu^2 z} \right] + \mathcal{O}(\alpha^2); \quad (6.56)$$

$$F_2(q^2) = \frac{\alpha}{2\pi} \int_0^1 dx dy dz \delta(x+y+z-1) \left[\frac{2m^2 z(1-z)}{m^2(1-z)^2 - q^2 xy} \right] + \mathcal{O}(\alpha^2). \quad (6.57)$$

Note that neither the ultraviolet nor the infrared divergence affects $F_2(q^2)$. We can therefore evaluate unambiguously

$$F_2(q^2 = 0) = \frac{\alpha}{2\pi} \int_0^1 dx dy dz \delta(x + y + z - 1) \frac{2m^2 z(1-z)}{m^2(1-z)^2} = \frac{\alpha}{\pi} \int_0^1 dz \int_0^{1-z} dy \frac{z}{1-z} = \frac{\alpha}{2\pi}. \quad (6.58)$$

Thus, we get a correction to the g -factor of the electron:

$$a_e \equiv \frac{g-2}{2} = \frac{\alpha}{2\pi} \approx .0011614. \quad (6.59)$$

This result was first obtained by Schwinger in 1948.* Experiments give $a_e = .0011597$. Apparently, the unambiguous value that we obtained for $F_2(0)$ is also, up to higher orders in α , unambiguously correct.

Precision Tests of QED

Building on the success of the order- α QED prediction for a_e , successive generations of physicists have improved the accuracy of both the theoretical and the experimental determination of this quantity. The coefficients of the QED formula for a_e are now known through order α^4 . The calculation of the order- α^2 and higher coefficients requires a systematic treatment of ultraviolet divergences.

These challenging theoretical calculations have been matched by increasingly imaginative experiments. The most recent measurement of a_e uses a technique, developed by Dehmelt and collaborators, in which individual electrons are trapped in a system of electrostatic and magnetostatic fields and

*J. Schwinger, *Phys. Rev.* **73**, 416L (1948).

excited to a spin resonance.[†] Today, the best theoretical and experimental values of a_e agree to eight significant figures.

High-order QED calculations have also been carried out for several other quantities. These include transition energies in hydrogen and hydrogen-like atoms, the anomalous magnetic moment of the muon, and the decay rates of singlet and triplet positronium. Many of these quantities have also been measured to high precision. The full set of these comparisons gives a detailed test of the validity of QED in a variety of settings. The results of these precision tests are summarized in Table 6.1.

There is some subtlety in reporting the results of precision comparisons between QED theory and experiment, since theoretical predictions require an extremely precise value of α , which can only be obtained from another precision QED experiment. We therefore quote each comparison between theory and experiment as an independent determination of α . Each value of α is assigned an error that is the composite of the expected uncertainties from theory and experiment. QED is confirmed to the extent that the values of α from different sources agree.

The first nine entries in Table 6.1 refer to QED calculations in atomic physics settings. Of these, the hydrogen hyperfine splitting, measured using Ramsey's hydrogen maser, is the most precisely known quantity in physics. Unfortunately, the influence of the internal structure of the proton leads to uncertainties that limit the accuracy with which this quantity can be predicted theoretically. The same difficulty applies to the Lamb shift, the splitting between the $j = 1/2$ $2S$ and $2P$ levels of hydrogen. The most accurate QED tests now come from systems that involve no strongly interacting particles, the electron $g-2$ and the hyperfine splitting in the $e^- \mu^+$ atom, muonium. The last entry in this group gives a new method for determining α , by converting a very accurate measurement of the neutron Compton wavelength, using accurately known mass ratios, to a value of the electron mass. This can be combined with the known value of the Rydberg energy and accurate QED formulae to determine α . The only serious discrepancy among these numbers comes in the triplet positronium decay rate; however, there is some evidence that diagrams of relative order α^2 give a large correction to the value quoted in the table.

The next two entries are determinations of α from higher-order QED reactions at high-energy electron colliders. These high-energy experiments typically achieve only percent-level accuracy, but their results are consistent with the precise information available at lower energies.

Finally, the last two entries in the table give two independent measurements of α from exotic quantum interference phenomena in condensed-matter systems. These two effects provide a standard resistance and a standard frequency, respectively, which are believed to measure the charge of the electron

[†]R. Van Dyck, Jr., P. Schwinberg, and H. Dehmelt, *Phys. Rev. Lett.* **59**, 26 (1987).

Table 6.1. Values of α^{-1} Obtained from Precision QED Experiments

Low-Energy QED:

Electron ($g - 2$)	137.035 992 35 (73)
Muon ($g - 2$)	137.035 5 (1 1)
Muonium hyperfine splitting	137.035 994 (18)
Lamb shift	137.036 8 (7)
Hydrogen hyperfine splitting	137.036 0 (3)
2^3S_1 - 1^3S_1 splitting in positronium	137.034 (16)
1S_0 positronium decay rate	137.00 (6)
3S_1 positronium decay rate	136.971 (6)
Neutron compton wavelength	137.036 010 1 (5 4)

High-Energy QED:

$\sigma(e^+e^- \rightarrow e^+e^-e^+e^-)$	136.5 (2.7)
$\sigma(e^+e^- \rightarrow e^+e^-\mu^+\mu^-)$	139.9 (1.2)

Condensed Matter:

Quantum Hall effect	137.035 997 9 (3 2)
AC Josephson effect	137.035 977 0 (7 7)

Each value of α displayed in this table is obtained by fitting an experimental measurement to a theoretical expression that contains α as a parameter. The numbers in parentheses are the standard errors in the last displayed digits, including both theoretical and experimental uncertainties. This table is based on results presented in the survey of precision QED of Kinoshita (1990). That book contains a series of lucid reviews of the remarkable theoretical and experimental technology that has been developed for the detailed analysis of QED processes. The five most accurate values are updated as given by T. Kinoshita in *History of Original Ideas and Basic Discoveries in Particle Physics*, H. Newman and T. Ypsilantis, eds. (Plenum Press, New York, 1995). This latter paper also gives an interesting perspective on the future of precision QED experiments.

with corrections that are strictly zero for macroscopic systems.[‡]

The entire picture fits together well beyond any reasonable expectation. On the evidence presented in this table, QED is the most stringently tested—and the most dramatically successful—of all physical theories.

[‡]For a discussion of these effects, and their exact relation to α , see D. R. Yennie, *Rev. Mod. Phys.* **59**, 781 (1987).

6.4 The Electron Vertex Function: Infrared Divergence

Now let us confront the infrared divergence in our result (6.56) for $F_1(q^2)$. The dominant part, in the $\mu \rightarrow 0$ limit, is

$$F_1(q^2) \approx \frac{\alpha}{2\pi} \int_0^1 dx dy dz \delta(x+y+z-1) \left[\frac{m^2(1-4z+z^2) + q^2(1-x)(1-y)}{m^2(1-z)^2 - q^2xy + \mu^2 z} - \frac{m^2(1-4z+z^2)}{m^2(1-z)^2 + \mu^2 z} \right]. \quad (6.60)$$

To understand this expression we must do some work to simplify it, extracting and evaluating the divergent part of the integral. Throughout this section we will retain only terms that diverge in the limit $\mu \rightarrow 0$.

First note that the divergence occurs in the corner of Feynman-parameter space where $z \approx 1$ (and therefore $x \approx y \approx 0$). In this region we can set $z = 1$ and $x = y = 0$ in the numerators of (6.60). We can also set $z = 1$ in the μ^2 terms in the denominators. Using the delta function to evaluate the x -integral, we then have

$$F_1(q^2) = \frac{\alpha}{2\pi} \int_0^1 dz \int_0^{1-z} dy \left[\frac{-2m^2 + q^2}{m^2(1-z)^2 - q^2y(1-z-y) + \mu^2} - \frac{-2m^2}{m^2(1-z)^2 + \mu^2} \right].$$

(The lower limit on the z -integral is unimportant.) Making the variable changes

$$y = (1-z)\xi, \quad w = (1-z),$$

this expression becomes

$$\begin{aligned} F_1(q^2) &= \frac{\alpha}{2\pi} \int_0^1 d\xi \frac{1}{2} \int_0^1 d(w^2) \left[\frac{-2m^2 + q^2}{[m^2 - q^2\xi(1-\xi)]w^2 + \mu^2} - \frac{-2m^2}{m^2w^2 + \mu^2} \right] \\ &= \frac{\alpha}{4\pi} \int_0^1 d\xi \left[\frac{-2m^2 + q^2}{m^2 - q^2\xi(1-\xi)} \log\left(\frac{m^2 - q^2\xi(1-\xi)}{\mu^2}\right) + 2 \log\left(\frac{m^2}{\mu^2}\right) \right]. \end{aligned}$$

In the limit $\mu \rightarrow 0$ we can ignore the details of the numerators inside the logarithms; anything proportional to m^2 or q^2 is effectively the same. We therefore write

$$F_1(q^2) = 1 - \frac{\alpha}{2\pi} f_{\text{IR}}(q^2) \log\left(\frac{-q^2 \text{ or } m^2}{\mu^2}\right) + \mathcal{O}(\alpha^2), \quad (6.61)$$

where the coefficient of the divergent logarithm is

$$f_{\text{IR}}(q^2) = \int_0^1 \left(\frac{m^2 - q^2/2}{m^2 - q^2\xi(1-\xi)} \right) d\xi - 1. \quad (6.62)$$

Since q^2 is negative and $\xi(1-\xi)$ has a maximum value of 1/4, the first term is greater than 1 and hence $f_{\text{IR}}(q^2)$ is positive.

How does this infinite term affect the cross section for electron scattering off a potential? Since $F_1(q^2)$ is just the quantity that multiplies γ^μ in the matrix element, we can find the new cross section by making the replacement $e \rightarrow e \cdot F_1(q^2)$. The cross section for the process $\mathbf{p} \rightarrow \mathbf{p}'$ is therefore

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{d\sigma}{d\Omega} \right)_0 \cdot \left[1 - \frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{-q^2 \text{ or } m^2}{\mu^2}\right) + \mathcal{O}(\alpha^2) \right], \quad (6.63)$$

where the first factor is the tree-level result. Note that the $\mathcal{O}(\alpha)$ correction to the cross section is not only infinite, but negative. Something is terribly wrong.

To gain a better understanding of the divergence, let us evaluate the coefficient of the divergent logarithm, $f_{\text{IR}}(q^2)$, in the limit $-q^2 \rightarrow \infty$. In this limit, we find a second logarithm:

$$\begin{aligned} \int_0^1 d\xi \frac{-q^2/2}{-q^2\xi(1-\xi) + m^2} &\simeq \frac{1}{2} \int_0^1 d\xi \frac{-q^2}{-q^2\xi + m^2} + \left(\begin{array}{c} \text{equal contribution} \\ \text{from } \xi \approx 1 \end{array} \right) \\ &= \log\left(\frac{-q^2}{m^2}\right). \end{aligned} \quad (6.64)$$

The form factor in this limit is therefore

$$F_1(-q^2 \rightarrow \infty) = 1 - \frac{\alpha}{2\pi} \log\left(\frac{-q^2}{m^2}\right) \log\left(\frac{-q^2}{\mu^2}\right) + \mathcal{O}(\alpha^2). \quad (6.65)$$

Note that the numerator in the second logarithm is $-q^2$, not m^2 ; this expression contains not only the correct coefficient of $\log(1/\mu^2)$, but also the correct coefficient of $\log^2(q^2)$.

The same double logarithm of $-q^2$ appeared in the cross section for soft bremsstrahlung, Eq. (6.26). This correspondence points to a resolution of the infrared divergence problem. Comparing (6.65) with (6.26), we find in the limit $-q^2 \rightarrow \infty$

$$\begin{aligned} \frac{d\sigma}{d\Omega}(p \rightarrow p') &= \left(\frac{d\sigma}{d\Omega} \right)_0 \left[1 - \frac{\alpha}{\pi} \log\left(\frac{-q^2}{m^2}\right) \log\left(\frac{-q^2}{\mu^2}\right) + \mathcal{O}(\alpha^2) \right]; \\ \frac{d\sigma}{d\Omega}(p \rightarrow p' + \gamma) &= \left(\frac{d\sigma}{d\Omega} \right)_0 \left[+ \frac{\alpha}{\pi} \log\left(\frac{-q^2}{m^2}\right) \log\left(\frac{-q^2}{\mu^2}\right) + \mathcal{O}(\alpha^2) \right]. \end{aligned} \quad (6.66)$$

The separate cross sections are divergent, but their sum is independent of μ and therefore finite.

In fact, neither the elastic cross section nor the soft bremsstrahlung cross section can be measured individually; only their sum is physically observable. In any real experiment, a photon detector can detect photons only down to

some minimum limiting energy E_ℓ . The probability that a scattering event occurs and this detector does not see a photon is the sum

$$\frac{d\sigma}{d\Omega}(p \rightarrow p') + \frac{d\sigma}{d\Omega}(p \rightarrow p' + \gamma(k < E_\ell)) \equiv \left(\frac{d\sigma}{d\Omega}\right)_{\text{measured}}. \quad (6.67)$$

The divergent part of this “measured” cross section is

$$\begin{aligned} \left(\frac{d\sigma}{d\Omega}\right)_{\text{measured}} &\approx \left(\frac{d\sigma}{d\Omega}\right)_0 \left[1 - \frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{-q^2 \text{ or } m^2}{\mu^2}\right) \right. \\ &\quad \left. + \frac{\alpha}{2\pi} \mathcal{I}(\mathbf{v}, \mathbf{v}') \log\left(\frac{E_\ell^2}{\mu^2}\right) + \mathcal{O}(\alpha^2) \right]. \end{aligned}$$

We have just seen that $\mathcal{I}(\mathbf{v}, \mathbf{v}') = 2f_{\text{IR}}(q^2)$ when $-q^2 \gg m^2$. If the same relation holds for general q^2 , the measured cross section becomes

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{measured}} \approx \left(\frac{d\sigma}{d\Omega}\right)_0 \left[1 - \frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{-q^2 \text{ or } m^2}{E_\ell^2}\right) + \mathcal{O}(\alpha^2) \right], \quad (6.68)$$

which depends on the experimental conditions, but no longer on μ^2 . The infrared divergences from soft bremsstrahlung and from $F_1(q^2)$ cancel each other, yielding a finite cross section for the quantity that can actually be measured.

We must still verify the identity $\mathcal{I}(\mathbf{v}, \mathbf{v}') = 2f_{\text{IR}}(q^2)$ for arbitrary values of q^2 . From (6.13) we have

$$\mathcal{I}(\mathbf{v}, \mathbf{v}') = \int \frac{d\Omega_{\mathbf{k}}}{4\pi} \left(\frac{2p \cdot p'}{(\hat{k} \cdot p')(\hat{k} \cdot p)} - \frac{m^2}{(\hat{k} \cdot p')^2} - \frac{m^2}{(\hat{k} \cdot p)^2} \right). \quad (6.69)$$

The last two terms are easy to evaluate:

$$\int \frac{d\Omega_{\mathbf{k}}}{4\pi} \frac{1}{(\hat{k} \cdot p)^2} = \frac{1}{2} \int_{-1}^1 d\cos\theta \frac{1}{(p^0 - p \cos\theta)^2} = \frac{1}{p^2} = \frac{1}{m^2}.$$

In the first term, we can combine the denominators with a Feynman parameter and perform the integral in the same way:

$$\begin{aligned} \int \frac{d\Omega_{\mathbf{k}}}{4\pi} \frac{1}{(\hat{k} \cdot p')(\hat{k} \cdot p)} &= \int_0^1 d\xi \int \frac{d\Omega_{\mathbf{k}}}{4\pi} \frac{1}{[\xi \hat{k} \cdot p' + (1-\xi) \hat{k} \cdot p]^2} \\ &= \int_0^1 d\xi \frac{1}{[\xi p' + (1-\xi)p]^2} = \int_0^1 d\xi \frac{1}{m^2 - \xi(1-\xi)q^2}. \end{aligned}$$

(In the last step we have used $2p \cdot p' = 2m^2 - q^2$.) Putting all the terms of (6.69) together, we find

$$\mathcal{I}(\mathbf{v}, \mathbf{v}') = \int_0^1 \left(\frac{2m^2 - q^2}{m^2 - \xi(1-\xi)q^2} \right) d\xi - 2 = 2f_{\text{IR}}(q^2), \quad (6.70)$$

just what we need to cancel the infrared divergence.

Although Eq. (6.68) demonstrates the cancellation of the infrared divergence, this result has little practical use. An experimentalist would want to know the precise dependence on q^2 , which we did not evaluate carefully. Recall from (6.65), however, that we were careful to obtain the correct coefficient of $\log^2(-q^2)$ in the limit $-q^2 \gg m^2$. In that limit, therefore, (6.68) becomes

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{measured}} \approx \left(\frac{d\sigma}{d\Omega}\right)_0 \left[1 - \frac{\alpha}{\pi} \log\left(\frac{-q^2}{m^2}\right) \log\left(\frac{-q^2}{E_\ell^2}\right) + \mathcal{O}(\alpha^2) \right]. \quad (6.71)$$

This result is unambiguous and useful. Note that the $\mathcal{O}(\alpha)$ correction again involves the Sudakov double logarithm.

6.5 Summation and Interpretation of Infrared Divergences

The discussion of infrared divergences in the previous section suffices for removing the infinities from our bremsstrahlung and vertex-correction calculations. There are still, however, three points that we have not addressed:

1. We have not demonstrated the cancellation of infrared divergences beyond the leading order.
2. The correction to the measured cross section that we found after the infrared cancellation (Eqs. (6.68) and (6.71)) can be made arbitrarily negative by making photon detectors with a sufficiently low threshold E_ℓ .
3. We have not yet reproduced the classical result (6.19) for the number of photons radiated during a collision.

The solutions of the second and third problems will follow immediately from that of the first, to which we now turn.

A complete treatment of infrared divergences to all orders is beyond the scope of this book.* We will discuss here only the terms with the largest logarithmic enhancement at each order of perturbation theory. In general, these terms are of order

$$\left[\frac{\alpha}{\pi} \log\left(\frac{-q^2}{\mu^2}\right) \log\left(\frac{-q^2}{m^2}\right) \right]^n \quad (6.72)$$

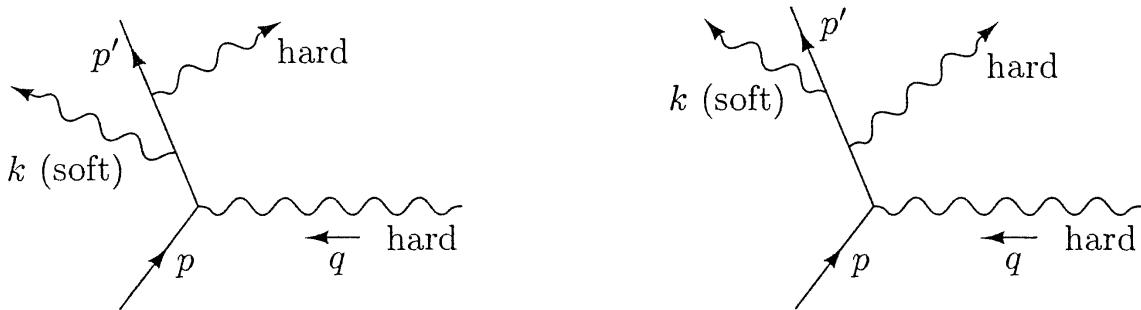
in the n th order of perturbation theory. Our final physical conclusions were first presented by Bloch and Nordsieck in a prescient paper written before the invention of relativistic perturbation theory.[†] We will follow a modern, and simplified, version of the analysis due to Weinberg.[‡]

*The definitive treatment is given in D. Yennie, S. Frautschi, and H. Suura, *Ann. Phys.* **13**, 379 (1961).

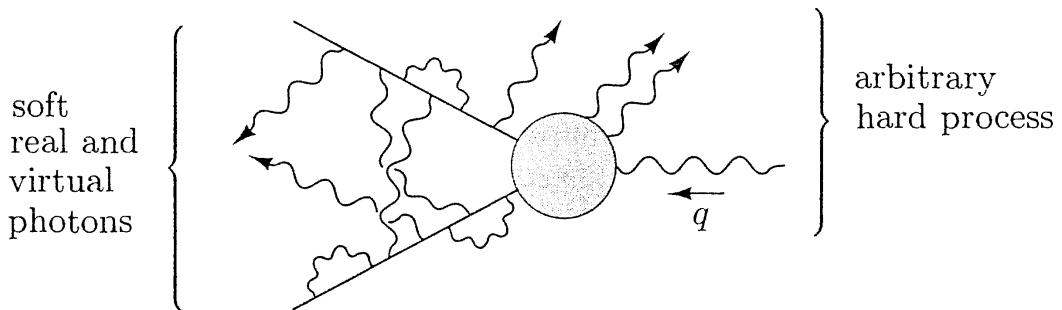
[†]F. Bloch and A. Nordsieck, *Phys. Rev.* **52**, 54 (1937).

[‡]S. Weinberg, *Phys. Rev.* **140**, B516 (1965).

Infrared divergences arise from photons with “soft” momenta: real photons with energy less than some cutoff E_ℓ , and virtual photons with (after Wick rotation) $k^2 < E_\ell^2$. A typical higher-order diagram will involve numerous real and virtual photons. But to find a divergence, we need more than a soft photon; we need a singular denominator in an electron propagator. Consider, for example, the following two diagrams:

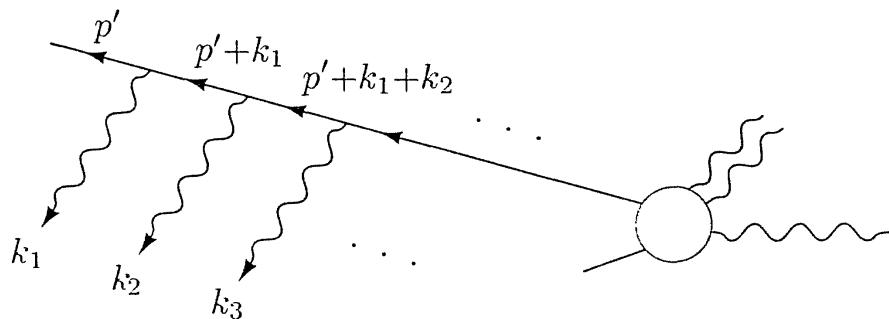


The first diagram, in which the electron emits a soft photon followed by a hard photon, has no infrared divergence, since the momenta in both electron propagators are far from the mass shell. If the soft photon is emitted last, however, the denominator of the adjacent propagator is $(p' + k)^2 - m^2 = 2p' \cdot k$, which vanishes as $k \rightarrow 0$. Thus the second diagram does contain a divergence. We would like, then, to consider diagrams in which an arbitrary hard process, possibly involving emission of hard and soft photons, is modified by the addition of soft real and virtual photons on the electron legs:



Following Weinberg, we will add up the contributions of all such diagrams. The only new difficulty in this calculation will be in the combinatorics of counting all the ways in which the photons can appear.

First consider the outgoing electron line:



We attach n photons to the line, with momenta $k_1 \dots k_n$. For the moment we do not care whether these are external photons, virtual photons connected to each other, or virtual photons connected to vertices on the incoming electron line. The Dirac structure of this diagram is

$$\bar{u}(p')(-ie\gamma^{\mu_1}) \frac{i(\not{p}'' + \not{k}_1 + m)}{2p' \cdot k_1} (-ie\gamma^{\mu_2}) \frac{i(\not{p}'' + \not{k}_1 + \not{k}_2 + m)}{2p' \cdot (k_1 + k_2) + \mathcal{O}(k^2)} \dots (-ie\gamma^{\mu_n}) \frac{i(\not{p}'' + \not{k}_1 + \dots + \not{k}_n + m)}{2p' \cdot (k_1 + \dots + k_n) + \mathcal{O}(k^2)} (i\mathcal{M}_{\text{hard}}) \dots \quad (6.73)$$

We will assume that all the k_i are small, dropping the $\mathcal{O}(k^2)$ terms in the denominators. We will also drop the \not{k}_i terms in the numerators, just as in our treatment of bremsstrahlung in Section 6.1. Also, as we did there, we can push the factors of $(\not{p}'' + m)$ to the left and use $\bar{u}(p')(-\not{p}'' + m) = 0$:

$$\begin{aligned} \bar{u}(p') \gamma^{\mu_1} (\not{p}'' + m) \gamma^{\mu_2} (\not{p}'' + m) \dots &= \bar{u}(p') 2p'^{\mu_1} \gamma^{\mu_2} (\not{p}'' + m) \dots \\ &= \bar{u}(p') 2p'^{\mu_1} 2p'^{\mu_2} \dots . \end{aligned}$$

This turns expression (6.73) into

$$\bar{u}(p') \left(e \frac{p'^{\mu_1}}{p' \cdot k_1} \right) \left(e \frac{p'^{\mu_2}}{p' \cdot (k_1 + k_2)} \right) \dots \left(e \frac{p'^{\mu_n}}{p' \cdot (k_1 + \dots + k_n)} \right) \dots \quad (6.74)$$

Still working with only the outgoing electron line, we must now sum over all possible orderings of the momenta $k_1 \dots k_n$. (This procedure will overcount when two of the photons are attached together to form a single virtual photon. We will deal with this overcounting later.) There are $n!$ different diagrams to sum, corresponding to the $n!$ permutations of the n photon momenta. Let π denote one such permutation, so that $\pi(i)$ is the number between 1 and n that i is taken to. (For example, if π denotes the permutation that takes $1 \rightarrow 3$, $2 \rightarrow 1$ and $3 \rightarrow 2$, then $\pi(1) = 3$, $\pi(2) = 1$, and $\pi(3) = 2$.)

Armed with this notation, we can perform the sum over permutations by means of the following identity:

$$\begin{aligned} \sum_{\text{all permutations } \pi} \frac{1}{p \cdot k_{\pi(1)}} \frac{1}{p \cdot (k_{\pi(1)} + k_{\pi(2)})} \dots \frac{1}{p \cdot (k_{\pi(1)} + k_{\pi(2)} + \dots + k_{\pi(n)})} \\ = \frac{1}{p \cdot k_1} \frac{1}{p \cdot k_2} \dots \frac{1}{p \cdot k_n}. \end{aligned} \quad (6.75)$$

The proof of this formula proceeds by induction on n . For $n = 2$ we have

$$\begin{aligned} \sum_{\pi} \frac{1}{p \cdot k_{\pi(1)}} \frac{1}{p \cdot (k_{\pi(1)} + k_{\pi(2)})} &= \frac{1}{p \cdot k_1} \frac{1}{p \cdot (k_1 + k_2)} + \frac{1}{p \cdot k_2} \frac{1}{p \cdot (k_2 + k_1)} \\ &= \frac{1}{p \cdot k_1} \frac{1}{p \cdot k_2}. \end{aligned}$$

For the induction step, notice that the last factor on the left-hand side of (6.75) is the same for every permutation π . Pulling this factor outside the sum, the left-hand side becomes

$$\text{LHS} = \frac{1}{p \cdot \sum k} \sum_{\pi} \frac{1}{p \cdot k_{\pi(1)}} \frac{1}{p \cdot (k_{\pi(1)} + k_{\pi(2)})} \cdots \frac{1}{p \cdot (k_{\pi(1)} + \cdots + k_{\pi(n-1)})}.$$

For any given π , the quantity being summed is independent of $k_{\pi(n)}$. Letting $i = \pi(n)$, we can now write

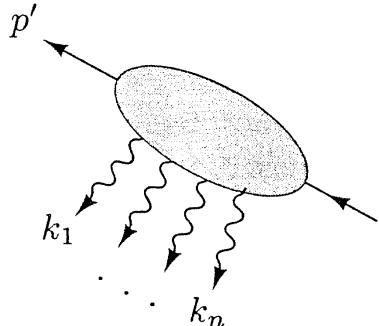
$$\sum_{\pi} = \sum_{i=1}^n \sum_{\pi'(i)},$$

where $\pi'(i)$ is the set of all permutations on the remaining $n - 1$ integers. Assuming by induction that (6.75) is true for $n - 1$, we have

$$\text{LHS} = \frac{1}{p \cdot \sum k} \sum_{i=1}^n \frac{1}{p \cdot k_1} \frac{1}{p \cdot k_2} \cdots \frac{1}{p \cdot k_{i-1}} \frac{1}{p \cdot k_{i+1}} \cdots \frac{1}{p \cdot k_n}.$$

If we now multiply and divide each term in this sum by $p \cdot k_i$, we easily obtain our desired result (6.75).

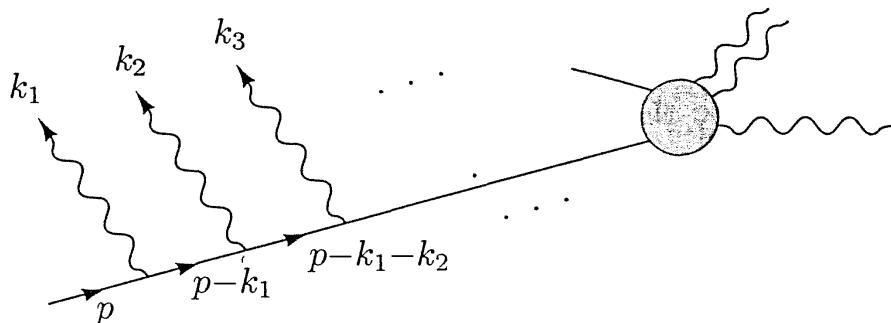
Applying (6.75) to (6.74), we find



$$= \bar{u}(p') \left(e^{\frac{p'^{\mu_1}}{p' \cdot k_1}} \right) \left(e^{\frac{p'^{\mu_2}}{p' \cdot k_2}} \right) \cdots \left(e^{\frac{p'^{\mu_n}}{p' \cdot k_n}} \right), \quad (6.76)$$

where the blob denotes a sum over all possible orders of inserting the n photon lines.

A similar set of manipulations simplifies the sum over soft photon insertions on the initial electron line. There, however, the propagator momenta are $p - k_1$, $p - k_1 - k_2$, and so on:



We therefore get an extra minus sign in the factor for each photon, since $(p - \sum k)^2 - m^2 \approx -2 \cdot \sum k$.

Now consider diagrams containing a total of n soft photons, connected in any possible order to the initial or final electron lines. The sum over all such diagrams can be written

$$\begin{aligned}
 & \text{Diagram: } \text{hard} \quad \cdot e\left(\frac{p'^{\mu_1}}{p' \cdot k_1} - \frac{p^{\mu_1}}{p \cdot k_1}\right) \cdot e\left(\frac{p'^{\mu_2}}{p' \cdot k_2} - \frac{p^{\mu_2}}{p \cdot k_2}\right) \\
 & \quad \quad \quad \cdots e\left(\frac{p'^{\mu_n}}{p' \cdot k_n} - \frac{p^{\mu_n}}{p \cdot k_n}\right). \quad (6.77)
 \end{aligned}$$

By multiplying out all the factors, you can see that we get the correct term for each possible way of dividing the n photons between the two lines.

Next we must decide which photons are real and which are virtual.

We can make a virtual photon by picking two photon momenta k_i and k_j , setting $k_j = -k_i \equiv k$, multiplying by the photon propagator, and integrating over k . For each virtual photon we then obtain the expression

$$\frac{e^2}{2} \int \frac{d^4 k}{(2\pi)^4} \frac{-i}{k^2 + i\epsilon} \left(\frac{p'}{p' \cdot k} - \frac{p}{p \cdot k} \right) \cdot \left(\frac{p'}{-p' \cdot k} - \frac{p}{-p \cdot k} \right) \equiv \mathbf{X}. \quad (6.78)$$

The factor of $1/2$ is required because our procedure has counted each Feynman diagram twice: interchanging k_i and k_j gives back the same diagram. It is possible to evaluate this expression by careful contour integration, but there is an easier way. Notice that this approximation scheme assigns to the diagram with one loop and no external photons the value

$$\bar{u}(p') (i\mathcal{M}_{\text{hard}}) u(p) \cdot \mathbf{X}.$$

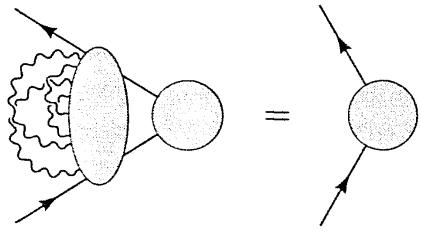
Thus, \mathbf{X} must be precisely the infrared limit of the one-loop correction to the form factor, as displayed in (6.61):

$$\mathbf{X} = -\frac{\alpha}{2\pi} f_{\text{IR}}(q^2) \log\left(\frac{-q^2}{\mu^2}\right). \quad (6.79)$$

A direct derivation of this result from (6.78) is given in Weinberg's paper cited above. Note that result (6.79) followed in our argument of the previous section only after the subtraction at $q^2 = 0$, and so we should worry whether (6.79) is consistent with the corresponding subtraction of the n th-order diagram. In addition, some of the diagrams we are summing contain external-leg corrections, which we have not discussed. Here we simply remark that neither of these subtleties affects the final answer; the proof requires the heavy machinery in the paper of Yennie, Frautschi, and Suura.

If there are m virtual photons we get m factors like (6.79), and also an additional symmetry factor of $1/m!$ since interchanging virtual photons with each other does not change the diagram. We can then sum over m to obtain

the complete correction due to the presence of arbitrarily many soft virtual photons:



$$= \sum_{m=0}^{\infty} \frac{\mathbf{X}^m}{m!} = \bar{u}(p') (i\mathcal{M}_{\text{hard}}) u(p) \exp(\mathbf{X}). \quad (6.80)$$

If in addition to the m virtual photons we also emit a real photon, we must multiply by its polarization vector, sum over polarizations, and integrate the squared matrix element over the photon's phase space. This gives an additional factor

$$\int \frac{d^3 k}{(2\pi)^3} \frac{1}{2k} e^2 (-g_{\mu\nu}) \left(\frac{p'^\mu}{p' \cdot k} - \frac{p^\mu}{p \cdot k} \right) \left(\frac{p'^\nu}{p' \cdot k} - \frac{p^\nu}{p \cdot k} \right) \equiv \mathbf{Y} \quad (6.81)$$

in the cross section. Assuming that the energy of the photon is greater than μ and less than E_ℓ (the detector threshold), this expression is simply

$$\mathbf{Y} = \frac{\alpha}{\pi} \mathcal{I}(\mathbf{v}, \mathbf{v}') \log\left(\frac{E_\ell}{\mu}\right) = \frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{E_\ell^2}{\mu^2}\right). \quad (6.82)$$

If n real photons are emitted we get n such factors, and also a symmetry factor of $1/n!$ since there are n identical bosons in the final state. The cross section for emission of any number of soft photons is therefore

$$\sum_{n=0}^{\infty} \frac{d\sigma}{d\Omega}(\mathbf{p} \rightarrow \mathbf{p}' + n\gamma) = \frac{d\sigma}{d\Omega}(\mathbf{p} \rightarrow \mathbf{p}') \cdot \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{Y}^n = \frac{d\sigma}{d\Omega}(\mathbf{p} \rightarrow \mathbf{p}') \cdot \exp(\mathbf{Y}). \quad (6.83)$$

Combining our results for virtual and real photons gives our final result for the measured cross section, to all orders in α , for the process $\mathbf{p} \rightarrow \mathbf{p}' +$ (any number of photons with $k < E_\ell$):

$$\begin{aligned} \left(\frac{d\sigma}{d\Omega} \right)_{\text{meas.}} &= \left(\frac{d\sigma}{d\Omega} \right)_0 \times \exp(2\mathbf{X}) \times \exp(\mathbf{Y}) \\ &= \left(\frac{d\sigma}{d\Omega} \right)_0 \times \exp\left[-\frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{-q^2}{\mu^2}\right)\right] \times \exp\left[\frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{E_\ell^2}{\mu^2}\right)\right] \\ &= \left(\frac{d\sigma}{d\Omega} \right)_0 \times \exp\left[-\frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{-q^2}{E_\ell^2}\right)\right]. \end{aligned} \quad (6.84)$$

The correction factor depends on the detector sensitivity E_ℓ , but is independent of the infrared cutoff μ . Note that if we expand this result to $\mathcal{O}(\alpha)$, we recover our earlier result (6.68). Now, however, the correction factor is controlled in magnitude—always between 0 and 1.

In the limit $-q^2 \gg m^2$, our result becomes

$$\left(\frac{d\sigma}{d\Omega} \right)_{\text{meas.}} = \left(\frac{d\sigma}{d\Omega} \right)_0 \times \left| \exp\left[-\frac{\alpha}{2\pi} \log\left(\frac{-q^2}{m^2}\right) \log\left(\frac{-q^2}{E_\ell^2}\right)\right] \right|^2. \quad (6.85)$$

In this limit, the probability of scattering without emitting a hard photon decreases faster than any power of q^2 . The exponential correction factor, containing the Sudakov double logarithm, is known as the *Sudakov form factor*.

To conclude this section, let us calculate the probability, in the same approximation, that some hard scattering process is accompanied by the production of n soft photons, all with energies between E_- and E_+ . The phase-space integral for these photons gives $\log(E_+/E_-)$ instead of $\log(E_\ell/\mu)$. If we assign photons with energy greater than E_+ to the “hard” part of the process, we find that the cross section is given by (6.84), times the additional factor

$$\begin{aligned} \text{Prob}(n\gamma \text{ with } E_- < E < E_+) &= \frac{1}{n!} \left[\frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{E_+^2}{E_-^2}\right) \right]^n \\ &\quad \times \exp\left[-\frac{\alpha}{\pi} f_{\text{IR}}(q^2) \log\left(\frac{E_+^2}{E_-^2}\right)\right]. \end{aligned} \quad (6.86)$$

This expression has the form of a Poisson distribution,

$$P(n) = \frac{1}{n!} \lambda^n e^{-\lambda},$$

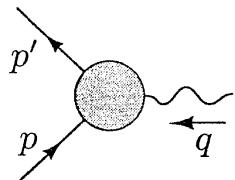
with

$$\lambda = \langle n \rangle = \frac{\alpha}{\pi} \log\left(\frac{E_+}{E_-}\right) \mathcal{I}(\mathbf{v}, \mathbf{v}').$$

This is precisely the semiclassical estimate of the number of radiated photons that we made in Eq. (6.19).

Problems

6.1 Rosenbluth formula. As discussed Section 6.2, the exact electromagnetic interaction vertex for a Dirac fermion can be written quite generally in terms of two form factors $F_1(q^2)$ and $F_2(q^2)$:



$$= \bar{u}(p') \left[\gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu} q_\nu}{2m} F_2(q^2) \right] u(p),$$

where $q = p' - p$ and $\sigma^{\mu\nu} = \frac{1}{2}i[\gamma^\mu, \gamma^\nu]$. If the fermion is a strongly interacting particle such as the proton, the form factors reflect the structure that results from the strong interactions and so are not easy to compute from first principles. However, these form factors can be determined experimentally. Consider the scattering of an electron with energy $E \gg m_e$ from a proton initially at rest. Show that the above expression for the vertex leads to the following expression (the Rosenbluth formula) for the elastic scattering cross section, computed to leading order in α but to all orders in the strong interactions:

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2 \left[(F_1^2 - \frac{q^2}{4m^2} F_2^2) \cos^2 \frac{\theta}{2} - \frac{q^2}{2m^2} (F_1 + F_2)^2 \sin^2 \frac{\theta}{2} \right]}{2E^2 \left[1 + \frac{2E}{m} \sin^2 \frac{\theta}{2} \right] \sin^4 \frac{\theta}{2},}$$

where θ is the lab-frame scattering angle and F_1 and F_2 are to be evaluated at the q^2 associated with elastic scattering at this angle. By measuring $(d\sigma/d\cos\theta)$ as a function of angle, it is thus possible to extract F_1 and F_2 . Note that when $F_1 = 1$ and $F_2 = 0$, the Rosenbluth formula reduces to the Mott formula (in the massless limit) for scattering off a point particle (see Problem 5.1).

6.2 Equivalent photon approximation. Consider the process in which electrons of very high energy scatter from a target. In leading order in α , the electron is connected to the target by one photon propagator. If the initial and final energies of the electron are E and E' , the photon will carry momentum q such that $q^2 \approx -2EE'(1 - \cos\theta)$. In the limit of forward scattering, whatever the energy loss, the photon momentum approaches $q^2 = 0$; thus the reaction is highly peaked in the forward direction. It is tempting to guess that, in this limit, the virtual photon becomes a real photon. Let us investigate in what sense that is true.

- (a) The matrix element for the scattering process can be written as

$$\mathcal{M} = (-ie)\bar{u}(p')\gamma^\mu u(p) \frac{-ig_{\mu\nu}}{q^2} \hat{\mathcal{M}}^\nu(q),$$

where $\hat{\mathcal{M}}^\nu$ represents the (in general, complicated) coupling of the virtual photon to the target. Let us analyze the structure of the piece $\bar{u}(p')\gamma^\mu u(p)$. Let $q = (q^0, \mathbf{q})$, and define $\tilde{q}^\mu = (q^0, -\mathbf{q})$. We can expand the spinor product as:

$$\bar{u}(p')\gamma^\mu u(p) = A \cdot q^\mu + B \cdot \tilde{q}^\mu + C \cdot \epsilon_1^\mu + D \cdot \epsilon_2^\mu,$$

where A, B, C, D are functions of the scattering angle and energy loss and ϵ_i are two unit vectors transverse to \mathbf{q} . By dotting this expression with q_μ , show that the coefficient B is at most of order θ^2 . This will mean that we can ignore it in the rest of the analysis. The coefficient A is large, but it is also irrelevant, since, by the Ward identity, $q^\mu \hat{\mathcal{M}}_\mu = 0$.

- (b) Working in the frame where $p = (E, 0, 0, E)$, compute explicitly

$$\bar{u}(p')\gamma \cdot \epsilon_i u(p)$$

using massless electrons, $u(p)$ and $u(p')$ spinors of definite helicity, and ϵ_1, ϵ_2 unit vectors parallel and perpendicular to the plane of scattering. We need this quantity only for scattering near the forward direction, and we need only the term of order θ . Note, however, that for ϵ in the plane of scattering, the small $\hat{3}$ component of ϵ also gives a term of order θ which must be taken into account.

- (c) Now write the expression for the electron scattering cross section, in terms of $|\hat{\mathcal{M}}^\mu|^2$ and the integral over phase space on the target side. This expression must be integrated over the final electron momentum p' . The integral over p'^3 is an integral over the energy loss of the electron. Show that the integral over p'_\perp diverges logarithmically as $p'_\perp \rightarrow 0$.
- (d) The divergence as $\theta \rightarrow 0$ appears because we have ignored the electron mass in too many places. Show that reintroducing the electron mass in the expression for q^2 ,

$$q^2 = -2(EE' - pp' \cos\theta) + 2m^2,$$

cuts off the divergence and yields a factor of $\log(s/m^2)$ in its place.

- (e) Assembling all the factors, and assuming that the target cross sections are independent of the photon polarization, show that the largest part of the electron-target scattering cross section is given by considering the electron to be the source of a beam of real photons with energy distribution ($x = E_\gamma/E$):

$$N_\gamma(x)dx = \frac{dx}{x} \frac{\alpha}{2\pi} [1 + (1 - x)^2] \log\left(\frac{s}{m^2}\right).$$

This is the Weizsäcker-Williams *equivalent photon approximation*. This phenomenon allows us, for example, to study photon-photon scattering using e^+e^- collisions. Notice that the distribution we have found here is the same one that appeared in Problem 5.5 when we considered soft photon emission before electron scattering. It should be clear that a parallel general derivation can be constructed for that case.

6.3 Exotic contributions to $g - 2$. Any particle that couples to the electron can produce a correction to the electron-photon form factors and, in particular, a correction to $g - 2$. Because the electron $g - 2$ agrees with QED to high accuracy, these corrections allow us to constrain the properties of hypothetical new particles.

- (a) The unified theory of weak and electromagnetic interactions contains a scalar particle h called the *Higgs boson*, which couples to the electron according to

$$H_{\text{int}} = \int d^3x \frac{\lambda}{\sqrt{2}} h \bar{\psi} \psi.$$

Compute the contribution of a virtual Higgs boson to the electron ($g - 2$), in terms of λ and the mass m_h of the Higgs boson.

- (b) QED accounts extremely well for the electron's anomalous magnetic moment. If $a = (g - 2)/2$,

$$|a_{\text{expt.}} - a_{\text{QED}}| < 1 \times 10^{-10}.$$

What limits does this place on λ and m_h ? In the simplest version of the electroweak theory, $\lambda = 3 \times 10^{-6}$ and $m_h > 60$ GeV. Show that these values are not excluded. The coupling of the Higgs boson to the muon is larger by a factor (m_μ/m_e): $\lambda = 6 \times 10^{-4}$. Thus, although our experimental knowledge of the muon anomalous magnetic moment is not as precise,

$$|a_{\text{expt.}} - a_{\text{QED}}| < 3 \times 10^{-8},$$

one can still obtain a stronger limit on m_h . Is it strong enough?

- (c) Some more complex versions of this theory contain a pseudoscalar particle called the *axion*, which couples to the electron according to

$$H_{\text{int}} = \int d^3x \frac{i\lambda}{\sqrt{2}} a \bar{\psi} \gamma^5 \psi.$$

The axion may be as light as the electron, or lighter, and may couple more strongly than the Higgs boson. Compute the contribution of a virtual axion to the $g - 2$ of the electron, and work out the excluded values of λ and m_a .

Radiative Corrections: Some Formal Developments

We cheated four times in the last three chapters,* stating (and sometimes motivating) a result but postponing its proof. Those results were:

1. The formula for decay rates in terms of S -matrix elements, Eq. (4.86).
2. The master formula for S -matrix elements in terms of Feynman diagrams, Eq. (4.103).
3. The Ward identity, Eq. (5.79).
4. The *ad hoc* subtraction to remove the ultraviolet divergence in the vertex-correction diagram, Eq. (6.55).

It is time now to return to these issues and give them a proper treatment. In Sections 7.2 through 7.4 we will derive all four of these results. The knowledge we gain along the way will help us interpret the three remaining loop corrections for electron scattering from a heavy target shown in (6.1): the external leg corrections and the vacuum polarization. We will evaluate the former in Section 7.1 and the latter in Section 7.5.

This chapter will be more abstract than the two preceding ones. Its main theme will be the singularities of Feynman diagrams viewed as analytic functions of their external momenta. We will find, however, that this apparently esoteric subject is rich in physical implications, and that it illuminates the relation between Feynman diagrams and the general principles of quantum theory.

7.1 Field-Strength Renormalization

In this section we will investigate the analytic structure of the two-point correlation function,

$$\langle \Omega | T\phi(x)\phi(y) | \Omega \rangle \quad \text{or} \quad \langle \Omega | T\psi(x)\bar{\psi}(y) | \Omega \rangle .$$

In a free field theory, the two-point function $\langle 0 | T\phi(x)\phi(y) | 0 \rangle$ has a simple interpretation: It is the amplitude for a particle to propagate from y to x . To what extent does this interpretation carry over into an interacting theory?

*A fifth cheat, postulating rather than deriving the photon propagator, will be remedied in Chapter 9.

Our analysis of the two-point function will rely only on general principles of relativity and quantum mechanics; it will not depend on the nature of the interactions or on an expansion in perturbation theory. We will, however, restrict our consideration to scalar fields. Similar results can be obtained for correlation functions of fields with spin; we will display the analogous result for Dirac fields at the end of the analysis.

To dissect the two-point function $\langle \Omega | T\phi(x)\phi(y) | \Omega \rangle$ we will insert the identity operator, in the form of a sum over a complete set of states, between $\phi(x)$ and $\phi(y)$. We choose these states to be eigenstates of the full interacting Hamiltonian, H . Since the momentum operator \mathbf{P} commutes with H , we can also choose the states to be eigenstates of \mathbf{P} . But we can also make a stronger use of Lorentz invariance. Let $|\lambda_0\rangle$ be an eigenstate of H with momentum zero: $\mathbf{P}|\lambda_0\rangle = 0$. Then all the boosts of $|\lambda_0\rangle$ are also eigenstates of H , and these have all possible 3-momenta. Conversely, any eigenstate of H with definite momentum can be written as the boost of some zero-momentum eigenstate $|\lambda_0\rangle$. The eigenvalues of the 4-momentum operator $P^\mu = (H, \mathbf{P})$ organize themselves into hyperboloids, as shown in Fig. 7.1.

Recall from Chapter 2 that the completeness relation for the one-particle states is

$$(1)_{\text{1-particle}} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} |\mathbf{p}\rangle \langle \mathbf{p}|. \quad (7.1)$$

We can write an analogous completeness relation for the entire Hilbert space with the aid of a bit of notation. Let $|\lambda_p\rangle$ be the boost of $|\lambda_0\rangle$ with momentum \mathbf{p} , and assume that the states $|\lambda_p\rangle$, like the one-particle states $|\mathbf{p}\rangle$, are relativistically normalized. Let $E_p(\lambda) \equiv \sqrt{|\mathbf{p}|^2 + m_\lambda^2}$, where m_λ is the “mass” of the states $|\lambda_p\rangle$, that is, the energy of the state $|\lambda_0\rangle$. Then the desired completeness relation is

$$1 = |\Omega\rangle \langle \Omega| + \sum_\lambda \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p(\lambda)} |\lambda_p\rangle \langle \lambda_p|, \quad (7.2)$$

where the sum runs over all zero-momentum states $|\lambda_0\rangle$.

We now insert this expansion between the operators in the two-point function. Assume for now that $x^0 > y^0$. Let us drop the uninteresting constant term $\langle \Omega | \phi(x) | \Omega \rangle \langle \Omega | \phi(y) | \Omega \rangle$. (This term is usually zero by symmetry; for higher-spin fields, it is zero by Lorentz invariance.) The two-point function is then

$$\langle \Omega | \phi(x)\phi(y) | \Omega \rangle = \sum_\lambda \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p(\lambda)} \langle \Omega | \phi(x) | \lambda_p \rangle \langle \lambda_p | \phi(y) | \Omega \rangle. \quad (7.3)$$

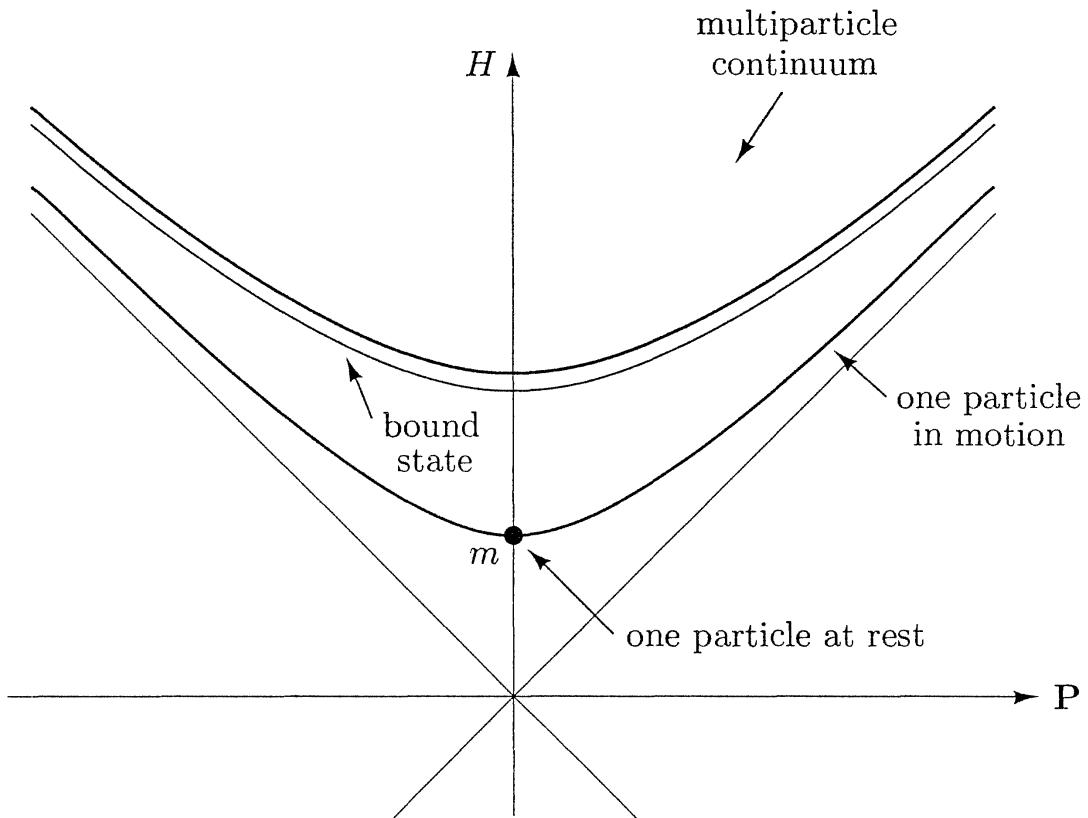


Figure 7.1. The eigenvalues of the 4-momentum operator $P^\mu = (H, \mathbf{P})$ occupy a set of hyperboloids in energy-momentum space. For a typical theory the states consist of one or more particles of mass m . Thus there is a hyperboloid of one-particle states and a continuum of hyperboloids of two-particle states, three-particle states, and so on. There may also be one or more bound-state hyperboloids below the threshold for creation of two free particles.

We can manipulate the matrix elements as follows:

$$\begin{aligned}
 \langle \Omega | \phi(x) | \lambda_p \rangle &= \langle \Omega | e^{iP \cdot x} \phi(0) e^{-iP \cdot x} | \lambda_p \rangle \\
 &= \langle \Omega | \phi(0) | \lambda_p \rangle e^{-ip \cdot x} \Big|_{p^0=E_p} \\
 &= \langle \Omega | \phi(0) | \lambda_0 \rangle e^{-ip \cdot x} \Big|_{p^0=E_p}.
 \end{aligned} \tag{7.4}$$

The last equality is a result of the Lorentz invariance of $\langle \Omega |$ and $\phi(0)$: Insert factors of $U^{-1}U$, where U is the unitary operator that implements a Lorentz boost from \mathbf{p} to 0, and use $U\phi(0)U^{-1} = \phi(0)$. (For a field with spin, we would need to keep track of its nontrivial Lorentz transformation.) Introducing an integration over p^0 , our expression for the two-point function (still for $x^0 > y^0$) becomes

$$\langle \Omega | \phi(x) \phi(y) | \Omega \rangle = \sum_\lambda \int \frac{d^4 p}{(2\pi)^4} \frac{i}{p^2 - m_\lambda^2 + i\epsilon} e^{-ip \cdot (x-y)} |\langle \Omega | \phi(0) | \lambda_0 \rangle|^2. \tag{7.5}$$

Note the appearance of the Feynman propagator, $D_F(x - y)$, but with m replaced by m_λ .

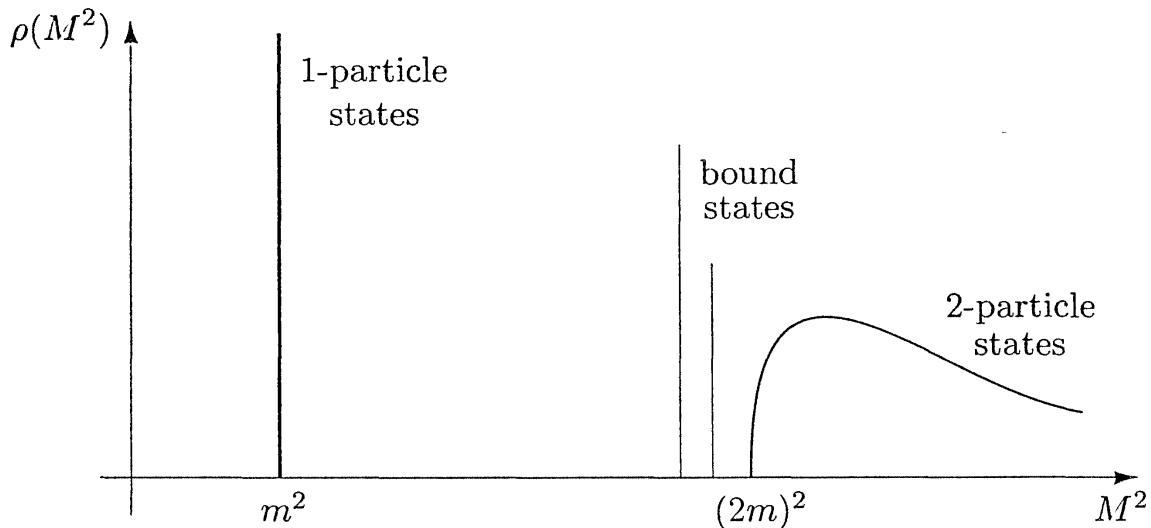


Figure 7.2. The spectral function $\rho(M^2)$ for a typical interacting field theory. The one-particle states contribute a delta function at m^2 (the square of the particle's mass). Multiparticle states have a continuous spectrum beginning at $(2m)^2$. There may also be bound states.

Analogous expressions hold for the case $y^0 > x^0$. Both cases can be summarized in the following general representation of the two-point function (the *Källén-Lehmann spectral representation*):

$$\langle \Omega | T\phi(x)\phi(y) | \Omega \rangle = \int_0^\infty \frac{dM^2}{2\pi} \rho(M^2) D_F(x - y; M^2), \quad (7.6)$$

where $\rho(M^2)$ is a positive *spectral density* function,

$$\rho(M^2) = \sum_\lambda (2\pi) \delta(M^2 - m_\lambda^2) |\langle \Omega | \phi(0) | \lambda_0 \rangle|^2. \quad (7.7)$$

The spectral density $\rho(M^2)$ for a typical theory is plotted in Fig. 7.2. Note that the one-particle states contribute an isolated delta function to the spectral density:

$$\rho(M^2) = 2\pi \delta(M^2 - m^2) \cdot Z + (\text{nothing else until } M^2 \gtrsim (2m)^2), \quad (7.8)$$

where Z is some number given by the squared matrix element in (7.7). We refer to Z as the *field-strength renormalization*. The quantity m is the exact mass of a single particle—the exact energy eigenvalue at rest. This quantity will in general differ from the value of the mass parameter that appears in the Lagrangian. We will refer to the parameter in the Lagrangian as m_0 , the *bare* mass, and refer to m as the *physical* mass of the ϕ boson. Only the physical mass m is directly observable.

The spectral decomposition (7.6) yields the following form for the Fourier

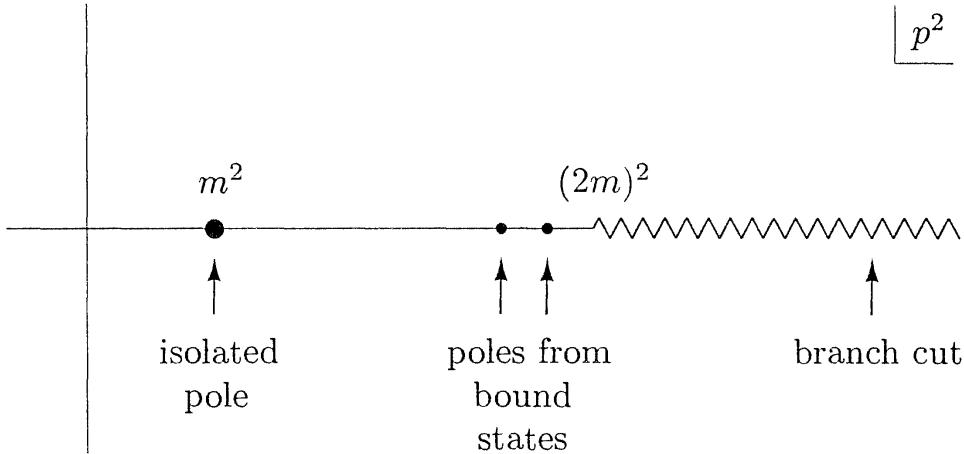


Figure 7.3. Analytic structure in the complex p^2 -plane of the Fourier transform of the two-point function for a typical theory. The one-particle states contribute an isolated pole at the square of the particle mass. States of two or more free particles give a branch cut, while bound states give additional poles.

transform of the two-point function:

$$\begin{aligned}
 \int d^4x e^{ip \cdot x} \langle \Omega | T\phi(x)\phi(0) | \Omega \rangle &= \int_0^\infty \frac{dM^2}{2\pi} \rho(M^2) \frac{i}{p^2 - M^2 + i\epsilon} \\
 &= \frac{iZ}{p^2 - m^2 + i\epsilon} + \int_{\sim 4m^2}^\infty \frac{dM^2}{2\pi} \rho(M^2) \frac{i}{p^2 - M^2 + i\epsilon}.
 \end{aligned} \tag{7.9}$$

The analytic structure of this function in the complex p^2 -plane is shown in Fig. 7.3. The first term gives an isolated simple pole at $p^2 = m^2$, while the second term contributes a branch cut beginning at $p^2 = (2m)^2$. If there are any two-particle bound states, these will appear as additional delta functions in $\rho(M^2)$ and thus as additional poles below the cut.

In Section 2.4, we found an explicit result for the two-point correlation function in the theory of a free scalar field:

$$\int d^4x e^{ip \cdot x} \langle 0 | T\phi(x)\phi(0) | 0 \rangle = \frac{i}{p^2 - m^2 + i\epsilon}. \tag{7.10}$$

We interpreted this formula, for $x^0 > 0$, as the amplitude for a particle to propagate from 0 to x . Equation (7.9) shows that the two-point function takes a similar form in the most general theory of an interacting scalar field. The general expression is essentially a sum of scalar propagation amplitudes for states created from the vacuum by the field operator $\phi(0)$. There are two differences between (7.9) and (7.10). First, Eq. (7.9) contains the field-strength renormalization factor $Z = |\langle \lambda_0 | \phi(0) | \Omega \rangle|^2$, the probability for $\phi(0)$ to create a given state from the vacuum. In (7.10), this factor is included implicitly, since $\langle p | \phi(0) | 0 \rangle = 1$ in free field theory. Second, Eq. (7.9) contains

contributions from multiparticle intermediate states with a continuous mass spectrum. In free field theory, $\phi(0)$ can create only a single particle from the vacuum. With these two differences, (7.9) is a direct generalization of (7.10).

It will be important in our later analysis that the contributions to (7.9) from one-particle and multiparticle intermediate states can be distinguished by the strength of their analytic singularities. The poles in p^2 come only from one-particle intermediate states, while multiparticle intermediate states give weaker branch cut singularities. We will see in the next section that this rather formal observation generalizes to higher-point correlation functions and plays a crucial role in our derivation of the diagrammatic formula for S -matrix elements.

The analysis of this section generalizes directly to two-point functions of higher-spin fields. The main complication comes in the generalization of the manipulation (7.4), since now the field has a nontrivial transformation law under boosts. In general, several invariant spectral functions are required to represent the multiparticle states. But this complication does not affect the major result that a pole in p^2 can arise only from the contribution of a single-particle state created by the field operator. The two-point function of Dirac fields, for example, has the structure

$$\begin{aligned} \int d^4x e^{ip \cdot x} \langle \Omega | T\psi(x)\bar{\psi}(0) | \Omega \rangle &= \frac{iZ_2 \sum_s u^s(p)\bar{u}^s(p)}{p^2 - m^2 + i\epsilon} + \dots \\ &= \frac{iZ_2(\not{p} + m)}{p^2 - m^2 + i\epsilon} + \dots, \end{aligned} \quad (7.11)$$

where the omitted terms give the multiparticle branch cut. As in the scalar case, the constant Z_2 is the probability for the quantum field to create or annihilate an exact one-particle eigenstate of H :

$$\langle \Omega | \psi(0) | p, s \rangle = \sqrt{Z_2} u^s(p). \quad (7.12)$$

(For an antiparticle, replace u with \bar{v} .) At the pole, the Dirac two-point function is exactly that of a free field with the physical mass, aside from the rescaling factor Z_2 .

An Example: The Electron Self-Energy

This nonperturbative analysis of the two-point correlation function has been very different from our usual direct analysis of Feynman diagrams. But since this derivation was done in complete generality, the singularity structure of the two-point function that it implies ought also to be visible in a Feynman diagram computation. In the rest of this section we will explicitly check our results for the electron two-point function in QED.

The electron two-point function is equal to the sum of diagrams

$$\langle \Omega | T\psi(x)\bar{\psi}(y) | \Omega \rangle = \text{---} \xleftarrow{x} \text{---} \xrightarrow{y} + \text{---} \xleftarrow{x} \text{---} \xrightarrow{y} \text{---} \text{---} + \dots \quad (7.13)$$

Each of these diagrams, according to the Feynman rules for correlation functions, contains a factor of $e^{ip \cdot (x-y)}$ for the two external points and an integration $\int (d^2 p / (2\pi)^4)$ over the momentum p carried by the initial and final propagators. We will consistently omit these factors in this section; in other words, each diagram will denote the corresponding term in the Fourier transform of the two-point function.

The first diagram is just the free-field propagator:

$$\begin{array}{c} \text{---} \\ \text{---} \\ p \end{array} = \frac{i(\not{p} + m_0)}{p^2 - m_0^2 + i\epsilon}. \quad (7.14)$$

Throughout this calculation, we will write m_0 instead of m as the mass in the electron propagator. This makes explicit the fact noted above that the mass appearing in the Lagrangian differs, in general, from the observable rest energy of a particle. However, if a perturbation expansion is applicable, the leading-order expression for the propagator should approximate the exact expression. Indeed, the function (7.14) has a pole, of just the form of (7.11), at $p^2 = m_0^2$. We therefore expect that the complete expression for the two-point function also has a pole of this form, at a slightly shifted location $m^2 = m_0^2 + \mathcal{O}(\alpha)$.

The second diagram in (7.13), called the *electron self-energy*, is somewhat more complicated:

$$\begin{array}{c} \text{---} \\ \text{---} \\ p \\ \text{---} \\ \text{---} \\ k \\ \text{---} \\ p \end{array} = \frac{i(\not{p} + m_0)}{p^2 - m_0^2} [-i\Sigma_2(p)] \frac{i(\not{p} + m_0)}{p^2 - m_0^2}, \quad (7.15)$$

where

$$-i\Sigma_2(p) = (-ie)^2 \int \frac{d^4 k}{(2\pi)^4} \gamma^\mu \frac{i(\not{k} + m_0)}{k^2 - m_0^2 + i\epsilon} \gamma_\mu \frac{-i}{(p-k)^2 - \mu^2 + i\epsilon}. \quad (7.16)$$

(The notation Σ_2 indicates that this is the second-order (in e) contribution to a quantity Σ that we will define below.) The integral Σ_2 has an infrared divergence, which we have regularized by adding a small photon mass μ . Outside this integral, the diagram seems to have a double pole at $p^2 = m_0^2$. All in all, the form of this correction is quite unpleasant. But let us press on and try to evaluate $\Sigma_2(p)$ using the calculational techniques developed for the vertex correction in the Section 6.3.

First introduce a Feynman parameter to combine the two denominators:

$$\frac{1}{k^2 - m_0^2 + i\epsilon} \frac{1}{(p-k)^2 - \mu^2 + i\epsilon} = \int_0^1 dx \frac{1}{[k^2 - 2xk \cdot p + xp^2 - x\mu^2 - (1-x)m_0^2 + i\epsilon]^2}.$$

Now complete the square and define a shifted momentum $\ell \equiv k - xp$. Dropping the term linear in ℓ from the numerator, we have

$$-i\Sigma_2(p) = -e^2 \int_0^1 dx \int \frac{d^4\ell}{(2\pi)^4} \frac{-2xp + 4m_0}{[\ell^2 - \Delta + i\epsilon]^2}, \quad (7.17)$$

where $\Delta = -x(1-x)p^2 + x\mu^2 + (1-x)m_0^2$. The integral over ℓ is divergent. To evaluate it, we first regulate it using the Pauli-Villars procedure (6.51):

$$\frac{1}{(p-k)^2 - \mu^2 + i\epsilon} \rightarrow \frac{1}{(p-k)^2 - \mu^2 + i\epsilon} - \frac{1}{(p-k)^2 - \Lambda^2 + i\epsilon}.$$

The second term will have the same form as (7.17), but with μ replaced by Λ . As in Section 6.3, we now Wick-rotate and substitute the Euclidean variable $\ell_E^0 = -i\ell^0$. This gives

$$\begin{aligned} \int \frac{d^4\ell}{(2\pi)^4} \frac{1}{[\ell^2 - \Delta]^2} &\rightarrow \frac{i}{(4\pi)^2} \int_0^\infty d\ell_E^2 \left(\frac{\ell_E^2}{[\ell_E^2 + \Delta]^2} - \frac{\ell_E^2}{[\ell_E^2 + \Delta_\Lambda]^2} \right) \\ &= \frac{i}{(4\pi)^2} \log\left(\frac{\Delta_\Lambda}{\Delta}\right), \end{aligned} \quad (7.18)$$

where

$$\Delta_\Lambda = -x(1-x)p^2 + x\Lambda^2 + (1-x)m_0^2 \xrightarrow{\Lambda \rightarrow \infty} x\Lambda^2.$$

The final result is therefore

$$\Sigma_2(p) = \frac{\alpha}{2\pi} \int_0^1 dx (2m_0 - xp) \log\left(\frac{x\Lambda^2}{(1-x)m_0^2 + x\mu^2 - x(1-x)p^2}\right). \quad (7.19)$$

Before discussing the divergences in this expression, let us work out its analytic behavior as a function of p^2 . The logarithm in (7.19) has a branch cut when its argument becomes negative, and for any fixed x this will occur for sufficiently large p^2 . More exactly, the cut begins at the point where

$$(1-x)m_0^2 + x\mu^2 - x(1-x)p^2 = 0.$$

Solving this equation for x , we find

$$\begin{aligned} x &= \frac{1}{2} + \frac{m_0^2}{2p^2} - \frac{\mu^2}{2p^2} \pm \sqrt{\frac{(p^2 + m_0^2 - \mu^2)^2}{4p^4} - \frac{m_0^2}{p^2}} \\ &= \frac{1}{2} + \frac{m_0^2}{2p^2} - \frac{\mu^2}{2p^2} \pm \frac{1}{2p^2} \sqrt{[p^2 - (m_0 + \mu)^2][p^2 - (m_0 - \mu)^2]}. \end{aligned} \quad (7.20)$$

The branch cut of $\Sigma_2(p^2)$ begins at the minimum value of p^2 such that this equation has a real solution for x between 0 and 1. This occurs when $p^2 = (m_0 + \mu)^2$, that is, at the threshold for creation of a two-particle (electron

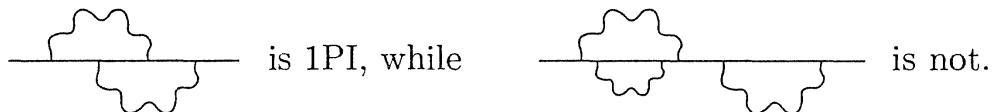
plus photon) state. In fact, it is a simple exercise in relativistic kinematics to show that the square root in (7.20), written in the form

$$k = \frac{1}{\sqrt{p^2}} \sqrt{[p^2 - (m_0 + \mu)^2][p^2 - (m_0 - \mu)^2]},$$

is precisely the momentum in the center-of-mass frame for two particles of mass m_0 and μ and total energy $\sqrt{p^2}$. It is natural that this momentum becomes real at the two-particle threshold. The location of the branch cut is exactly where we would expect from the Källén-Lehmann spectral representation.[†]

We have now located the two-particle branch cut predicted by the Källén-Lehmann representation, but we have not found the expected simple pole at $p^2 = m^2$. To find it we must actually include an infinite series of Feynman diagrams. Fortunately, this series will be easily summed.

Let us define a *one-particle irreducible* (1PI) diagram to be any diagram that cannot be split in two by removing a single line:



Let $-i\Sigma(p)$ denote the sum of all 1PI diagrams with two external fermion lines:

$$\begin{aligned} -i\Sigma(p) &= \text{---} \xleftarrow{\text{1PI}} \text{---} \\ &= \text{---} \xleftarrow{\text{1PI}} \text{---} + \text{---} \xleftarrow{\text{1PI}} \text{---} + \text{---} \xleftarrow{\text{1PI}} \text{---} + \dots \end{aligned} \quad (7.21)$$

(Although each diagram has two external lines, the Feynman propagators for these lines are not to be included in the expression for $\Sigma(p)$.) To leading order in α we see that $\Sigma = \Sigma_2$.

The Fourier transform of the two-point function can now be written as

$$\begin{aligned} \int d^4x \langle \Omega | T\psi(x)\bar{\psi}(0) | \Omega \rangle e^{ip \cdot x} &= \text{---} \xleftarrow{\text{1PI}} \text{---} \\ &= \text{---} \text{---} + \text{---} \xleftarrow{\text{1PI}} \text{---} + \text{---} \xleftarrow{\text{1PI}} \text{---} \xleftarrow{\text{1PI}} \text{---} + \dots \\ &= \frac{i(p + m_0)}{p^2 - m_0^2} + \frac{i(p + m_0)}{p^2 - m_0^2} (-i\Sigma) \frac{i(p + m_0)}{p^2 - m_0^2} + \dots \end{aligned} \quad (7.22)$$

[†]In real QED, $\mu = 0$ and the two-particle branch cut merges with the one-particle pole. This subtlety plays a role in the full treatment of the cancellation of infrared divergences, but it is beyond the scope of our present analysis.

The first diagram has a simple pole at $p^2 = m_0^2$. Each diagram in the second class has a double pole at $p^2 = m_0^2$. Each diagram in the third class has a triple pole. The behavior near $p^2 = m_0^2$ gets worse and worse as we include more and more diagrams. But fortunately, the sum of all the diagrams forms a geometric series. Note that $\Sigma(p)$ commutes with \not{p} , since $\Sigma(p)$ is a function only of pure numbers and \not{p} . In fact, we can consider $\Sigma(p)$ to be a function of \not{p} , writing $p^2 = (\not{p})^2$. Then we can rewrite each electron propagator as $i/(\not{p} - m_0)$ and express the above series as

$$\begin{aligned} & \int d^4x \langle \Omega | T\psi(x)\bar{\psi}(0) | \Omega \rangle e^{ip \cdot x} \\ &= \frac{i}{\not{p} - m_0} + \frac{i}{\not{p} - m_0} \left(\frac{\Sigma(\not{p})}{\not{p} - m_0} \right) + \frac{i}{\not{p} - m_0} \left(\frac{\Sigma(\not{p})}{\not{p} - m_0} \right)^2 + \dots \\ &= \frac{i}{\not{p} - m_0 - \Sigma(\not{p})}. \end{aligned} \quad (7.23)$$

The full propagator has a simple pole, which is shifted away from m_0 by $\Sigma(\not{p})$.

The location of this pole, the physical mass m , is the solution of the equation

$$[\not{p} - m_0 - \Sigma(\not{p})] \Big|_{\not{p}=m} = 0. \quad (7.24)$$

Notice that, if $\Sigma(\not{p})$ is defined by the convention (7.21), then a positive contribution to Σ yields a positive shift of the electron mass. Close to the pole, the denominator of (7.23) has the form

$$(\not{p} - m) \cdot \left(1 - \frac{d\Sigma}{d\not{p}} \Big|_{\not{p}=m} \right) + \mathcal{O}((\not{p} - m)^2). \quad (7.25)$$

Thus the full electron propagator has a single-particle pole of just the form (7.11), with m given by (7.24) and

$$Z_2^{-1} = 1 - \frac{d\Sigma}{d\not{p}} \Big|_{\not{p}=m}. \quad (7.26)$$

Our explicit calculation of Σ_2 allows us to compute the first corrections to m and Z_2 . Let us begin with m . To order α , the mass shift is

$$\delta m = m - m_0 = \Sigma_2(\not{p} = m) \approx \Sigma_2(\not{p} = m_0). \quad (7.27)$$

Thus, using (7.19),

$$\delta m = \frac{\alpha}{2\pi} m_0 \int_0^1 dx (2-x) \log \left(\frac{\Lambda^2}{(1-x)^2 m_0^2 + x\mu^2} \right). \quad (7.28)$$

The mass shift is ultraviolet divergent; the divergent term has the form

$$\delta m \xrightarrow[\Lambda \rightarrow \infty]{3\alpha}{4\pi} m_0 \log \left(\frac{\Lambda^2}{m_0^2} \right). \quad (7.29)$$

Is it a problem that m differs from m_0 by a divergent quantity? This question has two levels, those of concept and practice.

On the conceptual level, we should fully expect the mass of the electron to be modified by its coupling to the electromagnetic field. In classical electrodynamics, the rest energy of any charge is increased by the energy of its electrostatic field, and this energy shift diverges in the case of a point charge:

$$\int d^3r \frac{1}{2}|\mathbf{E}|^2 = \int d^3r \frac{1}{2} \left(\frac{e}{4\pi r^2} \right)^2 = \frac{\alpha}{2} \int \frac{dr}{r^2} \sim \alpha \Lambda. \quad (7.30)$$

In fact, it is puzzling why the divergence in (7.29) is so weak, logarithmic in Λ rather than linear as in (7.30). To understand this feature, suppose that m_0 were set to 0. Then the two helicity components of the electron field ψ_L and ψ_R would not be coupled by any term in the QED Hamiltonian. This would imply that perturbative corrections could never induce a coupling of ψ_L and ψ_R , nor, in particular, an electron mass term. In other words, δm must vanish when $m_0 = 0$. The mass shift must therefore be proportional to m_0 , and so, by dimensional analysis, it can depend only logarithmically on Λ .

On a practical level, the infinite mass shift casts doubt on our perturbative calculations. For example, all of the theoretical results in Chapter 5 should technically involve m_0 rather than m . To compare theory to experiment we must eliminate m_0 in favor of m , using the relation $m_0 = m + \mathcal{O}(\alpha)$. Since the “small” $\mathcal{O}(\alpha)$ correction is infinite, the validity of this procedure is far from obvious. The validity of perturbation theory would be more plausible if we could compute Feynman diagrams using the propagator $i/(\not{p} - m)$, which has the correct pole location, instead of $i/(\not{p} - m_0)$. In Chapter 10 we will see how to rearrange the perturbation series in such a way that m_0 is systematically eliminated in favor of m and the zeroth-order propagator has its pole at the physical mass. In the remainder of this chapter, we will continue to simply replace m_0 by m in expressions for order- α corrections.

Finally, let us examine the perturbative correction to Z_2 . From (7.26), we find that the order- α correction $\delta Z_2 = (Z_2 - 1)$ is

$$\begin{aligned} \delta Z_2 &= \frac{d\Sigma_2}{d\not{p}} \bigg|_{\not{p}=m} \\ &= \frac{\alpha}{2\pi} \int_0^1 dx \left[-x \log \frac{x\Lambda^2}{(1-x)^2 m^2 + x\mu^2} + 2(2-x) \frac{x(1-x)m^2}{(1-x)^2 m^2 + x\mu^2} \right]. \end{aligned} \quad (7.31)$$

This expression is also logarithmically ultraviolet divergent. We will discuss the observability of this divergent term at the end of Section 7.2. However, it is interesting to note, even before that discussion, that (7.31) is very similar in form to the value of the *ad hoc* subtraction that we made in our calculation of the electron vertex correction in Section 6.3. From Eq. (6.56), the value of

this subtraction was

$$\begin{aligned}
 \delta F_1(0) &= \frac{\alpha}{2\pi} \int_0^1 dx dy dz \delta(x+y+z-1) \\
 &\quad \times \left[\log\left(\frac{z\Lambda^2}{(1-z)^2 m^2 + z\mu^2}\right) + \frac{(1-4z+z^2)m^2}{(1-z)^2 m^2 + z\mu^2} \right] \\
 &= \frac{\alpha}{2\pi} \int_0^1 dz (1-z) \left[\log\left(\frac{z\Lambda^2}{(1-z)^2 m^2 + z\mu^2}\right) + \frac{(1-4z+z^2)m^2}{(1-z)^2 m^2 + z\mu^2} \right]. \quad (7.32)
 \end{aligned}$$

Using the integration by parts

$$\begin{aligned}
 \int_0^1 dz (1-2z) \log\left(\frac{\Lambda^2}{(1-z)^2 m^2 + z\mu^2}\right) &= - \int_0^1 dz z(1-z) \frac{2(1-z)m^2}{(1-z)^2 m^2 - \mu^2} \\
 &= \int_0^1 dz \left[(1-z) - \frac{(1-z)(1-z^2)m^2}{(1-z)^2 m^2 - \mu^2} \right],
 \end{aligned}$$

it is not hard to show that $\delta F_1(0) + \delta Z_2 = 0$. This identity will play a crucial role in justifying the *ad hoc* subtraction of Section 6.3.

7.2 The LSZ Reduction Formula

In the last section we saw that the Fourier transform of the two-point correlation function, considered as an analytic function of p^2 , has a simple pole at the mass of the one-particle state:

$$\int d^4x e^{ip \cdot x} \langle \Omega | T\phi(x)\phi(0) | \Omega \rangle \underset{p^2 \rightarrow m^2}{\sim} \frac{iZ}{p^2 - m^2 + i\epsilon}. \quad (7.33)$$

(Here and throughout this section we use the symbol \sim to mean that the poles of both sides are identical; there are additional finite terms, given in this case by Eq. (7.9).) In this section we will generalize this result to higher correlation functions. We will derive a general relation between correlation functions and S -matrix elements first obtained by Lehmann, Symanzik, and Zimmermann and known as the *LSZ reduction formula*.[†] This result, combined with our Feynman rules for computing correlation functions, will justify Eq. (4.103), our master formula for S -matrix elements in terms of Feynman diagrams. For simplicity, we will carry out the whole analysis for the case of scalar fields.

[†]H. Lehmann, K. Symanzik, and W. Zimmermann, *Nuovo Cimento* **1**, 1425 (1955).

The strategy of the argument will be as follows. To calculate the S -matrix element for a $2\text{-body} \rightarrow n\text{-body}$ scattering process, we begin with the correlation function of $n+2$ Heisenberg fields. Fourier-transforming with respect to the coordinate of any one of these fields, we will find a pole of the form (7.33) in the Fourier-transform variable p^2 . We will argue that the one-particle states associated with these poles are in fact asymptotic states, that is, states given by the limit of well-separated wavepackets as they become concentrated around definite momenta. Taking the limit in which all $n+2$ external particles go on-shell, we can then interpret the coefficient of the multiple pole as an S -matrix element.

To begin, let us Fourier-transform the $(n+2)$ -point correlation function with respect to one argument x . We must then analyze the integral

$$\int d^4x e^{ip \cdot x} \langle \Omega | T\{\phi(x)\phi(z_1)\phi(z_2)\dots\} | \Omega \rangle.$$

We would like to identify poles in the variable p^0 . To do this, divide the integral over x^0 into three regions:

$$\int dx^0 = \int_{T_+}^{\infty} dx^0 + \int_{T_-}^{T_+} dx^0 + \int_{-\infty}^{T_-} dx^0, \quad (7.34)$$

where T_+ is much greater than all the z_i^0 and T_- is much less than all the z_i^0 . Call these three intervals regions I, II, and III. Since region II is bounded and the integrand depends on p^0 through the analytic function $\exp(ip^0 x^0)$, the contribution from this region will be analytic in p^0 . However, regions I and III, which are unbounded, may develop singularities in p^0 .

Consider first region I. Here x^0 is the latest time, so $\phi(x)$ stands first in the time ordering. Insert a complete set of intermediate states in the form of (7.2):

$$1 = \sum_{\lambda} \int \frac{d^3q}{(2\pi)^3} \frac{1}{2E_{\mathbf{q}}(\lambda)} |\lambda_{\mathbf{q}}\rangle \langle \lambda_{\mathbf{q}}|.$$

The integral over region I then becomes

$$\begin{aligned} \int_{T_+}^{\infty} dx^0 \int d^3x e^{ip^0 x^0} e^{-i\mathbf{p} \cdot \mathbf{x}} \sum_{\lambda} \int \frac{d^3q}{(2\pi)^3} \frac{1}{2E_{\mathbf{q}}(\lambda)} \langle \Omega | \phi(x) | \lambda_{\mathbf{q}} \rangle \\ \times \langle \lambda_{\mathbf{q}} | T\{\phi(z_1)\phi(z_2)\dots\} | \Omega \rangle. \end{aligned} \quad (7.35)$$

Using Eq. (7.4),

$$\langle \Omega | \phi(x) | \lambda_{\mathbf{q}} \rangle = \langle \Omega | \phi(0) | \lambda_0 \rangle e^{-iq \cdot x} \Big|_{q^0 = E_{\mathbf{q}}(\lambda)},$$

and including a factor $e^{-\epsilon x^0}$ to insure that the integral is well defined, this integral becomes

$$\begin{aligned} & \sum_{\lambda} \int_{T_+}^{\infty} dx^0 \int \frac{d^3 q}{(2\pi)^3} \frac{1}{2E_{\mathbf{q}}(\lambda)} e^{ip^0 x^0} e^{-iq^0 x^0} e^{-\epsilon x^0} \langle \Omega | \phi(0) | \lambda_0 \rangle (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \\ & \quad \times \langle \lambda_{\mathbf{q}} | T\{\phi(z_1) \dots\} | \Omega \rangle \\ & = \sum_{\lambda} \frac{1}{2E_{\mathbf{p}}(\lambda)} \frac{i}{p^0 - E_{\mathbf{p}}(\lambda) + i\epsilon} \langle \Omega | \phi(0) | \lambda_0 \rangle \langle \lambda_{\mathbf{p}} | T\{\phi(z_1) \dots\} | \Omega \rangle. \end{aligned} \quad (7.36)$$

The denominator is just that of Eq. (7.5): $p^2 - m_{\lambda}^2$. There is an analytic singularity in p^0 ; as in Section 7.1, this singularity will be either a pole or a branch cut depending upon whether or not the rest energy m_{λ} is isolated. The one-particle state corresponds to an isolated energy value $p^0 = E_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}$, and at this point Eq. (7.36) has a pole:

$$\begin{aligned} & \int d^4 x e^{ip \cdot x} \langle \Omega | T\{\phi(x) \phi(z_1) \dots\} | \Omega \rangle \\ & \underset{p^0 \rightarrow +E_{\mathbf{p}}}{\sim} \frac{i}{p^2 - m^2 + i\epsilon} \sqrt{Z} \langle \mathbf{p} | T\{\phi(z_1) \dots\} | \Omega \rangle. \end{aligned} \quad (7.37)$$

The factor \sqrt{Z} is the same field strength renormalization factor as in Eq. (7.8), since it replaces the same matrix element as in (7.7).

To evaluate the contribution from region III, we would put $\phi(x)$ last in the operator ordering, then insert a complete set of states between $T\{\phi(z_1) \dots\}$ and $\phi(x)$. Repeating the above argument produces a pole as $p^0 \rightarrow -E_{\mathbf{p}}$:

$$\begin{aligned} & \int d^4 x e^{ip \cdot x} \langle \Omega | T\{\phi(x) \phi(z_1) \dots\} | \Omega \rangle \\ & \underset{p^0 \rightarrow -E_{\mathbf{p}}}{\sim} \langle \Omega | T\{\phi(z_1) \dots\} | -\mathbf{p} \rangle \sqrt{Z} \frac{i}{p^2 - m^2 + i\epsilon}. \end{aligned} \quad (7.38)$$

Next we would like to Fourier-transform with respect to the remaining field coordinates. To keep the various external particles from interfering, however, we must isolate them from each other in space. Let us therefore repeat the preceding calculation using a wavepacket rather than a simple Fourier transform. In Eq. (7.35), replace

$$\int d^4 x e^{ip^0 x^0} e^{-ip \cdot x} \rightarrow \int \frac{d^3 k}{(2\pi)^3} \int d^4 x e^{ip^0 x^0} e^{-ik \cdot x} \varphi(\mathbf{k}), \quad (7.39)$$

where $\varphi(\mathbf{k})$ is a narrow distribution centered on $\mathbf{k} = \mathbf{p}$. This distribution constrains x to lie within a band, whose spatial extent is that of the wavepacket, about the trajectory of a particle with momentum \mathbf{p} . With this modification, the right-hand side of (7.36) has a more complicated singularity structure:

$$\sum_{\lambda} \int \frac{d^3 k}{(2\pi)^3} \varphi(\mathbf{k}) \frac{1}{2E_{\mathbf{k}}(\lambda)} \frac{i}{p^0 - E_{\mathbf{k}}(\lambda) + i\epsilon} \langle \Omega | \phi(0) | \lambda_0 \rangle \langle \lambda_{\mathbf{k}} | T\{\phi(z_1) \dots\} | \Omega \rangle$$

$$p^0 \xrightarrow{+E_p} \int \frac{d^3 k}{(2\pi)^3} \varphi(\mathbf{k}) \frac{i}{\tilde{p}^2 - m^2 + i\epsilon} \sqrt{Z} \langle \mathbf{k} | T\{\phi(z_1) \dots\} | \Omega \rangle, \quad (7.40)$$

where, in the second line, $\tilde{p} = (p_0, \mathbf{k})$. The one-particle singularity is now a branch cut, whose length is the width in momentum space of the wavepacket $\varphi(\mathbf{k})$. However, if $\varphi(\mathbf{k})$ defines the momentum narrowly, this branch cut is very short, and (7.40) has a well-defined limit in which $\varphi(\mathbf{k})$ tends to $(2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{p})$ and the singularity of (7.40) sharpens up to the pole of (7.36). The singularity due to single-particle states in the far past, Eq. (7.38), is modified in the same way.

Now consider integrating each of the coordinates in the $(n + 2)$ -point correlation function against a wavepacket, to form*

$$\left(\prod_i \int \frac{d^3 k_i}{(2\pi)^3} \int d^4 x_i e^{i\tilde{p}_i \cdot x_i} \varphi_i(\mathbf{k}_i) \right) \langle \Omega | T\{\phi(x_1) \phi(x_2) \dots\} | \Omega \rangle. \quad (7.41)$$

The wavepackets should be chosen to overlap in a region around $x = 0$ and to separate in the far past and the far future. To analyze this integral, we choose a large positive time T_+ such that all of the wavepackets are well separated for $x_i^0 > T_+$, and we choose a large negative time T_- such that all of the wavepackets are well separated for $x_i^0 < T_-$. Then we can break up each of the integrals over x_i^0 into three regions as in (7.34). The integral of any x_i^0 over the bounded region II leads to an expression analytic in the corresponding energy p_0^i , so we can concentrate on the case in which all of the x_i^0 are placed at large past or future times.

For definiteness, consider the contribution in which only two of the time coordinates, x_1^0 and x_2^0 , are in the future. In this case, the fields $\phi(x_1)$ and $\phi(x_2)$ stand to the left of the other fields in time order. Inserting a complete set of states $|\lambda_K\rangle$, the integrations in (7.41) over the coordinates of these two fields take the form

$$\begin{aligned} \sum_{\lambda} \int \frac{d^3 K}{(2\pi)^3} \frac{1}{2E_K} & \left(\prod_{i=1,2} \int \frac{d^3 k_i}{(2\pi)^3} \int d^4 x_i e^{i\tilde{p}_i \cdot x_i} \varphi_i(\mathbf{k}_i) \right) \\ & \times \langle \Omega | T\{\phi(x_1) \phi(x_2)\} | \lambda_K \rangle \langle \lambda_K | T\{\phi(x_3) \dots\} | \Omega \rangle. \end{aligned}$$

The state $|\lambda_K\rangle$ is annihilated by two field operators constrained to lie in distant wavepackets. It must therefore consist of two distinct excitations of the vacuum at two distinct locations. If these excitations are well separated,

*As in Section 4.5, the product symbol applies symbolically to the integrations as well as to the other factors within the parentheses; the x_i integrals apply to what lies outside the parentheses as well.

they should be independent of one another, so we can approximate

$$\begin{aligned} & \sum_{\lambda} \int \frac{d^3 K}{(2\pi)^3} \frac{1}{2E_{\mathbf{K}}} \langle \Omega | T\{\phi(x_1)\phi(x_2)\} | \lambda_{\mathbf{K}} \rangle \langle \lambda_{\mathbf{K}} | \\ &= \sum_{\lambda_1, \lambda_2} \int \frac{d^3 q_1}{(2\pi)^3} \frac{1}{2E_{\mathbf{q}_1}} \int \frac{d^3 q_2}{(2\pi)^3} \frac{1}{2E_{\mathbf{q}_2}} \langle \Omega | \phi(x_1) | \lambda_{\mathbf{q}_1} \rangle \langle \Omega | \phi(x_2) | \lambda_{\mathbf{q}_2} \rangle \langle \lambda_{\mathbf{q}_1} \lambda_{\mathbf{q}_2} |. \end{aligned}$$

The sums over λ_1 and λ_2 in the this equation run over all zero-momentum states, but only single-particle states will contribute the poles we are looking for. In this case, the integrals over x_1^0 and \mathbf{q}_1 produce a sharp singularity in p_1^0 of the form of (7.40), and the integrals over x_2^0 and \mathbf{q}_2 produce the same singular behavior in p_2^0 . The term in (7.41) with both singularities is

$$\left(\prod_{i=1,2} \int \frac{d^3 k_i}{(2\pi)^3} \varphi_i(\mathbf{k}_i) \frac{i}{\tilde{p}_i^2 - m^2 + i\epsilon} \cdot \sqrt{Z} \right) \langle \mathbf{k}_1 \mathbf{k}_2 | T\{\phi(x_3) \dots\} | \Omega \rangle.$$

In the limit in which the wavepackets tend to delta functions concentrated at definite momenta \mathbf{p}_1 and \mathbf{p}_2 , this expression tends to

$$\left(\prod_{i=1,2} \frac{i}{p_i^2 - m^2 + i\epsilon} \cdot \sqrt{Z} \right)_{\text{out}} \langle \mathbf{p}_1 \mathbf{p}_2 | T\{\phi(x_3) \dots\} | \Omega \rangle.$$

The state $\langle \mathbf{p}_1 \mathbf{p}_2 |$ is precisely an *out* state as defined in Section 4.5, since it is the definite-momentum limit of a state of particles constrained to well-separated wavepackets. Applying the same analysis to the times x_i^0 in the far past gives the result that the coefficient of the maximally singular term in the corresponding p_i^0 is a matrix element with an *in* state. This most singular term in (7.41) thus has the form

$$\left(\prod_{i=1,2} \frac{i}{p_i^2 - m^2 + i\epsilon} \cdot \sqrt{Z} \right) \left(\prod_{i=3,\dots} \frac{i}{p_i^2 - m^2 + i\epsilon} \cdot \sqrt{Z} \right)_{\text{out}} \langle \mathbf{p}_1 \mathbf{p}_2 | -\mathbf{p}_3 \dots \rangle_{\text{in}}.$$

The last factor is just an *S*-matrix element.

We have now shown that we can extract the value of an *S*-matrix element by folding the corresponding vacuum expectation value of fields with wavepackets, extracting the leading singularities in the energies p_i^0 , and then taking the limit as these wavepackets become delta functions of momenta. However, the computation would be made much simpler if we could do these operations in the reverse order—first letting the wavepackets become delta functions, returning us to the simple Fourier transform, and then extracting the singularities. In fact, the result for the leading singularity is not changed by this switch of the order of operations. It is, however, rather subtle to argue this point. Roughly, the explanation is the following: In the language of the analysis just completed, new singularities might arise because, in the Fourier transform, x_1 and x_2 can become close together in the far future. However, in this region, the exponential factor is close to $\exp[i(p_1 + p_2) \cdot x_1]$, and thus the new singularities are single poles in the variable $(p_1^0 + p_2^0)$, rather than

being products of poles in the two separate energy variables. A more careful argument (unfortunately, couched in a rather different language) can be found in the original paper of Lehmann, Symanzik, and Zimmermann cited at the beginning of this section.

Given the ability to make this reversal in the order of operations, we obtain a precise relation between Fourier transforms of correlation functions and S -matrix elements. This is the LSZ reduction formula:

$$\prod_1^n \int d^4 x_i e^{ip_i \cdot x_i} \prod_1^m \int d^4 y_j e^{-ik_j \cdot y_j} \langle \Omega | T\{\phi(x_1) \cdots \phi(x_n) \phi(y_1) \cdots \phi(y_m)\} | \Omega \rangle$$

$$\stackrel{\text{each } p_i^0 \rightarrow +E_{p_i}}{\sim} \left(\prod_{i=1}^n \frac{\sqrt{Z} i}{p_i^2 - m^2 + i\epsilon} \right) \left(\prod_{j=1}^m \frac{\sqrt{Z} i}{k_j^2 - m^2 + i\epsilon} \right) \langle p_1 \cdots p_n | S | k_1 \cdots k_m \rangle. \quad (7.42)$$

The quantity Z that appears in this equation is exactly the field-strength renormalization constant, defined in Section 7.1 as the residue of the single-particle pole in the two-point function of fields. Each distinct particle will have a distinct renormalization factor Z , obtained from its own two-point function. For higher-spin fields, each factor of \sqrt{Z} comes with a polarization factor such as $u^s(p)$, as in Eq. (7.12). The polarization s must be summed over in the second line of (7.42).

In words, the LSZ formula says that an S -matrix element can be computed as follows. Compute the appropriate Fourier-transformed correlation function, look at the region of momentum space where the external particles are near the mass shell, and identify the coefficient of the multiparticle pole. For fields with spin, one must also multiply by a polarization spinor (like $u^s(p)$) or vector (like $\epsilon^r(k)$) to project out the desired spin state.

Our next goal is to express this procedure in the language of Feynman diagrams. Let us analyze the relation between the diagrammatic expansion of the scalar field four-point function and the S -matrix element for 2-particle \rightarrow 2-particle scattering. We will consider explicitly the fully connected Feynman diagrams contributing to the correlator. By a similar analysis, it is easy to confirm that disconnected diagrams should be disregarded because they do not have the singularity structure, with a product of four poles, indicated on the right-hand side of (7.42).

The exact four-point function

$$\left(\prod_1^2 \int d^4 x_i e^{ip_i \cdot x_i} \right) \left(\prod_1^2 \int d^4 y_i e^{-ik_i \cdot y_i} \right) \langle \Omega | T\{\phi(x_1) \phi(x_2) \phi(y_1) \phi(y_2)\} | \Omega \rangle$$

has the general form shown in Fig. 7.4. In this figure, we have indicated explicitly the diagrammatic corrections on each leg; the shaded circle in the center represents the sum of all amputated four-point diagrams.

We can sum up the corrections to each external leg just as we did for the electron propagator in the previous section. Let $-iM(p^2)$ denote the sum of

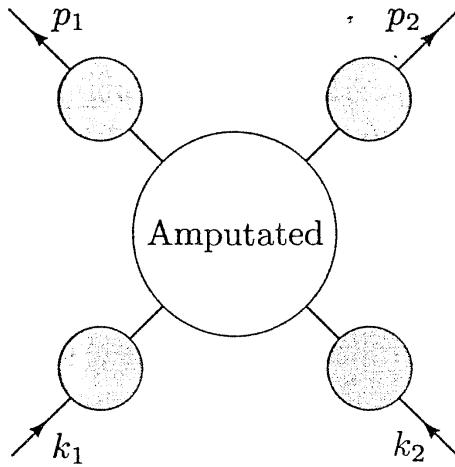


Figure 7.4. Structure of the exact four-point function in scalar field theory.

all one-particle-irreducible (1PI) insertions into the scalar propagator:

$$-iM(p^2) = \text{---} \textcircled{0} \text{---} + \text{---} \textcircled{8} \text{---} + \text{---} \textcircled{0} \text{---} + \dots = \text{---} \textcircled{1PI} \text{---}$$

Then the exact propagator can be written as a geometric series and summed as in Eq. (7.23):

$$\begin{aligned}
 \text{---} \textcircled{0} \text{---} &= \text{---} \textcircled{1PI} \text{---} + \text{---} \textcircled{1PI} \text{---} \textcircled{1PI} \text{---} + \dots \\
 &= \frac{i}{p^2 - m_0^2} + \frac{i}{p^2 - m_0^2} (-iM^2) \frac{i}{p^2 - m_0^2} + \dots \\
 &= \frac{i}{p^2 - m_0^2 - M^2(p^2)}. \tag{7.43}
 \end{aligned}$$

Notice that, as in the case of the electron propagator, our sign convention for the 1PI self-energy $M(p^2)$ implies that a positive contribution to $M(p^2)$ corresponds to a positive shift of the scalar particle mass. If we expand each resummed propagator about the physical particle pole, we see that each external leg of the four-point amplitude contributes

$$\frac{i}{p^2 - m_0^2 - M^2} \underset{p^0 \rightarrow E_p}{\sim} \frac{iZ}{p^2 - m^2} + (\text{regular}). \tag{7.44}$$

Thus, the sum of diagrams contains a product of four poles:

$$\frac{iZ}{p_1^2 - m^2} \frac{iZ}{p_2^2 - m^2} \frac{iZ}{k_1^2 - m^2} \frac{iZ}{k_2^2 - m^2}.$$

This is exactly the singularity on the second line of (7.42). Comparing the

coefficients of this product of poles, we find the relation

$$\langle \mathbf{p}_1 \mathbf{p}_2 | S | \mathbf{k}_1 \mathbf{k}_2 \rangle = (\sqrt{Z})^4 \quad \text{Amp.} \quad ,$$

where the shaded circle is the sum of amputated four-point diagrams and Z is the field-strength renormalization factor.

An identical analysis can be applied to the Fourier transform of the $(n+2)$ -point correlator in a general field theory. The relation between S -matrix elements and Feynman diagrams then takes the form

$$\langle \mathbf{p}_1 \dots \mathbf{p}_n | S | \mathbf{k}_1 \mathbf{k}_2 \rangle = (\sqrt{Z})^{n+2} \quad \text{Amp.} \quad . \quad (7.45)$$

(If the external particles are of different species, each has its own renormalization factor \sqrt{Z} ; if the particles have nonzero spin, there will be additional polarization factors such as $u^s(k)$ on the right-hand side.) This is almost precisely the diagrammatic formula for the S -matrix that we wrote down in Section 4.6. The only new feature is the appearance of the renormalization factors \sqrt{Z} . The Z factors are irrelevant for calculations at the leading order of perturbation theory, but are important in the calculation of higher-order corrections.

Up to this point, we have performed only one full calculation of a higher-order correction, the computation of the order- α corrections to the electron form factors. We did not take into account the effects of the electron field-strength renormalization. Let us now add in this factor and examine its effects.

Since the expressions (6.28) and (6.30) for electron scattering from a heavy target were derived using our previous, incorrect formula for S -matrix elements, we should correct these formulae by inserting factors of $\sqrt{Z_2}$ for the initial and final electrons. Equation (6.33) for the structure of the exact vertex should then read

$$Z_2 \Gamma^\mu(p', p) = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu} q_\nu}{2m} F_2(q^2), \quad (7.46)$$

with $\Gamma^\mu(p', p)$ the sum of amputated electron-photon vertex diagrams.

We can use this equation to reevaluate the form factors to order α . Since $Z_2 = 1 + \mathcal{O}(\alpha)$ and F_2 begins in order α , our previous computation of F_2 is unaffected. To compute F_1 , write the left-hand side of (7.46) as

$$Z_2 \Gamma^\mu = (1 + \delta Z_2)(\gamma^\mu + \delta \Gamma^\mu) = \gamma^\mu + \delta \Gamma^\mu + \gamma^\mu \cdot \delta Z_2,$$

where δZ_2 and $\delta \Gamma^\mu$ denote the order- α corrections to these quantities. Comparing to the right-hand side of (7.46), we see that $F_1(q^2)$ receives a new

contribution equal to δZ_2 . Now let $\delta F_1(q^2)$ denote the (unsubtracted) correction to the form factor that we computed in Section 6.3, and recall from the end of Section 7.1 that $\delta Z_2 = -\delta F_1(0)$. Then

$$F_1(q^2) = 1 + \delta F_1(q^2) + \delta Z_2 = 1 + [\delta F_1(q^2) - \delta F_1(0)].$$

This is exactly the result we claimed, but did not prove, in Section 6.3. The inclusion of field-strength renormalization justifies the subtraction procedure that we applied on an *ad hoc* basis there.

At this level of analysis, it is difficult to see how the cancellation of divergences in F_1 will persist to higher orders. Worse, though we argued in Section 6.3 for the general result $F_1(0) = 1$, our verification of this result in order α seems to depend on a numerical coincidence.

We can state this problem carefully as follows: Define a second rescaling factor Z_1 by the relation

$$\Gamma^\mu(q = 0) = Z_1^{-1} \gamma^\mu, \quad (7.47)$$

where Γ^μ is the complete amputated vertex function. To find $F_1(0) = 1$, we must prove the identity $Z_1 = Z_2$, so that the vertex rescaling exactly compensates the electron field-strength renormalization. We will prove this identity to all orders in perturbation theory at the end Section 7.4.

We conclude our discussion of the LSZ reduction formula with one further formal observation. The LSZ formula distinguishes in and out particles only by the sign of the Fourier transform momentum p_i^0 or k_i^0 . This means that, by analytically continuing the residue of the pole in p^2 from positive to negative p^0 , we can convert the S -matrix element with $\phi(\mathbf{p})$ in the final state into the S -matrix element with the antiparticle $\phi^*(-\mathbf{p})$ in the initial state. This is exactly the statement of *crossing symmetry*, which we derived diagrammatically in Section 5.4:

$$\langle \cdots \phi(p) | S | \cdots \rangle \Big|_{p=-k} = \langle \cdots | S | \phi^*(k) \cdots \rangle.$$

Since the proof of the LSZ formula does not depend on perturbation theory, we see that the crossing symmetry of the S -matrix is a general result of quantum theory, not merely a property of Feynman diagrams.

7.3 The Optical Theorem

In Section 7.1 we saw that the two-point correlation function of quantum fields, viewed as an analytic function of the momentum p^2 , has branch cut singularities associated with multiparticle intermediate states. This conclusion should not be surprising to those familiar with nonrelativistic scattering theory, since it is already true there that the scattering amplitude, as a function of energy, has a branch cut on the positive real axis. The imaginary part of the scattering amplitude appears as a discontinuity across this branch cut. By the *optical theorem*, the imaginary part of the forward scattering amplitude is

$$2\text{Im} = \sum_f \int d\Pi_f \left(\begin{array}{c} k_2 \\ k_1 \end{array} \right) \left(\begin{array}{c} p_2 \\ p_1 \end{array} \right)$$

Figure 7.5. The optical theorem: The imaginary part of any scattering amplitude arises from a sum of contributions from all possible intermediate-state particles.

proportional to the total cross section. We will now prove the field-theoretic version of the optical theorem and illustrate how it arises in Feynman diagram calculations.

The optical theorem is a straightforward consequence of the unitarity of the S -matrix: $S^\dagger S = 1$. Inserting $S = 1 + iT$ as in (4.72), we have

$$-i(T - T^\dagger) = T^\dagger T. \quad (7.48)$$

Let us take the matrix element of this equation between two-particle states $|\mathbf{p}_1 \mathbf{p}_2\rangle$ and $|\mathbf{k}_1 \mathbf{k}_2\rangle$. To evaluate the right-hand side, insert a complete set of intermediate states:

$$\langle \mathbf{p}_1 \mathbf{p}_2 | T^\dagger T | \mathbf{k}_1 \mathbf{k}_2 \rangle = \sum_n \left(\prod_{i=1}^n \int \frac{d^3 q_i}{(2\pi)^3} \frac{1}{2E_i} \right) \langle \mathbf{p}_1 \mathbf{p}_2 | T^\dagger | \{q_i\} \rangle \langle \{q_i\} | T | \mathbf{k}_1 \mathbf{k}_2 \rangle.$$

Now express the T -matrix elements as invariant matrix elements \mathcal{M} times 4-momentum-conserving delta functions. Identity (7.48) then becomes

$$\begin{aligned} 2 \text{Im } \mathcal{M}(k_1 k_2 \rightarrow p_1 p_2) &= \sum_n \left(\prod_{i=1}^n \int \frac{d^3 q_i}{(2\pi)^3} \frac{1}{2E_i} \right) \mathcal{M}^*(p_1 p_2 \rightarrow \{q_i\}) \\ &\quad \times \mathcal{M}(k_1 k_2 \rightarrow \{q_i\}) (2\pi)^4 \delta^{(4)}(k_1 + k_2 - \sum_i q_i), \end{aligned}$$

times an overall delta function $(2\pi)^4 \delta^{(4)}(k_1 + k_2 - p_1 - p_2)$. Let us abbreviate this identity as

$$2 \text{Im } \mathcal{M}(a \rightarrow b) = \sum_f \int d\Pi_f \mathcal{M}^*(b \rightarrow f) \mathcal{M}(a \rightarrow f), \quad (7.49)$$

where the sum runs over all possible sets f of final-state particles. This identity is written pictorially in Fig. 7.5. Although we have so far assumed that a and b are two-particle states, they could equally well be one-particle or multiparticle asymptotic states.

Setting $k_i = p_i$ and supplying the kinematic factors required by (4.79) to build a cross section, we obtain the standard form of the optical theorem,

$$\text{Im } \mathcal{M}(k_1, k_2 \rightarrow k_1, k_2) = 2E_{\text{cm}} p_{\text{cm}} \sigma_{\text{tot}}(k_1, k_2 \rightarrow \text{anything}), \quad (7.50)$$

where E_{cm} is the total center-of-mass energy and p_{cm} is the momentum of either particle in the center-of-mass frame. This equation relates the forward

scattering amplitude to the total cross section for production of all final states. Since the imaginary part of the forward scattering amplitude gives the attenuation of the forward-going wave as the beam passes through the target, it is natural that this quantity should be proportional to the probability of scattering. Identity (7.50) gives the precise connection.

The Optical Theorem for Feynman Diagrams

Let us now investigate how this identity for the imaginary part of an S -matrix element arises in the Feynman diagram expansion. It is easily checked (in QED, for example) that each diagram contributing to an S -matrix element \mathcal{M} is purely real unless some denominators vanish, so that the $i\epsilon$ prescription for treating the poles becomes relevant. A Feynman diagram thus yields an imaginary part for \mathcal{M} only when the virtual particles in the diagram go on-shell. We will now show how to isolate and compute this imaginary part.

For our present purposes, let us *define* \mathcal{M} by the Feynman rules for perturbation theory. This allows us to consider $\mathcal{M}(s)$ as an analytic function of the complex variable $s = E_{\text{cm}}^2$, even though S -matrix elements are defined only for external particles with real momenta.

We first demonstrate that the appearance of an imaginary part of $\mathcal{M}(s)$ always requires a branch cut singularity. Let s_0 be the threshold energy for production of the lightest multiparticle state. For real s below s_0 the intermediate state cannot go on-shell, so $\mathcal{M}(s)$ is real. Thus, for real $s < s_0$, we have the identity

$$\mathcal{M}(s) = [\mathcal{M}(s^*)]^*. \quad (7.51)$$

Each side of this equation is an analytic function of s , so it can be analytically continued to the entire complex s plane. In particular, near the real axis for $s > s_0$, Eq. (7.51) implies

$$\begin{aligned} \text{Re } \mathcal{M}(s + i\epsilon) &= \text{Re } \mathcal{M}(s - i\epsilon); \\ \text{Im } \mathcal{M}(s + i\epsilon) &= -\text{Im } \mathcal{M}(s - i\epsilon). \end{aligned}$$

There is a branch cut across the real axis, starting at the threshold energy s_0 ; the discontinuity across the cut is

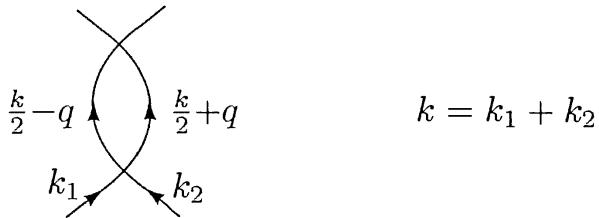
$$\text{Disc } \mathcal{M}(s) = 2i \text{Im } \mathcal{M}(s + i\epsilon).$$

Usually it is easier to compute the discontinuity of a diagram than to compute the imaginary part directly. The $i\epsilon$ prescription in the Feynman propagator indicates that physical scattering amplitudes should be evaluated above the cut, at $s + i\epsilon$.

We already saw in Section 7.1 that the electron self-energy diagram has a branch cut beginning at the physical electron-photon threshold. Let us now study more general one-loop diagrams, and show that their discontinuities

give precisely the imaginary parts required by (7.49). The generalization of this result to multiloop diagrams has been proven by Cutkosky,[†] who showed in the process that the discontinuity of a Feynman diagram across its branch cut is always given by a simple set of *cutting rules*.[‡]

We begin by checking (7.49) in ϕ^4 theory. Since the right-hand side of (7.49) begins in order λ^2 , we expect that $\text{Im } \mathcal{M}$ should also receive its first contribution from higher-order diagrams. Consider, then, the order- λ^2 diagram



with a loop in the s -channel. (It is easy to check that the corresponding t - and u -channel diagrams have no branch cut singularities for s above threshold.) The total momentum is $k = k_1 + k_2$, and for simplicity we have chosen the symmetrical routing of momenta shown above. The value of this Feynman diagram is

$$i\delta\mathcal{M} = \frac{\lambda^2}{2} \int \frac{d^4q}{(2\pi)^4} \frac{1}{(k/2 - q)^2 - m^2 + i\epsilon} \frac{1}{(k/2 + q)^2 - m^2 + i\epsilon}. \quad (7.52)$$

When this integral is evaluated using the methods of Section 6.3, the Wick rotation produces an extra factor of i , so that, below threshold, $\delta\mathcal{M}$ is purely real.

We would like to verify that the integral (7.52) has a discontinuity across the real axis in the physical region $k^0 > 2m$. It is easiest to identify this discontinuity by computing the integral for $k^0 < 2m$, then increasing k^0 by analytic continuation. It is not difficult to compute the integral directly using Feynman parameters (see Problem 7.1). However, it is illuminating to use a less direct approach, as follows.

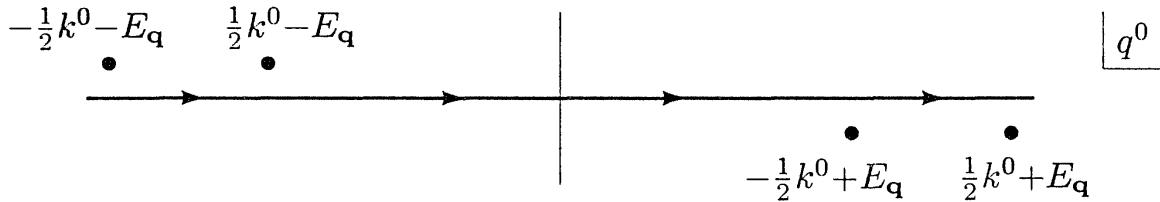
Let us work in the center-of-mass frame, where $k = (k^0, \mathbf{0})$. Then the integrand of (7.52) has four poles in the integration variable q^0 , at the locations

$$q^0 = \frac{1}{2}k^0 \pm (E_{\mathbf{q}} - i\epsilon), \quad q^0 = -\frac{1}{2}k^0 \pm (E_{\mathbf{q}} - i\epsilon).$$

[†]R. E. Cutkosky, *J. Math. Phys.* **1**, 429 (1960).

[‡]These rules are simple only for singularities in the physical region. Away from the physical region, the singularities of three- and higher-point amplitudes can become quite intricate. This subject is reviewed in R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix* (Cambridge University Press, 1966).

Two of these poles lie above the real q^0 axis and two lie below, as shown:



We will close the integration contour downward and pick up the residues of the poles in the lower half-plane. Of these, only the pole at $q^0 = -(1/2)k^0 + E_q$ will contribute to the discontinuity. Note that picking up the residue of this pole is equivalent to replacing

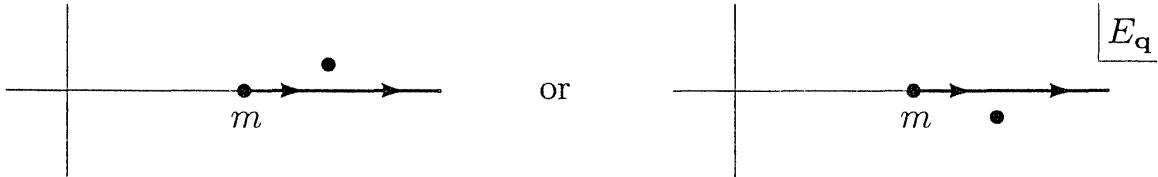
$$\frac{1}{(k/2 + q)^2 - m^2 + i\epsilon} \rightarrow -2\pi i \delta((k/2 + q)^2 - m^2) \quad . \quad (7.53)$$

under the dq^0 integral.

The contribution of this pole yields the integral

$$\begin{aligned} i\delta\mathcal{M} &= -2\pi i \frac{\lambda^2}{2} \int \frac{d^3q}{(2\pi)^4} \frac{1}{2E_q} \frac{1}{(k^0 - E_q)^2 - E_q^2} \\ &= -2\pi i \frac{\lambda^2}{2} \frac{4\pi}{(2\pi)^4} \int_m^\infty dE_q E_q |\mathbf{q}| \frac{1}{2E_q} \frac{1}{k^0(k^0 - 2E_q)}. \end{aligned} \quad (7.54)$$

The integrand in the second line has a pole at $E_q = k^0/2$. When $k^0 < 2m$, this pole does not lie on the integration contour, so $\delta\mathcal{M}$ is manifestly real. When $k^0 > 2m$, however, the pole lies just above or below the contour of integration, depending upon whether k^0 is given a small positive or negative imaginary part:



Thus the integral acquires a discontinuity between $k^2 + i\epsilon$ and $k^2 - i\epsilon$. To compute this discontinuity, apply

$$\frac{1}{k^0 - 2E_q \pm i\epsilon} = P \frac{1}{k^0 - 2E_q} \mp i\pi \delta(k^0 - 2E_q)$$

(where P denotes the principal value). The discontinuity is given by replacing the pole with a delta function. This in turn is equivalent to replacing the original propagator by a delta function:

$$\frac{1}{(k/2 - q)^2 - m^2 + i\epsilon} \rightarrow -2\pi i \delta((k/2 - q)^2 - m^2). \quad (7.55)$$

$$\begin{aligned}
 (a) \quad 2\text{Im} \left(\text{Diagram with two vertical lines and a central horizontal line} \right) &= \int d\Pi \left| \text{Diagram with two vertical lines and a central horizontal line} \right|^2 \\
 (b) \quad 2\text{Im} \left(\text{Diagram with two vertical lines and a central circle} \right) &= \int d\Pi \left| \text{Diagram with two vertical lines and a central circle} \right|^2
 \end{aligned}$$

Figure 7.6. Two contributions to the optical theorem for Bhabha scattering.

Let us now retrace our steps and see what we have proved. Go back to the original integral (7.52), relabel the momenta on the two propagators as p_1, p_2 and substitute

$$\int \frac{d^4 q}{(2\pi)^4} = \int \frac{d^4 p_1}{(2\pi)^4} \int \frac{d^4 p_2}{(2\pi)^4} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - k).$$

We have shown that the discontinuity of the integral is computed by replacing each of the two propagators by a delta function:

$$\frac{1}{p_i^2 - m^2 + i\epsilon} \rightarrow -2\pi i \delta(p_i^2 - m^2). \quad (7.56)$$

The discontinuity of \mathcal{M} comes only from the region of the $d^4 q$ integral in which the two delta functions are simultaneously satisfied. By integrating over the delta functions, we put the momenta p_i on shell and convert the integrals $d^4 p_i$ into integrals over relativistic phase space. What is left over in expression (7.52) is just the factor λ^2 , the square of the leading-order scattering amplitude, and the symmetry factor $(1/2)$, which can be reinterpreted as the symmetry factor for identical bosons in the final state. Thus we have shown that, to order λ^2 on each side of the equation,

$$\begin{aligned}
 \text{Disc } \mathcal{M}(k) &= 2i \text{Im } \mathcal{M}(k) \\
 &= \frac{i}{2} \int \frac{d^3 p_1}{(2\pi)^3} \frac{1}{2E_1} \frac{d^3 p_2}{(2\pi)^3} \frac{1}{2E_2} |\mathcal{M}(k)|^2 (2\pi)^4 \delta^{(4)}(p_1 + p_2 - k).
 \end{aligned}$$

This explicitly verifies (7.49) to order λ^2 in ϕ^4 theory.

The preceding argument made no essential use of the fact that the two propagators in the diagram had equal masses, or of the fact that these propagators connected to a simple point vertex. Indeed, the analysis can be applied to an arbitrary one-loop diagram. Whenever, in the region of momentum integration of the diagram, two propagators can simultaneously go on-shell, we can follow the argument above to compute a nonzero discontinuity of \mathcal{M} . The value of this discontinuity is given by making the substitution (7.56) for

each of the two propagators. For example, in the order- α^2 Bhabha scattering diagrams shown in Fig. 7.6, we can compute the imaginary parts by cutting through the diagrams as shown and putting the cut propagators on shell using (7.56). The poles of the additional propagators in the diagrams do not contribute to the discontinuities. By integrating over the delta functions as in the previous paragraph, we derive the indicated relations between the imaginary parts of these diagrams and contributions to the total cross section.

Cutkosky proved that this method of computing discontinuities is completely general. The physical discontinuity of any Feynman diagram is given by the following algorithm:

1. Cut through the diagram in all possible ways such that the cut propagators can simultaneously be put on shell.
2. For each cut, replace $1/(p^2 - m^2 + i\epsilon) \rightarrow -2\pi i\delta(p^2 - m^2)$ in each cut propagator, then perform the loop integrals.
3. Sum the contributions of all possible cuts.

Using these *cutting rules*, it is possible to prove the optical theorem (7.49) to all orders in perturbation theory.

Unstable Particles

The cutting rules imply that the generalized optical theorem (7.49) is true not only for S -matrix elements, but for any amplitudes \mathcal{M} that we can define in terms of Feynman diagrams. This fact is extremely useful for dealing with unstable particles, which never appear in asymptotic states.

Recall from Eq. (7.43) that the exact two-point function for a scalar particle has the form

$$\text{---} \circ \text{---} = \frac{i}{p^2 - m_0^2 - M^2(p^2)}.$$

We defined the quantity $-iM^2(p^2)$ as the sum of all 1PI insertions into the boson propagator, but we can equally well think of it as the sum of all amputated diagrams for 1-particle \rightarrow 1-particle “scattering”. The LSZ formula then implies

$$\mathcal{M}(p \rightarrow p) = -ZM^2(p^2). \quad (7.57)$$

We can use this relation and the generalized optical theorem (7.49) to discuss the imaginary part of $M^2(p^2)$.

First consider the familiar case where the scalar boson is stable. In this case, there is no possible final state that can contribute to the right-hand side of (7.49). Thus $M^2(p^2)$ is real. The position of the pole in the propagator is determined by the equation $m^2 - m_0^2 - M^2(m^2) = 0$, which has a real-valued solution m . The pole therefore lies on the real p^2 axis, below the multiparticle branch cut.

Often, however, a particle can decay into two or more lighter particles. In this case $M^2(p^2)$ will acquire an imaginary part, so we must modify our

definitions slightly. Let us define the particle's mass m by the condition

$$m^2 - m_0^2 - \text{Re } M^2(m^2) = 0. \quad (7.58)$$

Then the pole in the propagator is displaced from the real axis:

$$\text{---} \circ \text{---} \sim \frac{iZ}{p^2 - m^2 - iZ \text{Im } M^2(p^2)}.$$

If this propagator appears in the s channel of a Feynman diagram, the cross section one computes, in the vicinity of the pole, will have the form

$$\sigma \propto \left| \frac{1}{s - m^2 - iZ \text{Im } M^2(s)} \right|^2. \quad (7.59)$$

This expression closely resembles the relativistic Breit-Wigner formula (4.64) for the cross section in the region of a resonance:

$$\sigma \propto \left| \frac{1}{p^2 - m^2 + im\Gamma} \right|^2. \quad (7.60)$$

The mass m defined by (7.58) is the position of the resonance. If $\text{Im } M(m^2)$ is small, so that the resonance in (7.59) is narrow, we can approximate $\text{Im } M^2(s)$ as $\text{Im } M^2(m^2)$ over the width of the resonance; then (7.59) will have precisely the Breit-Wigner form. In this case, we can identify

$$\Gamma = -\frac{Z}{m} \text{Im } M^2(m^2). \quad (7.61)$$

If the resonance is broad, it will show deviations from the Breit-Wigner shape, generally becoming narrower on the leading edge and broader on the trailing edge.

To compute $\text{Im } M^2$, and hence Γ , we could use the definition of M^2 as the sum of 1PI insertions into the propagator. The imaginary parts of the relevant loop diagrams give the decay rate. But the optical theorem (7.49), generalized to Feynman diagrams by the Cutkosky rules, simplifies this procedure. If we take (7.57) as the definition of the matrix element $\mathcal{M}(p \rightarrow p)$, and similarly define the decay matrix elements $\mathcal{M}(p \rightarrow f)$ through their Feynman diagram expansions, then (7.49) implies

$$Z \text{Im } M^2(p^2) = -\text{Im } \mathcal{M}(p \rightarrow p) = -\frac{1}{2} \sum_f \int d\Pi_f |\mathcal{M}(p \rightarrow f)|^2, \quad (7.62)$$

where the sum runs over all possible final states f . The decay rate is therefore

$$\Gamma = \frac{1}{2m} \sum_f \int d\Pi_f |\mathcal{M}(p \rightarrow f)|^2, \quad (7.63)$$

as quoted in Eq. (4.86).

We stress once again that our derivation of this equation applies only to the case of a long-lived unstable particle, so that $\Gamma \ll m$. For a broad resonance, the full energy dependence of $M^2(p^2)$ must be taken into account.

7.4 The Ward-Takahashi Identity

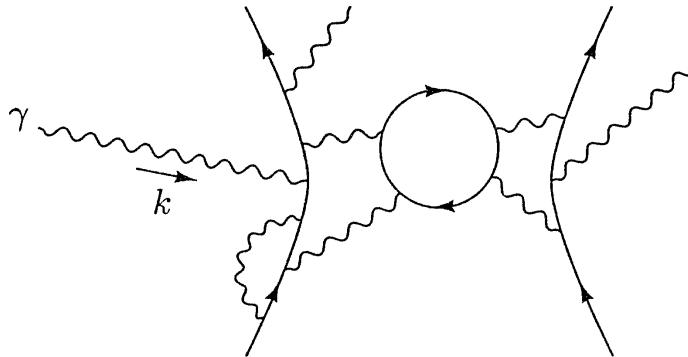
Of the loose ends listed at the beginning of this chapter, only one remains, the proof of the Ward identity. Recall from Section 5.5 that this identity states the following: If $\mathcal{M}(k) = \epsilon_\mu(k)\mathcal{M}^\mu(k)$ is the amplitude for some QED process involving an external photon with momentum k , then this amplitude vanishes if we replace ϵ_μ with k_μ :

$$k_\mu \mathcal{M}^\mu(k) = 0. \quad (7.64)$$

To prove this assertion, it is useful to prove a more general identity for QED correlation functions, called the *Ward-Takahashi identity*. To discuss this more general case we will let \mathcal{M} denote a Fourier-transformed correlation function, in which the external momenta are not necessarily on-shell. The right-hand side of (7.64) will contain nonzero terms in this case; but when we apply the LSZ formula to extract an S -matrix element, those terms will not contribute.

We will prove the Ward-Takahashi identity order by order in α , by looking directly at the Feynman diagrams that contribute to $\mathcal{M}(k)$. The identity is generally not true for individual Feynman diagrams; we must sum over the diagrams for $\mathcal{M}(k)$ at any given order.

Consider a typical diagram for a typical amplitude $\mathcal{M}(k)$:

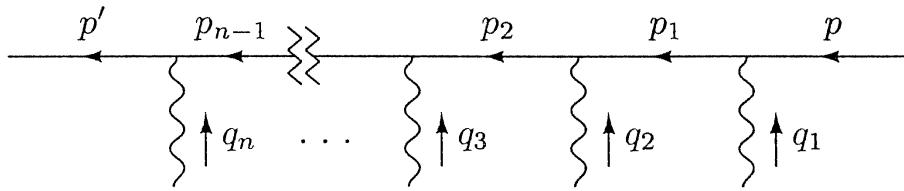


If we remove the photon $\gamma(k)$, we obtain a simpler diagram which is part of a simpler amplitude \mathcal{M}_0 . If we reinsert the photon somewhere else inside the simpler diagram, we again obtain a contribution to $\mathcal{M}(k)$. The crucial observation is that by summing over all the diagrams that contribute to \mathcal{M}_0 , then summing over all possible points of insertion in each of these diagrams, we obtain $\mathcal{M}(k)$. The Ward-Takahashi identity is true individually for each diagram contributing to \mathcal{M}_0 , once we sum over insertion points; this is what we will prove.

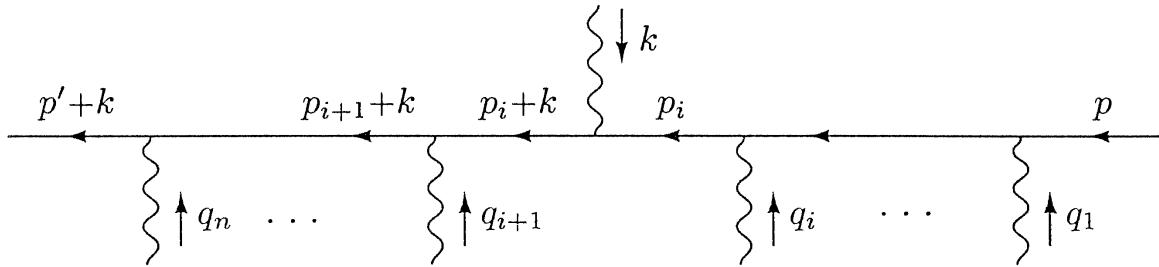
When we insert our photon into one of the diagrams of \mathcal{M}_0 , it must attach either to an electron line that runs out of the diagram to two external points, or to an internal electron loop. Let us consider each of these cases in turn.

First suppose that the electron line runs between external points. Before

we insert our photon $\gamma(k)$, the line looks like this:



The electron propagators have momenta p , $p_1 = p + q_1$, $p_2 = p_1 + q_2$, and so on up to $p' = p_{n-1} + q_n$. If there are n vertices, we can insert our photon in $n + 1$ different places. Suppose we insert it after the i th vertex:



The electron propagators to the left of the new photon then have their momenta increased by k .

Let us now look at the values of these diagrams, with the polarization vector $\epsilon_\mu(k)$ replaced by k_μ . The product of k_μ with the new vertex is conveniently written:

$$-ie k_\mu \gamma^\mu = -ie [(\not{p}_i + \not{k} - m) - (\not{p}_i - m)].$$

Multiplying by the adjacent electron propagators, we obtain

$$\frac{i}{\not{p}_i + \not{k} - m} (-ie \not{k}) \frac{i}{\not{p}_i - m} = e \left(\frac{i}{\not{p}_i - m} - \frac{i}{\not{p}_i + \not{k} - m} \right). \quad (7.65)$$

The diagram with the photon inserted at position i therefore has the structure

$$\dots \xleftarrow{\quad} \begin{array}{c} \not{k} \\ \swarrow \quad \nwarrow \\ \not{q}_{i+1} \quad \not{q}_i \end{array} \xleftarrow{\quad} \dots = \dots \left(\frac{i}{\not{p}_{i+1} + \not{k} - m} \right) \gamma^{\lambda_{i+1}} \left(\frac{i}{\not{p}_i - m} - \frac{i}{\not{p}_i + \not{k} - m} \right) \gamma^{\lambda_i} \\ \times \left(\frac{i}{\not{p}_{i-1} - m} \right) \gamma^{\lambda_{i-1}} \dots$$

Similarly, the diagram with the photon inserted at position $i - 1$ has the structure

$$\dots \xleftarrow{\quad} \begin{array}{c} \not{k} \\ \swarrow \quad \nwarrow \\ \not{q}_{i+1} \quad \not{q}_i \end{array} \xleftarrow{\quad} \dots = \dots \left(\frac{i}{\not{p}_{i+1} + \not{k} - m} \right) \gamma^{\lambda_{i+1}} \left(\frac{i}{\not{p}_i + \not{k} - m} \right) \gamma^{\lambda_i} \\ \times \left(\frac{i}{\not{p}_{i-1} - m} - \frac{i}{\not{p}_{i-1} + \not{k} - m} \right) \gamma^{\lambda_{i-1}} \dots$$

Note that the first term of this expression cancels the second term of the previous expression. A similar cancellation occurs between any other pair of

diagrams with adjacent insertions. When we sum over all possible insertion points along the line, everything cancels except the unpaired terms at the ends. The unpaired term coming from insertion after the last vertex (on the far left) is just e times the value of the original diagram; the other unpaired term, from insertion before the first vertex, is identical except for a minus sign and the replacement of p by $p + k$ everywhere. Diagrammatically, our result is

$$\sum_{\text{insertion points}} k_\mu \cdot \left(\mu \begin{array}{c} \text{---} \\ \text{---} \\ \vdots \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \\ \vdots \\ \text{---} \end{array} \right) = e \left(\begin{array}{c} \text{---} \\ \text{---} \\ \vdots \\ \text{---} \end{array} - \begin{array}{c} \text{---} \\ \text{---} \\ \vdots \\ \text{---} \end{array} \right), \quad (7.66)$$

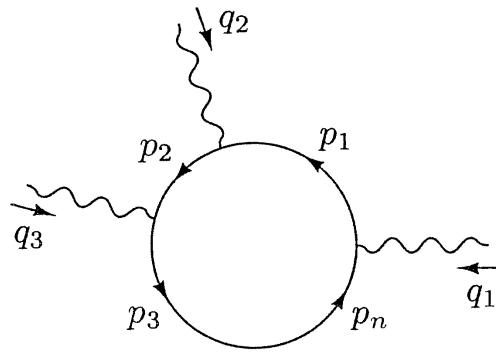
where we have renamed $p' + k \rightarrow q$ for the sake of symmetry.

In each diagram on the left-hand side of (7.66), the momentum entering the electron line is p and the momentum exiting is q . According to the LSZ formula, we can extract from each diagram a contribution to an S -matrix element by taking the coefficient of the product of poles

$$\left(\frac{i}{q-m} \right) \left(\frac{i}{p-m} \right).$$

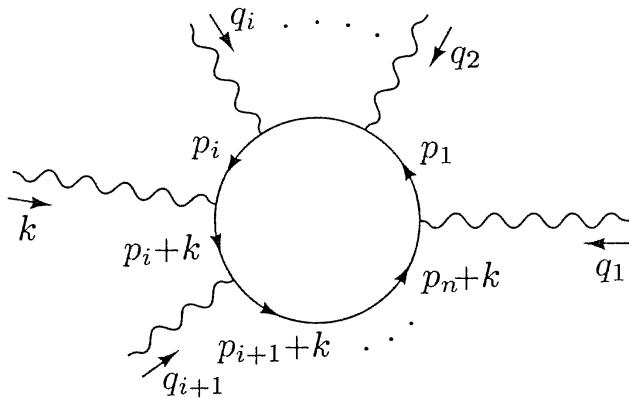
The terms on the right-hand side of (7.66) each contain one of these poles, but neither contains both poles. Thus the right-hand side of (7.66) contributes nothing to the S -matrix.*

To complete the proof of the Ward-Takahashi identity, we must consider the case in which our photon attaches to an internal electron loop. Before the insertion of the photon, a typical loop looks like this:



*This step of the argument is straightforward only if we have arranged the perturbation series so that the propagator contains m rather than m_0 . We will do this in Chapter 10.

The electron propagators have momenta $p_1, p_1 + q_2 = p_2$, and so on up to p_n . Suppose now that we insert the photon $\gamma(k)$ between vertices i and $i + 1$:



We now have an additional momentum k running around the loop from the new vertex; by convention, this momentum exits at vertex 1.

To evaluate the sum over all such insertions into the loop, apply identity (7.65) to each diagram. For the diagram in which the photon is inserted between vertices 1 and 2, we obtain

$$-e \int \frac{d^4 p_1}{(2\pi)^4} \text{tr} \left[\left(\frac{i}{\not{p}_n + \not{k} - m} \right) \gamma^{\lambda_n} \cdots \left(\frac{i}{\not{p}_2 + \not{k} - m} \right) \gamma^{\lambda_2} \times \left(\frac{i}{\not{p}_1 - m} - \frac{i}{\not{p}_1 + \not{k} - m} \right) \gamma^{\lambda_1} \right].$$

The first term will be canceled by one of the terms from the diagram with the photon inserted between vertices 2 and 3. Similar cancellations take place between terms from other pairs of adjacent insertions. When we sum over all n insertion points we are left with

$$-e \int \frac{d^4 p_1}{(2\pi)^4} \text{tr} \left[\left(\frac{i}{\not{p}_n - m} \right) \gamma^{\lambda_n} \left(\frac{i}{\not{p}_{n-1} - m} \right) \gamma^{\lambda_{n-1}} \cdots \left(\frac{i}{\not{p}_1 - m} \right) \gamma^{\lambda_1} - \left(\frac{i}{\not{p}_n + \not{k} - m} \right) \gamma^{\lambda_n} \left(\frac{i}{\not{p}_{n-1} + \not{k} - m} \right) \gamma^{\lambda_{n-1}} \cdots \left(\frac{i}{\not{p}_1 + \not{k} - m} \right) \gamma^{\lambda_1} \right]. \quad (7.67)$$

Shifting the integration variable from p_1 to $p_1 + k$ in the second term, we see that the two remaining terms cancel. Thus the diagrams in which the photon is inserted along a closed loop add up to zero.

We are now ready to assemble the pieces of the proof. Suppose that the amplitude \mathcal{M} has $2n$ external electron lines, n incoming and n outgoing. Label

the incoming momenta p_i and the outgoing momenta q_i :

$$\mathcal{M}(k; p_1 \cdots p_n; q_1 \cdots q_n) = \text{Diagram with a shaded circle, incoming wavy line } k \text{ from the left, outgoing wavy lines } q_1, \dots, q_n \text{ and } p_1, \dots, p_n \text{ from the right.}$$

(The amplitude can also involve an arbitrary number of additional external photons.) The amplitude \mathcal{M}_0 lacks the photon $\gamma(k)$ but is otherwise identical. To form $k_\mu \mathcal{M}^\mu$ from \mathcal{M}_0 we must sum over all diagrams that contribute to \mathcal{M}_0 , and for each diagram, sum over all points at which the photon could be inserted. Summing over insertion points along an internal loop in any diagram gives zero. Summing over insertion points along a through-going line in any diagram gives a contribution of the form (7.66). Summing over *all* insertion points for any particular diagram, we obtain

$$\sum_{\text{insertion points}} k_\mu \cdot \left(\text{Diagram with a shaded circle, incoming wavy line } k, \text{ outgoing wavy lines } (q_1 \cdots q_n) \text{ and } (p_1 \cdots p_n) \right) = e \sum_i \left(\text{Diagram with a shaded circle, incoming wavy line } k, \text{ outgoing wavy lines } (q_1 \cdots (q_i - k) \cdots) \text{ and } (p_1 \cdots p_n) - \text{Diagram with a shaded circle, incoming wavy line } k, \text{ outgoing wavy lines } (q_1 \cdots q_n) \text{ and } (p_1 \cdots (p_i + k) \cdots) \right),$$

where the shaded circle represents any particular diagram that contributes to \mathcal{M}_0 . Summing over all such diagrams, we finally obtain

$$k_\mu \mathcal{M}^\mu(k; p_1 \cdots p_n; q_1 \cdots q_n) = e \sum_i \left[\mathcal{M}_0(p_1 \cdots p_n; q_1 \cdots (q_i - k) \cdots) - \mathcal{M}_0(p_1 \cdots (p_i + k) \cdots; q_1 \cdots q_n) \right]. \quad (7.68)$$

This is the Ward-Takahashi identity for correlation functions in QED. We saw below (7.66) that the right-hand side does not contribute to the S -matrix; thus in the special case where \mathcal{M} is an S -matrix element, Eq. (7.68) reduces to the Ward identity (7.64).

Before discussing this identity further, we should mention a potential flaw in the above proof. In order to find the necessary cancellation in Eq. (7.67), we had to shift the integration variable by a constant. If the integral diverges, however, this shift is not permissible. Similarly, there may be divergent loop-momentum integrals in the expressions leading to Eq. (7.66). Here there is no explicit shift in the proof, but in practice one does generally perform a shift while evaluating the integrals. In either case, ultraviolet divergences can potentially invalidate the Ward-Takahashi identity. We will see an example of this problem, as well as a general solution to it, in the next section.

The simplest example of the Ward-Takahashi identity involves, on the left-hand side, the three-point function with one entering and one exiting electron and one external photon:

$$k_\mu \cdot \left(\mu \sim \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right) = e \left(\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} - \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right)$$

The quantities on the right-hand side are exact electron propagators, evaluated at p and $(p + k)$ respectively. Label these quantities $S(p)$ and $S(p + k)$; from Eq. (7.23),

$$S(p) = \frac{i}{\not{p} - m - \Sigma(p)}.$$

The full three-point amplitude on the left-hand side can be rewritten, just as in Eq. (7.44), as a product of full propagators for the entering and exiting electrons, times an amputated scattering diagram. In this case, the amputated function is just the vertex $\Gamma^\mu(p + k, p)$. Then the Ward-Takahashi identity reads:

$$S(p + k) [-ie k_\mu \Gamma^\mu(p + k, p)] S(p) = e(S(p) - S(p + k)).$$

To simplify this equation, multiply, on the left and right respectively, by the Dirac matrices $S^{-1}(p + k)$ and $S^{-1}(p)$. This gives

$$-ik_\mu \Gamma^\mu(p + k, p) = S^{-1}(p + k) - S^{-1}(p). \quad (7.69)$$

Often the term *Ward-Takahashi identity* is used to mean only this special case.

We can use identity (7.69) to obtain the general relation between the renormalization factors Z_1 and Z_2 . We defined Z_1 in (7.47) by the relation

$$\Gamma^\mu(p + k, p) \rightarrow Z_1^{-1} \gamma^\mu \quad \text{as } k \rightarrow 0.$$

We defined Z_2 as the residue of the pole in $S(p)$:

$$S(p) \sim \frac{iZ_2}{\not{p} - m}.$$

Setting p near mass shell and expanding (7.69) about $k = 0$, we find for the first-order terms on the left and right

$$-iZ_1^{-1} \not{k} = -iZ_2^{-1} \not{k},$$

that is,

$$Z_1 = Z_2. \quad (7.70)$$

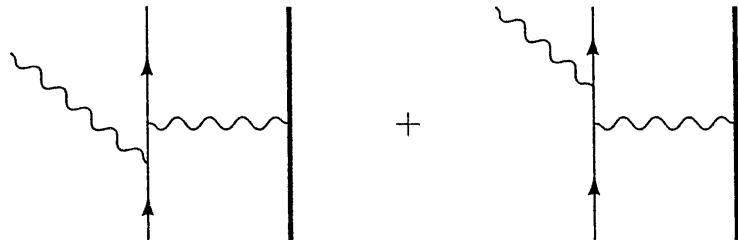
Thus, the Ward-Takahashi identity guarantees the exact cancellation of infinite rescaling factors in the electron scattering amplitude that we found at the end of Section 7.2. When combined with the correct formal expression

(7.46) for the electron form factors, this identity guarantees that $F_1(0) = 1$ to all orders in perturbation theory.

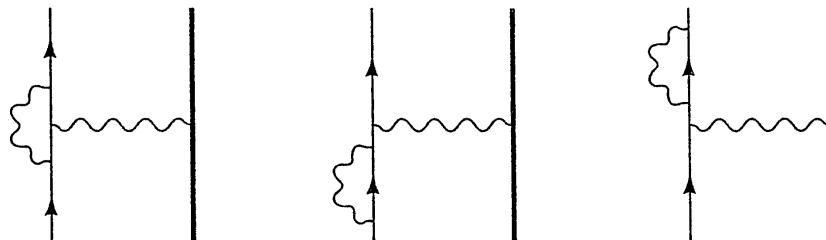
Often, in the literature, the terms *Ward identity*, *current conservation*, and *gauge invariance* are used interchangeably. This is quite natural, since the Ward identity is the diagrammatic expression of the conservation of the electric current, which is in turn a consequence of gauge invariance. In this book, however, we will distinguish these three concepts. By *gauge invariance* we mean the fundamental symmetry of the Lagrangian; by *current conservation* we mean the equation of motion that follows from this symmetry; and by the *Ward identity* we mean the diagrammatic identity that imposes the symmetry on quantum mechanical amplitudes.

7.5 Renormalization of the Electric Charge

At the beginning of Chapter 6 we set out to study the order- α radiative corrections to electron scattering from a heavy target. We evaluated (at least in the classical limit) the bremsstrahlung diagrams,

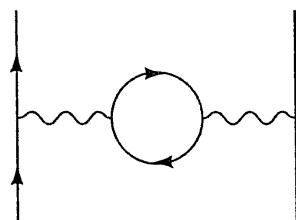


and also the corrections due to virtual photons:



Our discussion of field-strength renormalization in this chapter has finally clarified the role of the last two diagrams. In particular, we have seen that the Ward identity, through the relation $Z_1 = Z_2$, insures that the sum of the virtual photon corrections vanishes as the momentum transfer q goes to zero.

There is one remaining type of radiative correction to this process:



This is the order- α *vacuum polarization* diagram, also known as the *photon self-energy*. It can be viewed as a modification to the photon structure by a

virtual electron-positron pair. This diagram will alter the effective field $A^\mu(x)$ seen by the scattered electron. It can potentially shift the overall strength of this field, and can also change its dependence on x (or in Fourier space, on q). In this section we will compute this diagram, and see that it has both of these effects.

Overview of Charge Renormalization

Before beginning a detailed calculation, let's ask what kind of an answer we expect and what its interpretation will be. The interesting part of the diagram is the electron loop:

$$(-ie)^2(-1) \int \frac{d^4k}{(2\pi)^4} \text{tr} \left[\gamma^\mu \frac{i}{k-m} \gamma^\nu \frac{i}{k+q-m} \right] \equiv i\Pi_2^{\mu\nu}(q). \quad (7.71)$$

(The fermion loop factor of (-1) was derived in Eq. (4.120).) More generally, let us define $i\Pi^{\mu\nu}(q)$ to be the sum of all 1-particle-irreducible insertions into the photon propagator,

$$\equiv i\Pi^{\mu\nu}(q), \quad (7.72)$$

so that $\Pi_2^{\mu\nu}(q)$ is the second-order (in e) contribution to $\Pi^{\mu\nu}(q)$.

The only tensors that can appear in $\Pi^{\mu\nu}(q)$ are $g^{\mu\nu}$ and $q^\mu q^\nu$. The Ward identity, however, tells us that $q_\mu \Pi^{\mu\nu}(q) = 0$. This implies that $\Pi^{\mu\nu}(q)$ is proportional to the projector $(g^{\mu\nu} - q^\mu q^\nu/q^2)$. Furthermore, we expect that $\Pi^{\mu\nu}(q)$ will not have a pole at $q^2 = 0$; the only obvious source of such a pole would be a single-massless-particle intermediate state, which cannot occur in any 1PI diagram.[†] It is therefore convenient to extract the tensor structure from $\Pi^{\mu\nu}$ in the following way:

$$\Pi^{\mu\nu}(q) = (q^2 g^{\mu\nu} - q^\mu q^\nu) \Pi(q^2), \quad (7.73)$$

where $\Pi(q^2)$ is regular at $q^2 = 0$.

Using this notation, the exact photon two-point function is

$$\begin{aligned} \text{---} \circ \text{---} \nu &= \text{---} \text{---} + \text{---} \circ \text{1PI} \text{---} + \text{---} \circ \text{1PI} \text{---} \circ \text{1PI} \text{---} + \dots \\ &= \frac{-ig_{\mu\nu}}{q^2} + \frac{-ig_{\mu\rho}}{q^2} \left[i(q^2 g^{\rho\sigma} - q^\rho q^\sigma) \Pi(q^2) \right] \frac{-ig_{\sigma\nu}}{q^2} + \dots \end{aligned}$$

[†]One can prove that there is no such pole, but the proof is nontrivial. Schwinger has shown that, in two spacetime dimensions, the singularity in Π_2 due to a pair of massless fermions is a pole rather than a cut; this is a famous counterexample to our argument. There is no such problem in four dimensions.

$$= \frac{-ig_{\mu\nu}}{q^2} + \frac{-ig_{\mu\rho}}{q^2} \Delta_\nu^\rho \Pi(q^2) + \frac{-ig_{\mu\rho}}{q^2} \Delta_\sigma^\rho \Delta_\nu^\sigma \Pi^2(q^2) + \dots,$$

where $\Delta_\nu^\rho \equiv \delta_\nu^\rho - q^\rho q_\nu/q^2$. Noting that $\Delta_\sigma^\rho \Delta_\nu^\sigma = \Delta_\nu^\rho$, we can simplify this expression to

$$\begin{aligned}
\text{Diagram with a shaded loop} &= \frac{-ig_{\mu\nu}}{q^2} + \frac{-ig_{\mu\rho}}{q^2} \left(\delta_\nu^\rho - \frac{q^\rho q_\nu}{q^2} \right) (\Pi(q^2) + \Pi^2(q^2) + \dots) \\
&= \frac{-i}{q^2(1 - \Pi(q^2))} \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) + \frac{-i}{q^2} \left(\frac{q_\mu q_\nu}{q^2} \right). \quad (7.74)
\end{aligned}$$

In any S -matrix element calculation, at least one end of this exact propagator will connect to a fermion line. When we sum over all places along the line where it could connect, we must find, according to the Ward identity, that terms proportional to q_μ or q_ν vanish. For the purposes of computing S -matrix elements, therefore, we can abbreviate

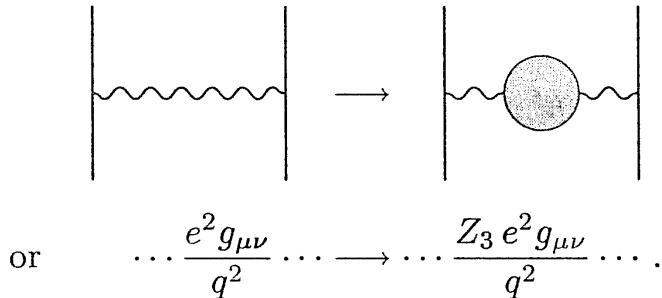
$$\mu \text{---} \text{---} \nu = \frac{-ig_{\mu\nu}}{q^2(1 - \Pi(q^2))}. \quad (7.75)$$

Notice that as long as $\Pi(q^2)$ is regular at $q^2 = 0$, the exact propagator always has a pole at $q^2 = 0$. In other words, the photon remains absolutely massless at all orders in perturbation theory. This claim depends strongly on our use of the Ward identity in (7.73). If, for example, $\Pi^{\mu\nu}(q)$ contained a term $M^2 g^{\mu\nu}$ (with no compensating $q^\mu q^\nu$ term), the photon mass would be shifted to M .

The residue of the $q^2 = 0$ pole is

$$\frac{1}{1 - \Pi(0)} \equiv Z_3.$$

The amplitude for any low- q^2 scattering process will be shifted by this factor, relative to the tree-level approximation:



Since a factor of e lies at each end of the photon propagator, we can conveniently account for this shift by making the replacement $e \rightarrow \sqrt{Z_3} e$. This replacement is called *charge renormalization*; it is in many ways analogous to the mass renormalization introduced in Section 7.1. Note in particular that the “physical” electron charge measured in experiments is $\sqrt{Z_3} e$. We will therefore shift our notation and call this quantity simply e . From now on we

will refer to the “bare” charge (the quantity that multiplies $A_\mu \bar{\psi} \gamma^\mu \psi$ in the Lagrangian) as e_0 . We then have

$$(\text{physical charge}) = e = \sqrt{Z_3} e_0 = \sqrt{Z_3} \cdot (\text{bare charge}). \quad (7.76)$$

To lowest order, $Z_3 = 1$ and $e = e_0$.

In addition to this constant shift in the strength of the electric charge, $\Pi(q^2)$ has another effect. Consider a scattering process with nonzero q^2 , and suppose that we have computed $\Pi(q^2)$ to leading order in α : $\Pi(q^2) \approx \Pi_2(q^2)$. The amplitude for the process will then involve the quantity

$$\frac{-ig_{\mu\nu}}{q^2} \left(\frac{e_0^2}{1 - \Pi(q^2)} \right) \underset{\mathcal{O}(\alpha)}{=} \frac{-ig_{\mu\nu}}{q^2} \left(\frac{e^2}{1 - [\Pi_2(q^2) - \Pi_2(0)]} \right).$$

(Swapping e^2 for e_0^2 does not matter to lowest order.) The quantity in parentheses can be interpreted as a q^2 -dependent electric charge. The full effect of replacing the tree-level photon propagator with the exact photon propagator is therefore to replace

$$\alpha_0 \rightarrow \alpha_{\text{eff}}(q^2) = \frac{e_0^2/4\pi}{1 - \Pi(q^2)} \underset{\mathcal{O}(\alpha)}{=} \frac{\alpha}{1 - [\Pi_2(q^2) - \Pi_2(0)]}. \quad (7.77)$$

(To leading order we could just as well bring the Π -terms into the numerator; but we will see in Chapter 12 that in this form, the expression is true to all orders when Π_2 is replaced by Π .)

Computation of Π_2

Having worked so hard to interpret $\Pi_2(q^2)$, we had better calculate it. Going back to (7.71), we have

$$\begin{aligned} i\Pi_2^{\mu\nu}(q) &= -(-ie)^2 \int \frac{d^4 k}{(2\pi)^4} \text{tr} \left[\gamma^\mu \frac{i(k+m)}{k^2 - m^2} \gamma^\nu \frac{i(k+q+m)}{(k+q)^2 - m^2} \right] \\ &= -4e^2 \int \frac{d^4 k}{(2\pi)^4} \frac{k^\mu (k+q)^\nu + k^\nu (k+q)^\mu - g^{\mu\nu} (k \cdot (k+q) - m^2)}{(k^2 - m^2)((k+q)^2 - m^2)}. \end{aligned} \quad (7.78)$$

We have written e and m instead of e_0 and m_0 for convenience, since the difference would give only an order- α^2 contribution to $\Pi^{\mu\nu}$.

Now introduce a Feynman parameter to combine the denominator factors:

$$\begin{aligned} \frac{1}{(k^2 - m^2)((k+q)^2 - m^2)} &= \int_0^1 dx \frac{1}{(k^2 + 2xk \cdot q + xq^2 - m^2)^2} \\ &= \int_0^1 dx \frac{1}{(\ell^2 + x(1-x)q^2 - m^2)^2}, \end{aligned}$$

where $\ell = k + xq$. In terms of ℓ , the numerator of (7.78) is

$$\begin{aligned} \text{Numerator} = & 2\ell^\mu\ell^\nu - g^{\mu\nu}\ell^2 - 2x(1-x)q^\mu q^\nu + g^{\mu\nu}(m^2 + x(1-x)q^2) \\ & + (\text{terms linear in } \ell). \end{aligned}$$

Performing a Wick rotation and substituting $\ell^0 = i\ell_E^0$, we obtain

$$\begin{aligned} i\Pi_2^{\mu\nu}(q) = & -4ie^2 \int_0^1 dx \int \frac{d^4\ell_E}{(2\pi)^4} \\ & \times \frac{-\frac{1}{2}g^{\mu\nu}\ell_E^2 + g^{\mu\nu}\ell_E^2 - 2x(1-x)q^\mu q^\nu + g^{\mu\nu}(m^2 + x(1-x)q^2)}{(\ell_E^2 + \Delta)^2}, \end{aligned} \quad (7.79)$$

where $\Delta = m^2 - x(1-x)q^2$. This integral is very badly ultraviolet divergent. If we were to cut it off at $\ell_E = \Lambda$, we would find for the leading term,

$$i\Pi_2^{\mu\nu}(q) \propto e^2 \Lambda^2 g^{\mu\nu},$$

with no compensating $q^\mu q^\nu$ term. This result violates the Ward identity; it would give the photon an infinite mass $M \propto e\Lambda$.

Our proof of the Ward identity has failed, in precisely the way anticipated at the end of the previous section: The shift of the integration variable in (7.67) is not permissible when the integral is divergent. In our present calculation, the failure of the Ward identity is catastrophic: It leads to an infinite photon mass,[‡] in conflict with experiment. Fortunately, there is a way to rescue the Ward identity.

In the above analysis we regulated the divergent integral in the most straightforward and most naive way: by cutting it off at a large momentum Λ . But other regulators are possible, and some will in fact preserve the Ward identity. In our computations of the vertex and electron self-energy diagrams, we used a Pauli-Villars regulator. This regulator preserved the relation $Z_1 = Z_2$, a consequence of the Ward identity; a naive cutoff does not (see Problem 7.2). We could fix our present computation by introducing Pauli-Villars fermions. Unfortunately, several sets of fermions are required, making the method rather complicated.* We will use a simpler method, *dimensional regularization*, due to 't Hooft and Veltman.[†] Dimensional regularization preserves the symmetries of QED and also of a wide class of more general theories.

The question of which regulator to use has no *a priori* answer in quantum field theory. Often the choice has no effect on the predictions of the theory.

[‡]We could still make the observed photon mass zero by adding a compensating infinite photon mass term to the Lagrangian. More generally, we could add terms to the Lagrangian to make the Ward identity valid for any n -point correlation function. This procedure would give the same results as the one we are about to follow, but would be much more complicated.

*This method is presented in Bjorken and Drell (1964), p. 154.

[†]G. 't Hooft and M. J. G. Veltman, *Nucl. Phys.* B44, 189 (1972).

When two regulators give different answers for observable quantities, it is generally because some symmetry (such as the Ward identity) is being violated by one (or both) of them. In these cases we take the symmetry to be fundamental and demand that it be preserved by the regulator. But the validity of this choice cannot be proven; we are adopting the symmetry as a new axiom.

Dimensional Regularization

The idea of dimensional regularization is very simple to state: Compute the Feynman diagram as an analytic function of the dimensionality of space-time, d . For sufficiently small d , any loop-momentum integral will converge and therefore the Ward identity can be proved. The final expression for any observable quantity should have a well-defined limit as $d \rightarrow 4$.

Let us do a practice calculation to see how this technique works. We consider spacetime to have one time dimension and $(d - 1)$ space dimensions. Then we can Wick-rotate Feynman integrals as before, to give integrals over a d -dimensional Euclidean space. A typical example is

$$\int \frac{d^d \ell_E}{(2\pi)^d} \frac{1}{(\ell_E^2 + \Delta)^2} = \int \frac{d\Omega_d}{(2\pi)^d} \cdot \int_0^\infty d\ell_E \frac{\ell_E^{d-1}}{(\ell_E^2 + \Delta)^2}. \quad (7.80)$$

The first factor in (7.80) contains the area of a unit sphere in d dimensions. To compute it, use the following trick:

$$\begin{aligned} (\sqrt{\pi})^d &= \left(\int dx e^{-x^2} \right)^d = \int d^d x \exp\left(-\sum_{i=1}^d x_i^2\right) \\ &= \int d\Omega_d \int_0^\infty dx x^{d-1} e^{-x^2} = \left(\int d\Omega_d \right) \cdot \frac{1}{2} \int_0^\infty d(x^2) (x^2)^{\frac{d}{2}-1} e^{-(x^2)} \\ &= \left(\int d\Omega_d \right) \cdot \frac{1}{2} \Gamma(d/2). \end{aligned}$$

So the area of a d -dimensional unit sphere is

$$\int d\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}. \quad (7.81)$$

This formula reproduces the familiar special cases:

d	$\Gamma(d/2)$	$\int d\Omega_d$
1	$\sqrt{\pi}$	2
2	1	2π
3	$\sqrt{\pi}/2$	4π
4	1	$2\pi^2$

The second factor in (7.80) is

$$\begin{aligned} \int_0^\infty d\ell \frac{\ell^{d-1}}{(\ell^2 + \Delta)^2} &= \frac{1}{2} \int_0^\infty d(\ell^2) \frac{(\ell^2)^{\frac{d}{2}-1}}{(\ell^2 + \Delta)^2} \\ &= \frac{1}{2} \left(\frac{1}{\Delta} \right)^{2-\frac{d}{2}} \int_0^1 dx x^{1-\frac{d}{2}} (1-x)^{\frac{d}{2}-1}, \end{aligned}$$

where we have substituted $x = \Delta/(\ell^2 + \Delta)$ in the second line. Using the definition of the beta function,

$$\int_0^1 dx x^{\alpha-1} (1-x)^{\beta-1} = B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad (7.82)$$

we can easily evaluate the integral over x . The final result for the d -dimensional integral is

$$\int \frac{d^d \ell_E}{(2\pi)^d} \frac{1}{(\ell_E^2 + \Delta)^2} = \frac{1}{(4\pi)^{d/2}} \frac{\Gamma(2-\frac{d}{2})}{\Gamma(2)} \left(\frac{1}{\Delta} \right)^{2-\frac{d}{2}}.$$

Since $\Gamma(z)$ has isolated poles at $z = 0, -1, -2, \dots$, this integral has isolated poles at $d = 4, 6, 8, \dots$. To find the behavior near $d = 4$, define $\epsilon = 4 - d$, and use the approximation[†]

$$\Gamma(2-\frac{d}{2}) = \Gamma(\epsilon/2) = \frac{2}{\epsilon} - \gamma + \mathcal{O}(\epsilon), \quad (7.83)$$

where $\gamma \approx .5772$ is the Euler-Mascheroni constant. (This constant will always cancel in observable quantities.) The integral is then

$$\int \frac{d^d \ell_E}{(2\pi)^d} \frac{1}{(\ell_E^2 + \Delta)^2} \xrightarrow{d \rightarrow 4} \frac{1}{(4\pi)^2} \left(\frac{2}{\epsilon} - \log \Delta - \gamma + \mathcal{O}(\epsilon) \right). \quad (7.84)$$

When we defined this integral with a Pauli-Villars regulator in Eq. (7.18), we found

$$\int \frac{d^4 \ell_E}{(2\pi)^4} \frac{1}{(\ell_E^2 + \Delta)^2} \xrightarrow{\Lambda \rightarrow \infty} \frac{1}{(4\pi)^2} \left(\log \frac{x\Lambda^2}{\Delta} + \mathcal{O}(\Lambda^{-1}) \right).$$

Thus the $1/\epsilon$ pole in dimensional regularization corresponds to a logarithmic divergence in the momentum integral. Note the curious fact that (7.84)

[†]This expansion follows immediately from the infinite product representation

$$\frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n} \right) e^{-z/n}.$$

involves the logarithm of Δ , a dimensionful quantity. The scale of the logarithm is hidden in the $1/\epsilon$ term, and appears explicitly when the divergence is canceled.

You can easily verify the more general integration formulae,

$$\int \frac{d^d \ell_E}{(2\pi)^d} \frac{1}{(\ell_E^2 + \Delta)^n} = \frac{1}{(4\pi)^{d/2}} \frac{\Gamma(n - \frac{d}{2})}{\Gamma(n)} \left(\frac{1}{\Delta}\right)^{n - \frac{d}{2}}; \quad (7.85)$$

$$\int \frac{d^d \ell_E}{(2\pi)^d} \frac{\ell_E^2}{(\ell_E^2 + \Delta)^n} = \frac{1}{(4\pi)^{d/2}} \frac{d}{2} \frac{\Gamma(n - \frac{d}{2} - 1)}{\Gamma(n)} \left(\frac{1}{\Delta}\right)^{n - \frac{d}{2} - 1}. \quad (7.86)$$

In d dimensions, $g^{\mu\nu}$ obeys $g_{\mu\nu}g^{\mu\nu} = d$. Thus, if the numerator of a symmetric integrand contains $\ell^\mu\ell^\nu$, we should replace

$$\ell^\mu\ell^\nu \rightarrow \frac{1}{d} \ell^2 g^{\mu\nu}, \quad (7.87)$$

in analogy with Eq. (6.46). In QED, the Dirac matrices can be manipulated as a set of d matrices satisfying

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \quad \text{tr}[1] = 4. \quad (7.88)$$

In manipulating Eq. (7.78), these rules give the same result as the purely four-dimensional rules. However, in the evaluation of other diagrams, there are additional contributions of order ϵ . In particular, the contraction identities (5.9) are modified in $d = 4 - \epsilon$ to

$$\begin{aligned} \gamma^\mu\gamma^\nu\gamma_\mu &= -(2 - \epsilon)\gamma^\nu \\ \gamma^\mu\gamma^\nu\gamma^\rho\gamma_\mu &= 4g^{\nu\rho} - \epsilon\gamma^\nu\gamma^\rho \\ \gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_\mu &= -2\gamma^\sigma\gamma^\rho\gamma^\nu + \epsilon\gamma^\nu\gamma^\rho\gamma^\sigma. \end{aligned} \quad (7.89)$$

These extra terms can contribute to the final value of the Feynman diagram if they multiply a factor ϵ^{-1} from a divergent integral. In QED at one-loop order, such extra terms appear in the vertex and self-energy diagrams but cancel when these diagrams are combined to compute an observable quantity.

Computation of Π_2 , Continued

Now let us apply these dimensional regularization formulae to the momentum integral in (7.79). The unpleasant terms with ℓ^2 in the numerator give

$$\begin{aligned} \int \frac{d^d \ell_E}{(2\pi)^d} \frac{(-\frac{2}{d} + 1)g^{\mu\nu}\ell_E^2}{(\ell_E^2 + \Delta)^2} &= \frac{-1}{(4\pi)^{d/2}} (1 - \frac{d}{2}) \Gamma(1 - \frac{d}{2}) \left(\frac{1}{\Delta}\right)^{1 - \frac{d}{2}} g^{\mu\nu} \\ &= \frac{1}{(4\pi)^{d/2}} \Gamma(2 - \frac{d}{2}) \left(\frac{1}{\Delta}\right)^{2 - \frac{d}{2}} \cdot (-\Delta g^{\mu\nu}). \end{aligned}$$

We would have expected a pole at $d = 2$, since the quadratic divergence in 4 dimensions becomes a logarithmic divergence in 2 dimensions. But the pole cancels. The Ward identity is working.

Evaluating the remaining terms in (7.79) and using $\Delta = m^2 - x(1-x)q^2$, we obtain

$$\begin{aligned} i\Pi_2^{\mu\nu}(q) &= -4ie^2 \int_0^1 dx \frac{1}{(4\pi)^{d/2}} \frac{\Gamma(2-\frac{d}{2})}{\Delta^{2-d/2}} \\ &\quad \times [g^{\mu\nu}(-m^2 + x(1-x)q^2) + g^{\mu\nu}(m^2 + x(1-x)q^2) - 2x(1-x)q^\mu q^\nu] \\ &= (q^2 g^{\mu\nu} - q^\mu q^\nu) \cdot i\Pi_2(q^2), \end{aligned}$$

where

$$\begin{aligned} \Pi_2(q^2) &= \frac{-8e^2}{(4\pi)^{d/2}} \int_0^1 dx x(1-x) \frac{\Gamma(2-\frac{d}{2})}{\Delta^{2-d/2}} \\ &\xrightarrow{d \rightarrow 4} -\frac{2\alpha}{\pi} \int_0^1 dx x(1-x) \left(\frac{2}{\epsilon} - \log \Delta - \gamma \right) \quad (\epsilon = 4 - d). \end{aligned} \quad (7.90)$$

With dimensional regularization, $\Pi_2^{\mu\nu}(q)$ indeed takes the form required by the Ward identity. But it is still logarithmically divergent.

We can now compute the order- α shift in the electric charge:

$$e^2 - e_0^2 = \delta Z_3 \underset{\mathcal{O}(\alpha)}{=} \Pi_2(0) \approx -\frac{2\alpha}{3\pi\epsilon}.$$

The bare charge is infinitely larger than the observed charge. But this difference is not observable. What can be observed is the q^2 dependence of the effective electric charge (7.77). This quantity depends on the difference

$$\widehat{\Pi}_2(q^2) \equiv \Pi_2(q^2) - \Pi_2(0) = -\frac{2\alpha}{\pi} \int_0^1 dx x(1-x) \log\left(\frac{m^2}{m^2 - x(1-x)q^2}\right), \quad (7.91)$$

which is independent of ϵ in the limit $\epsilon \rightarrow 0$. For the rest of this section we will investigate what physics this expression contains.

Interpretation of Π_2

First consider the analytic structure of $\widehat{\Pi}_2(q^2)$. For $q^2 < 0$, as is the case when the photon propagator is in the t - or u -channel, $\widehat{\Pi}_2(q^2)$ is manifestly real and analytic. But for an s -channel process, q^2 will be positive. The logarithm function has a branch cut when its argument becomes negative, that is, when

$$m^2 - x(1-x)q^2 < 0.$$

The product $x(1-x)$ is at most $1/4$, so $\widehat{\Pi}_2(q^2)$ has a branch cut beginning at

$$q^2 = 4m^2,$$

at the threshold for creation of a real electron-positron pair.

Let us calculate the imaginary part of $\widehat{\Pi}_2$ for $q^2 > 4m^2$. For any fixed q^2 , the x -values that contribute are between the points $x = \frac{1}{2} \pm \frac{1}{2}\beta$, where $\beta = \sqrt{1 - 4m^2/q^2}$. Since $\text{Im}[\log(-X \pm i\epsilon)] = \pm\pi$, we have

$$\begin{aligned} \text{Im}[\widehat{\Pi}_2(q^2 \pm i\epsilon)] &= -\frac{2\alpha}{\pi} (\pm\pi) \int_{\frac{1}{2}-\frac{1}{2}\beta}^{\frac{1}{2}+\frac{1}{2}\beta} dx x(1-x) \\ &= \mp 2\alpha \int_{-\beta/2}^{\beta/2} dy \left(\frac{1}{4} - y^2\right) \quad (y \equiv x - \frac{1}{2}) \\ &= \mp \frac{\alpha}{3} \sqrt{1 - \frac{4m^2}{q^2}} \left(1 + \frac{2m^2}{q^2}\right). \end{aligned} \quad (7.92)$$

This dependence on q^2 is exactly the same as in Eq. (5.13), the cross section for production of a fermion-antifermion pair. That is just what we would expect from the unitarity relation shown in Fig. 7.6(b); the cut through the diagram for forward Bhabha scattering gives the total cross section for $e^+e^- \rightarrow f\bar{f}$. The parameter β is precisely the velocity of the fermions in the center-of-mass frame.

Next let us examine how $\widehat{\Pi}_2(q^2)$ modifies the electromagnetic interaction, as determined by Eq. (7.77). In the nonrelativistic limit it makes sense to compute the potential $V(r)$. For the interaction between unlike charges, we have, in analogy with Eq. (4.126),

$$V(\mathbf{x}) = \int \frac{d^3q}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{x}} \frac{-e^2}{|\mathbf{q}|^2 [1 - \widehat{\Pi}_2(-|\mathbf{q}|^2)]}. \quad (7.93)$$

Expanding $\widehat{\Pi}$ for $|q^2| \ll m^2$, we obtain

$$V(\mathbf{x}) = -\frac{\alpha}{r} - \frac{4\alpha^2}{15m^2} \delta^{(3)}(\mathbf{x}). \quad (7.94)$$

The correction term indicates that the electromagnetic force becomes much stronger at small distances. This effect can be measured in the hydrogen atom, where the energy levels are shifted by

$$\Delta E = \int d^3x |\psi(\mathbf{x})|^2 \cdot \left(-\frac{4\alpha^2}{15m^2} \delta^{(3)}(\mathbf{x})\right) = -\frac{4\alpha^2}{15m^2} |\psi(0)|^2.$$

The wavefunction $\psi(\mathbf{x})$ is nonzero at the origin only for s -wave states. For the $2S$ state, the shift is

$$\Delta E = -\frac{4\alpha^2}{15m^2} \cdot \frac{\alpha^3 m^3}{8\pi} = -\frac{\alpha^5 m}{30\pi} = -1.123 \times 10^{-7} \text{ eV}.$$

This is a (small) part of the Lamb shift splitting listed in Table 6.1.

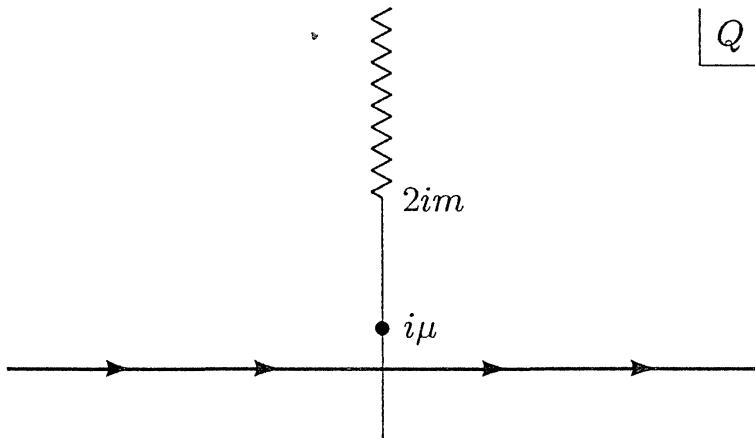


Figure 7.7. Contour for evaluating the effective strength of the electromagnetic interaction in the nonrelativistic limit. The pole at $Q = i\mu$ gives the Coulomb potential. The branch cut gives the order- α correction due to vacuum polarization.

The delta function in Eq. (7.94) is only an approximation; to find the true range of the correction term, we write Eq. (7.93) in the form

$$V(x) = \frac{ie^2}{(2\pi)^2 r} \int_{-\infty}^{\infty} dQ \frac{Q e^{iQr}}{Q^2 + \mu^2} [1 + \hat{\Pi}_2(-Q^2)] \quad (Q \equiv |\mathbf{q}|),$$

where we have inserted a photon mass μ to regulate the Coulomb potential. To perform this integral we push the contour upward (see Fig. 7.7). The leading contribution comes from the pole at $Q = i\mu$, giving the Coulomb potential, $-\alpha/r$. But there is an additional contribution from the branch cut, which begins at $Q = 2mi$. The real part of the integrand is the same on both sides of the cut, so the only contribution to the integral comes from the imaginary part of $\hat{\Pi}_2$. Defining $q = -iQ$, we find that the contribution from the cut is

$$\begin{aligned} \delta V(r) &= \frac{-e^2}{(2\pi)^2 r} \cdot 2 \int_{2m}^{\infty} dq \frac{e^{-qr}}{q} \text{Im}[\hat{\Pi}_2(q^2 - i\epsilon)] \\ &= -\frac{\alpha}{r} \frac{2}{\pi} \int_{2m}^{\infty} dq \frac{e^{-qr}}{q} \frac{\alpha}{3} \sqrt{1 - \frac{4m^2}{q^2}} \left(1 + \frac{2m^2}{q^2}\right). \end{aligned}$$

When $r \gg 1/m$, this integral is dominated by the region where $q \approx 2m$. Approximating the integrand in this region and substituting $t = q - 2m$, we find

$$\begin{aligned} \delta V(r) &= -\frac{\alpha}{r} \cdot \frac{2}{\pi} \int_0^{\infty} dt \frac{e^{-(t+2m)r}}{2m} \frac{\alpha}{3} \sqrt{\frac{t}{m}} \left(\frac{3}{2}\right) + \mathcal{O}(t) \\ &\approx -\frac{\alpha}{r} \cdot \frac{\alpha}{4\sqrt{\pi}} \frac{e^{-2mr}}{(mr)^{3/2}}, \end{aligned}$$

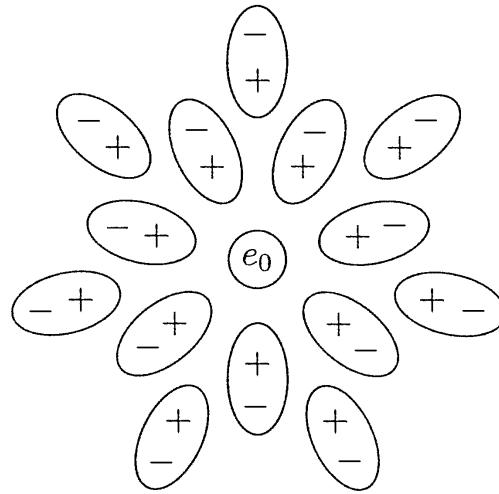


Figure 7.8. Virtual e^+e^- pairs are effectively dipoles of length $\sim 1/m$, which screen the bare charge of the electron.

so that

$$V(r) = -\frac{\alpha}{r} \left(1 + \frac{\alpha}{4\sqrt{\pi}} \frac{e^{-2mr}}{(mr)^{3/2}} + \dots \right). \quad (7.95)$$

Thus the range of the correction term is roughly the electron Compton wavelength, $1/m$. Since hydrogen wavefunctions are nearly constant on this scale, the delta function in Eq. (7.94) was a good approximation. The radiative correction to $V(r)$ is called the *Uehling potential*.

We can interpret the correction as being due to screening. At $r \gtrsim 1/m$, virtual e^+e^- pairs make the vacuum a dielectric medium in which the apparent charge is less than the true charge (see Fig. 7.8). At smaller distances we begin to penetrate the polarization cloud and see the bare charge. This phenomenon is known as *vacuum polarization*.

Now consider the opposite limit: small distance or $-q^2 \gg m^2$. Equation (7.91) then becomes

$$\begin{aligned} \hat{\Pi}_2(q^2) &\approx \frac{2\alpha}{\pi} \int_0^1 dx x(1-x) \left[\log\left(\frac{-q^2}{m^2}\right) + \log(x(1-x)) + \mathcal{O}\left(\frac{m^2}{q^2}\right) \right] \\ &= \frac{\alpha}{3\pi} \left[\log\left(\frac{-q^2}{m^2}\right) - \frac{5}{3} + \mathcal{O}\left(\frac{m^2}{q^2}\right) \right]. \end{aligned}$$

The effective coupling constant in this limit is therefore

$$\alpha_{\text{eff}}(q^2) = \frac{\alpha}{1 - \frac{\alpha}{3\pi} \log\left(\frac{-q^2}{Am^2}\right)}, \quad (7.96)$$

where $A = \exp(5/3)$. The effective electric charge becomes much larger at small distances, as we penetrate the screening cloud of virtual electron-positron pairs.

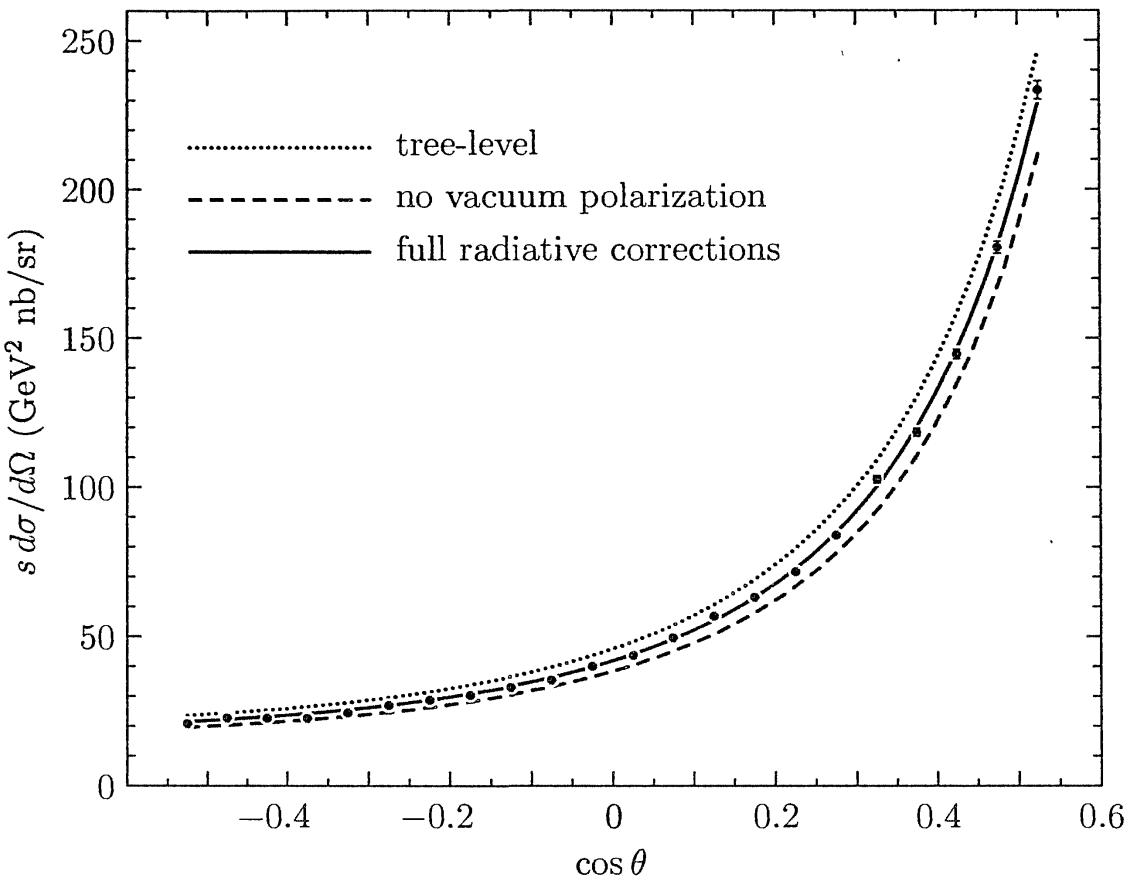


Figure 7.9. Differential cross section for Bhabha scattering, $e^+e^- \rightarrow e^+e^-$, at $E_{\text{cm}} = 29$ GeV, as measured by the HRS collaboration, M. Derrick, et. al., *Phys. Rev. D34*, 3286 (1986). The upper curve is the order- α^2 prediction derived in Problem 5.2, plus a very small ($\sim 2\%$) correction due to the weak interaction. The lower curve includes all QED radiative corrections to order α^3 *except* the vacuum polarization contribution; note that these corrections depend on the experimental conditions, as explained in Chapter 6. The middle curve includes the vacuum polarization contribution as well, which increases the effective value of α^2 by about 10% at this energy.

The combined vacuum polarization effect of the electron and of heavier quarks and leptons causes the value of $\alpha_{\text{eff}}(q^2)$ to increase by about 5% from $q = 0$ to $q = 30$ GeV, and this effect is observed in high-energy experiments. Figure 7.9 shows the cross section for Bhabha scattering at $E_{\text{cm}} = 29$ GeV, and a comparison to QED with and without the vacuum polarization diagram.

We can write α_{eff} as a function of r by setting $q = 1/r$. The behavior of $\alpha_{\text{eff}}(r)$ for all r is sketched in Fig. 7.10. The idea of a distance-dependent (or “scale-dependent” or “running”) coupling constant will be a major theme of the rest of this book.

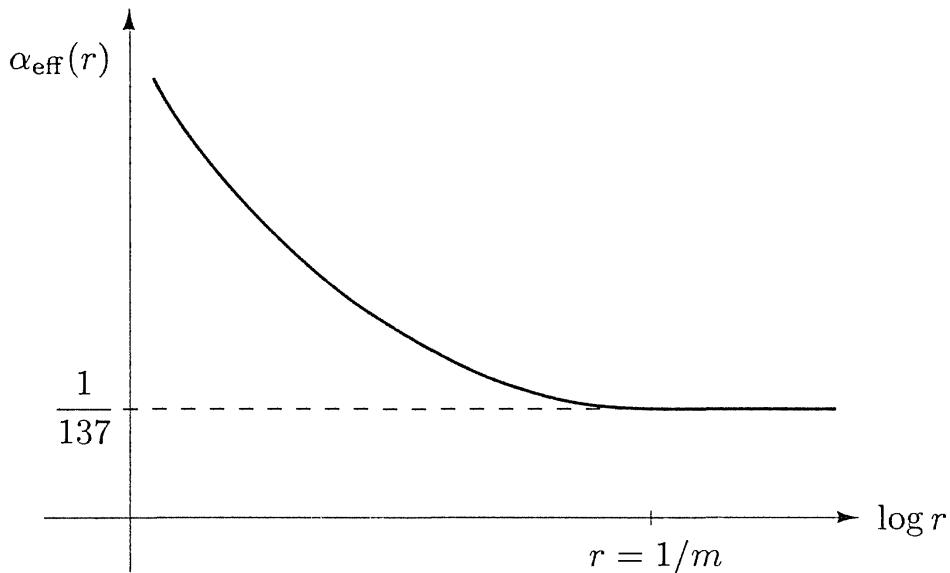


Figure 7.10. A qualitative sketch of the effective electromagnetic coupling constant generated by the one-loop vacuum polarization diagram, as a function of distance. The horizontal scale covers many orders of magnitude.

Problems

7.1 In Section 7.3 we used an indirect method to analyze the one-loop *s*-channel diagram for boson-boson scattering in ϕ^4 theory. To verify our indirect analysis, evaluate all three one-loop diagrams, using the standard method of Feynman parameters. Check the validity of the optical theorem.

7.2 Alternative regulators in QED. In Section 7.5, we saw that the Ward identity can be violated by an improperly chosen regulator. Let us check the validity of the identity $Z_1 = Z_2$, to order α , for several choices of the regulator. We have already verified that the relation holds for Pauli-Villars regularization.

- (a) Recompute δZ_1 and δZ_2 , defining the integrals (6.49) and (6.50) by simply placing an upper limit Λ on the integration over ℓ_E . Show that, with this definition, $\delta Z_1 \neq \delta Z_2$.
- (b) Recompute δZ_1 and δZ_2 , defining the integrals (6.49) and (6.50) by dimensional regularization. You may take the Dirac matrices to be 4×4 as usual, but note that, in d dimensions,

$$g_{\mu\nu}\gamma^\mu\gamma^\nu = d.$$

Show that, with this definition, $\delta Z_1 = \delta Z_2$.

7.3 Consider a theory of elementary fermions that couple both to QED and to a Yukawa field ϕ :

$$H_{\text{int}} = \int d^3x \frac{\lambda}{\sqrt{2}} \phi \bar{\psi} \psi + \int d^3x e A_\mu \bar{\psi} \gamma^\mu \psi.$$

- (a) Verify that the contribution to Z_1 from the vertex diagram with a virtual ϕ equals the contribution to Z_2 from the diagram with a virtual ϕ . Use dimensional regularization. Is the Ward identity generally true in this theory?

- (b) Now consider the renormalization of the $\phi\bar{\psi}\psi$ vertex. Show that the rescaling of this vertex at $q^2 = 0$ is *not* canceled by the correction to Z_2 . (It suffices to compute the ultraviolet-divergent parts of the diagrams.) In this theory, the vertex and field-strength rescaling give additional shifts of the observable coupling constant relative to its bare value.

Radiation of Gluon Jets

Although we have discussed QED radiative corrections at length in the last two chapters, so far we have made no attempt to compute a full radiatively corrected cross section. The reason is of course that such calculations are quite lengthy. Nevertheless it would be dishonest to pretend that one understands radiative corrections after computing only isolated effects as we have done. This “final project” is an attempt to remedy this situation. The project is the computation of one of the simplest, but most important, radiatively corrected cross sections. You should finish Chapter 6 before starting this project, but you need not have read Chapter 7.

Strongly-interacting particles—pions, kaons, and protons—are produced in e^+e^- annihilation when the virtual photon creates a pair of quarks. If one ignores the effects of the strong interactions, it is easy to calculate the total cross section for quark pair production. In this final project, we will analyze the first corrections to this formula due to the strong interactions.

Let us represent the strong interactions by the following simple model: Introduce a new massless vector particle, the *gluon*, which couples universally to quarks:

$$\Delta H = \int d^3x g \bar{\psi}_{fi} \gamma^\mu \psi_{fi} B_\mu.$$

Here f labels the type (“flavor”) of the quark (u , d , s , c , etc.) and $i = 1, 2, 3$ labels the color. The strong coupling constant g is independent of flavor and color. The electromagnetic coupling of quarks depends on the flavor, since the u and c quarks have charge $Q_f = +2/3$ while the d and s quarks have charge $Q_f = -1/3$. By analogy to α , let us define

$$\alpha_g = \frac{g^2}{4\pi}.$$

In this exercise, we will compute the radiative corrections to quark pair production proportional to α_g .

This model of the strong interactions of quarks does not quite agree with the currently accepted theory of the strong interactions, quantum chromodynamics (QCD). However, all of the results that we will derive here are also

correct in QCD with the replacement

$$\alpha_g \rightarrow \frac{4}{3}\alpha_s.$$

We will verify this claim in Chapter 17.

Throughout this exercise, you may ignore the masses of quarks. You may also ignore the mass of the electron, and average over electron and positron polarizations. To control infrared divergences, it will be necessary to assume that the gluons have a small nonzero mass μ , which can be taken to zero only at the end of the calculation. However (as we discussed in Problem 5.5), it is consistent to sum over polarization states of this massive boson by the replacement:

$$\sum \epsilon^\mu \epsilon^{\nu*} \rightarrow -g^{\mu\nu};$$

this also implies that we may use the propagator

$$\overline{B^\mu} B^\nu = \frac{-ig^{\mu\nu}}{k^2 - \mu^2 + i\epsilon}.$$

- (a) Recall from Section 5.1 that, to lowest order in α and neglecting the effects of gluons, the total cross section for production of a pair of quarks of flavor f is

$$\sigma(e^+e^- \rightarrow \bar{q}q) = \frac{4\pi\alpha^2}{3s} \cdot 3Q_f^2.$$

Compute the diagram contributing to $e^+e^- \rightarrow \bar{q}q$ involving one virtual gluon. Reduce this expression to an integral over Feynman parameters, and renormalize it by subtraction at $q^2 = 0$, following the prescription used in Eq. (6.55). Notice that the resulting expression can be considered as a correction to $F_1(q^2)$ for the quark. Argue that, for massless quarks, to all orders in α_g , the total cross section for production of a quark pair unaccompanied by gluons is

$$\sigma(e^+e^- \rightarrow \bar{q}q) = \frac{4\pi\alpha^2}{3s} \cdot 3|F_1(q^2 = s)|^2,$$

with $F_1(q^2 = 0) = Q_f$.

- (b) Before we attempt to evaluate the Feynman parameter integrals in part (a), let us put this contribution aside and study the process $e^+e^- \rightarrow \bar{q}qg$, quark pair production with an additional gluon emitted. Before we compute the cross section, it will be useful to work out some kinematics. Let q be the total 4-momentum of the reaction, let k_1 and k_2 be the 4-momenta of the final quark and antiquark, and let k_3 be the 4-momentum of the gluon. Define

$$x_i = \frac{2k_i \cdot q}{q^2}, \quad i = 1, 2, 3;$$

this is the ratio of the center-of-mass energy of particle i to the maximum available energy. Then show (i) $\sum x_i = 2$, (ii) all other Lorentz scalars involving only the final-state momenta can be computed in terms of the x_i and the particle masses, and (iii) the complete integral over 3-body phase space can be written as

$$\int d\Pi_3 = \prod_i \int \frac{d^3 k_i}{(2\pi)^3} \frac{1}{2E_i} (2\pi)^4 \delta^{(4)}(q - \sum_i k_i) = \frac{q^2}{128\pi^3} \int dx_1 dx_2.$$

Find the region of integration for x_1 and x_2 if the quark and antiquark are massless but the gluon has mass μ .

- (c) Draw the Feynman diagrams for the process $e^+e^- \rightarrow \bar{q}qg$, to leading order in α and α_g , and compute the differential cross section. You may throw away the information concerning the correlation between the initial beam axis and the directions of the final particles. This is conveniently done as follows: The usual trace tricks for evaluating the square of the matrix element give for this process a result of the structure

$$\int d\Pi_3 \frac{1}{4} \sum |\mathcal{M}|^2 = L_{\mu\nu} \int d\Pi_3 H^{\mu\nu},$$

where $L_{\mu\nu}$ represents the electron trace and $H^{\mu\nu}$ represents the quark trace. If we integrate over all parameters of the final state except x_1 and x_2 , which are scalars, the only preferred 4-vector characterizing the final state is q^μ . On the other hand, $H_{\mu\nu}$ satisfies

$$q^\mu H_{\mu\nu} = H_{\mu\nu} q^\nu = 0.$$

Why is this true? (There is an argument based on general principles; however, you might find it a useful check on your calculation to verify this property explicitly.) Since, after integrating over final-state vectors, $\int H^{\mu\nu}$ depends only on q^μ and scalars, it can only have the form

$$\int d\Pi_3 H^{\mu\nu} = \left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2} \right) \cdot H,$$

where H is a scalar. With this information, show that

$$L_{\mu\nu} \int d\Pi_3 H^{\mu\nu} = \frac{1}{3} (g^{\mu\nu} L_{\mu\nu}) \cdot \int d\Pi_3 (g^{\rho\sigma} H_{\rho\sigma}).$$

Using this trick, derive the differential cross section

$$\frac{d\sigma}{dx_1 dx_2} (e^+e^- \rightarrow \bar{q}qg) = \frac{4\pi\alpha^2}{3s} \cdot 3Q_f^2 \cdot \frac{\alpha_g}{2\pi} \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}$$

in the limit $\mu \rightarrow 0$. If we assume that each original final-state particle is realized physically as a jet of strongly interacting particles, this formula gives the probability for observing three-jet events in e^+e^- annihilation and the kinematic distribution of these events. The form of the distribution in the x_i is an absolute prediction, and it agrees with experiment. The

normalization of this distribution of a measure of the strong-interaction coupling constant.

- (d) Now replace $\mu \neq 0$ in the formula of part (c) for the differential cross section, and carefully integrate over the region found in part (b). You may assume $\mu^2 \ll q^2$. In this limit, you will find infrared-divergent terms of order $\log(q^2/\mu^2)$ and also $\log^2(q^2/\mu^2)$, finite terms of order 1, and terms explicitly suppressed by powers of (μ^2/q^2) . You may drop terms of the last type throughout this calculation. For the moment, collect and evaluate only the infrared-divergent terms.
- (e) Now analyze the Feynman parameter integral obtained in part (a), again working in the limit $\mu^2 \ll q^2$. Note that this integral has singularities in the region of integration. These should be controlled by evaluating the integral for q spacelike and then analytically continuing into the physical region. That is, write $Q^2 = -q^2$, evaluate the integral for $Q^2 > 0$, and then carefully analytically continue the result to $Q^2 = -q^2 - i\epsilon$. Combine the result with the answer from part (d) to form the total cross section for $e^+e^- \rightarrow$ strongly interacting particles, to order α_g . Show that all infrared-divergent logarithms cancel out of this quantity, so that this total cross section is well-defined in the limit $\mu \rightarrow 0$.
- (f) Finally, collect the terms of order 1 from the integrations of parts (d) and (e) and combine them. To evaluate certain of these terms, you may find the following formula useful:

$$\int_0^1 dx \frac{\log(1-x)}{x} = -\frac{\pi^2}{6}.$$

(It is not hard to prove this.) Show that the total cross section is given, to this order in α_g , by

$$\sigma(e^+e^- \rightarrow \bar{q}q \text{ or } \bar{q}qg) = \frac{4\pi\alpha}{3s} \cdot 3Q_f^2 \cdot \left(1 + \frac{3\alpha_g}{4\pi} + \dots\right).$$

This formula gives a second way of measuring the strong-interaction coupling constant. The experimental results agree (within the current experimental errors) with the results obtained by the method of part (c). We will discuss the measurement of α_s more fully in Section 17.6.

Part II

Renormalization

Invitation: Ultraviolet Cutoffs and Critical Fluctuations

The main purpose of Part II of this book is to develop a general theory of renormalization. This theory will explain the origin of ultraviolet divergences in field theory and will indicate when these divergences can be removed systematically. It will also give a way to convert the divergences of Feynman diagrams from a problem into a tool. We will apply this tool to study the asymptotic large- or small-momentum behavior of field theory amplitudes.

When we first encountered an ultraviolet divergence in the calculation of the one-loop vertex correction in Section 6.3, it seemed an aberration that ought to disappear before it caused us too much discomfort. In Chapter 7 we saw further examples of ultraviolet-divergent diagrams, enough to convince us that such divergences occur ubiquitously in Feynman diagram computations. Thus it is necessary for anyone studying field theory to develop a point of view toward these divergences. Most people begin with the belief that any theory that contains divergences must be nonsense. But this viewpoint is overly restrictive, since it excludes not only quantum field theory but even the classical electrodynamics of point particles.

With some experience, one might adopt a more permissive attitude of peaceful coexistence with the divergences: One can accept a theory with divergences, as long as they do not appear in physical predictions. In Chapter 7 we saw that all of the divergences that appear in the one-loop radiative corrections to electron scattering from a heavy target can be eliminated by consistently eliminating the bare values of the mass and charge of the electron in favor of their measured physical values. In Chapter 10, we will argue that all of the ultraviolet divergences of QED, in all orders of perturbation theory, can be eliminated in this way. Thus, as long as one is willing to consider the mass and charge of the electron as measured parameters, the predictions of QED perturbation theory will always be free of divergences. We will also show in Chapter 10 that QED belongs to a well-defined class of field theories in which all ultraviolet divergences are removed after a fixed small number of physical parameters are taken from experiment. These theories, called *renormalizable* quantum field theories, are the only ones in which perturbation theory gives well-defined predictions.

Ideally, though, one should take the further step of trying to understand

physically why the divergences appear and why their effects are more severe in some theories than in others. This direct approach to the divergence problem was pioneered in the 1960s by Kenneth Wilson. The crucial insights needed to solve this problem emerged from a correspondence, discovered by Wilson and others, between quantum field theory and the statistical physics of magnets and fluids. Wilson's approach to renormalization is the subject of Chapter 12. The present chapter gives a brief introduction to the issues in condensed matter physics that have provided insight into the problem of ultraviolet divergences.

Formal and Physical Cutoffs

Ultraviolet divergences signal that quantities calculated in a quantum field theory depend on some very large momentum scale, the ultraviolet cutoff. Equivalently, in position space, divergent quantities depend on some very small distance scale.

The idea of a small-distance cutoff in the continuum description of a system occurs in classical field theories as well. Typically the cutoff is at the scale of atomic distances, where the continuum description no longer applies. However, the size of the cutoff manifests itself in certain parameters of the continuum theory. In fluid dynamics, for instance, parameters such as the viscosity and the speed of sound are of just the size one would expect by combining typical atomic radii and velocities. Similarly, in a magnet, the magnetic susceptibility can be estimated by assuming that the energy cost of flipping an electron spin is on the order of a tenth of an eV, as we would expect from atomic physics. Each of these systems possesses a natural ultraviolet cutoff at the scale of an atom; by understanding the physics at the atomic scale, we can compute the parameters that determine the physics on larger scales.

In quantum field theory, however, we have no precise knowledge of the fundamental physics at very short distance scales. Thus, we can only measure parameters such as the physical charge and mass of the electron, not compute them from first principles. The presence of ultraviolet divergences in the relations between these physical parameters and their bare values is a sign that these parameters are controlled by the unknown short-distance physics.

Whether we know the fundamental physics at small distance scales or not, we need two kinds of information in order to write an effective theory for large-distance phenomena. First, we must know how many parameters from the small distance scale are relevant to large-distance physics. Second, and more importantly, we must know what degrees of freedom from the underlying theory appear at large distances.

In fluid mechanics, it is something of a miracle, from the atomic point of view, that any large-distance degrees of freedom even exist. Nevertheless, the equations that express the transport of energy and mass over large distances do have smooth, coherent solutions. The large-distance degrees of freedom are

the flows that transport these conserved quantities, and sound waves of long wavelength.

In quantum field theory, the large-distance physics involves only those particles that have masses that are very small compared to the fundamental cutoff scale. These particles and their dynamics are the quantum analogues of the large-scale flows in fluid mechanics. The simplest way to naturally arrange for such particles to appear is to make use of particles that naturally have zero mass. So far in this book, we have encountered two types of particles whose mass is precisely zero, the photon and the chiral fermion. (In Chapter 11 we will meet one further naturally massless particle, the *Goldstone boson*.) We might argue that QED exists as a theory on scales much larger than its cutoff because the photon is naturally massless and because the left- and right-handed electrons are very close to being chiral fermions.

There is another way that particles of zero or almost zero mass can arise in quantum field theory: We can simply tune the parameters of a scalar field theory so that the scalar particles have masses small compared to the cutoff. This method of introducing particles with small mass seems arbitrary and unnatural. Nevertheless, it has an analogue in statistical mechanics that is genuinely interesting in that discipline and can teach us some important lessons.

Normally, in a condensed matter system, the thermal fluctuations are correlated only over atomic distances. Under special circumstances, however, they can have much longer range. The clearest example of this phenomenon occurs in a ferromagnet. At high temperature, the electron spins in a magnet are disorganized and fluctuating; but at low temperature, these spins align to a fixed direction.* Let us think about how this alignment builds up as the temperature of the magnet is lowered. As the magnet cools from high temperature, clusters of correlated spins become larger and larger. At a certain point—the temperature of magnetization—the entire sample becomes a single large cluster with a well-defined macroscopic orientation. Just above this temperature, the magnet contains large clusters of spins with a common orientation, which in turn belong to still larger clusters, such that the orientations on the very largest scale are still randomized through the sample. This situation is illustrated in Fig. 8.1. Similar behavior occurs in the vicinity of any other second-order phase transition, for example, the order-disorder transition in binary alloys, the critical point in fluids, or the superfluid transition in Helium-4.

The natural description of these very long wavelength fluctuations is in terms of a fluctuating continuum field. At the lowest intuitive level, we might

*In a real ferromagnet, the long-range magnetic dipole-dipole interaction causes the state of uniform magnetization to break up into an array of magnetic domains. In this book, we will ignore this interaction and think of a magnetic spin as a pure orientation. It is this idealized system that is directly analogous to a quantum field theory.

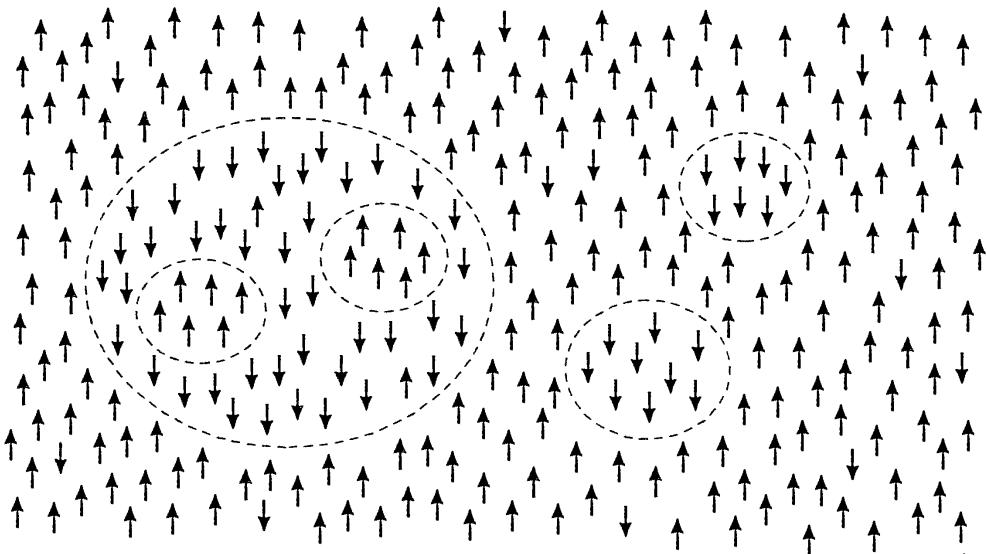


Figure 8.1. Clusters of oriented spins near the critical point of a ferromagnet.

substitute quantum for statistical fluctuations and try to describe this system as a quantum field theory. In Section 9.3 we will derive a somewhat more subtle relation that makes a precise connection between the statistical and the quantum systems. Through this connection, the behavior of any statistical system near a second-order phase transition can be translated into the behavior of a particular quantum field theory. This quantum field theory has a field with a mass that is very small compared to the basic atomic scale and that goes to zero precisely at the phase transition.

But this connection seems to compound the problem of ultraviolet divergences in quantum field theory: If the wealth of phase transitions observed in Nature generates a similar wealth of quantum field theories, how can we possibly define a quantum field theory without detailed reference to its origins in physics at the scale of its ultraviolet cutoff? Saying that a quantum field theory makes predictions independent of the cutoff would be equivalent to saying that the statistical fluctuations in the neighborhood of a critical point are independent of whether the system is a magnet, a fluid, or an alloy. But is this statement so obviously incorrect? By reversing the logic, we would find that quantum field theory makes a remarkably powerful prediction for condensed matter systems, a prediction of *universality* for the statistical fluctuations near a critical point. In fact, this prediction is verified experimentally.

A major theme of Part II of this book will be that these two ideas—cutoff independence in quantum field theory and universality in the theory of critical phenomena—are naturally the same idea, and that understanding either of these ideas gives insight into the other.

Landau Theory of Phase Transitions

To obtain a first notion of what could be universal in the phenomena of phase transitions, let us examine the simplest continuum theory of second-order phase transitions, due to Landau.

First we should review a little thermodynamics and clarify our nomenclature. In thermodynamics, a *first-order phase transition* is a point across which some thermodynamic variable (the density of a fluid, or the magnetization of a ferromagnet) changes discontinuously. At a phase transition point, two quite distinct thermodynamic states (liquid and gas, or magnetization parallel and antiparallel to a given axis) are in equilibrium. The thermodynamic quantity that changes discontinuously across the transition, and that characterizes the difference of the two competing phases, is called the *order parameter*. In most circumstances, it is possible to change a second thermodynamic parameter in such a way that the two competing states move closer together in the thermodynamic space, so that at some value of this parameter, these two states become identical and the discontinuity in the order parameter disappears. This endpoint of the line of first-order transitions is called a *second-order phase transition*, or, more properly, a *critical point*. Viewed from the other direction, a critical point is a point at which a single thermodynamic state bifurcates into two macroscopically distinct states. It is this bifurcation that leads to the long-ranged thermal fluctuations discussed in the previous section.

A concrete example of this behavior is exhibited by a ferromagnet. Let us assume for simplicity that the material we are discussing has a preferred axis of magnetization, so that at low temperature, the system will have its spins ordered either parallel or antiparallel to this axis. The total magnetization along this axis, M , is the order parameter. At low temperature, application of an external magnetic field H will favor one or the other of the two possible states. At $H = 0$, the two states will be in equilibrium; if H is changed from a small negative to a small positive value, the thermodynamic state and the value of M will change discontinuously. Thus, for any fixed (low) temperature, there is a first-order transition at $H = 0$. Now consider the effect of raising the temperature: The fluctuation of the spins increases and the value of $|M|$ decreases. At some temperature T_C the system ceases to be magnetized at $H = 0$. At this point, the first-order phase transition disappears and the two competing thermodynamic states coalesce. The system thus has a critical point at $T = T_C$. The location of these various transitions in the H - T plane is shown in Fig. 8.2.

Landau described this behavior by the use of the Gibbs free energy G ; this is the thermodynamic potential that depends on M and T , such that

$$\left. \frac{\partial G}{\partial M} \right|_T = -H. \quad (8.1)$$

He suggested that we concentrate our attention on the region of the critical

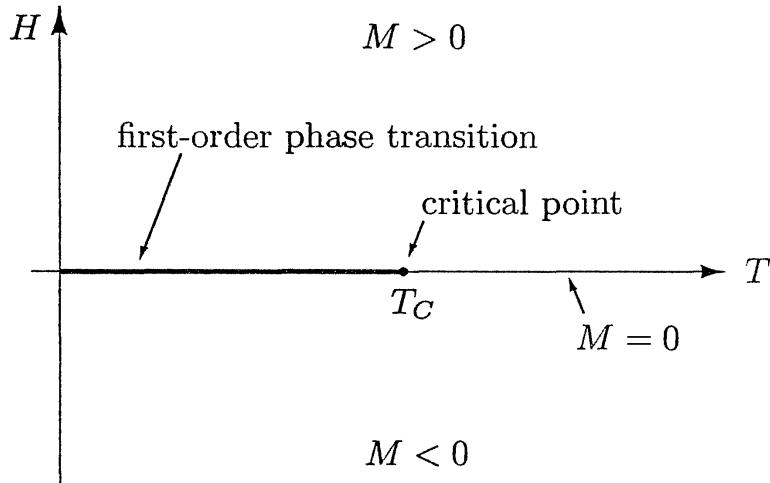


Figure 8.2. Phase diagram in the H - T plane for a uniaxial ferromagnet.

point: $T \approx T_C$, $M \approx 0$. Then it is reasonable to expand $G(M)$ as a Taylor series in M . For $H = 0$, we can write

$$G(M) = A(T) + B(T)M^2 + C(T)M^4 + \dots \quad (8.2)$$

Because the system has a symmetry under $M \rightarrow -M$, $G(M)$ can contain only even powers of M . Since M is small, we will ignore the higher terms in the expansion. Given Eq. (8.2), we can find the possible values of M at $H = 0$ by solving

$$0 = \frac{\partial G}{\partial M} = 2B(T)M + 4C(T)M^3. \quad (8.3)$$

If B and C are positive, the only solution is $M = 0$. However, if $C > 0$ but B is negative below some temperature T_C , we have a nontrivial solution for $T < T_C$, as shown in Fig. 8.3. More concretely, approximate for $T \approx T_C$:

$$B(T) = b(T - T_C), \quad C(T) = c. \quad (8.4)$$

Then the solution to Eq. (8.3) is

$$M = \begin{cases} 0 & \text{for } T > T_C; \\ \pm [(b/2c)(T_C - T)]^{1/2} & \text{for } T < T_C. \end{cases} \quad (8.5)$$

This is just the qualitative behavior that we expect at a critical point.

To find the value of M at nonzero external field, we could solve Eq. (8.1) with the left-hand side given by (8.2). An equivalent procedure is to minimize a new function, related to (8.2). Define

$$G(M, H) = A(T) + B(T)M^2 + C(T)M^4 - HM. \quad (8.6)$$

Then the minimum of $G(M, H)$ with respect to M at fixed H gives the value of M that satisfies Eq. (8.1). The minimum is unique except when $H = 0$ and $T < T_C$, where we find the double minimum in the second line of (8.5). This is consistent with the phase diagram shown in Fig. 8.2.

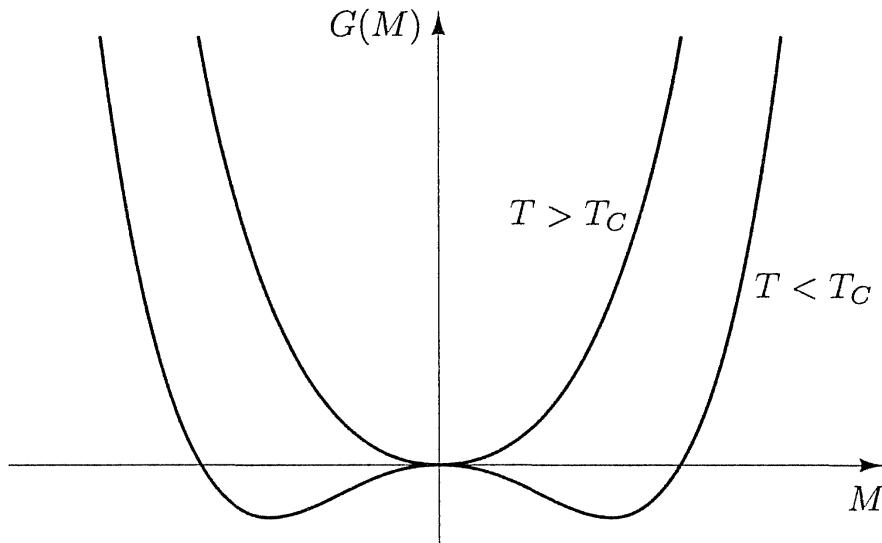


Figure 8.3. Behavior of the Gibbs free energy $G(M)$ in Landau theory, at temperatures above and below the critical temperature.

To study correlations in the vicinity of the phase transition, Landau generalized this description further by considering the magnetization M to be the integral of a local spin density:

$$M = \int d^3x s(\mathbf{x}). \quad (8.7)$$

Then the Gibbs free energy (8.6) becomes the integral of a local function of $s(\mathbf{x})$,

$$G = \int d^3x \left[\frac{1}{2}(\nabla s)^2 + b(T - T_C)s^2 + cs^4 - Hs \right], \quad (8.8)$$

which must be minimized with respect to the field configuration $s(\mathbf{x})$. The first term is the simplest possible way to introduce the tendency of nearby spins to align with one another. We have rescaled $s(\mathbf{x})$ so that the coefficient of this term is set to $1/2$. In writing this free energy integral, we could even consider H to vary as a function of position. In fact, it is useful to do that; we can turn on $H(\mathbf{x})$ near $x = 0$ and see what response we find at another point.

The minimum of the free energy expression (8.8) with respect to $s(\mathbf{x})$ is given by the solution to the variational equation

$$0 = \delta G[s(\mathbf{x})] = -\nabla^2 s + 2b(T - T_C)s + 4cs^3 - H(\mathbf{x}). \quad (8.9)$$

For $T > T_C$, where the macroscopic magnetization vanishes and so $s(\mathbf{x})$ should be small, we can find the qualitative behavior by ignoring the s^3 term. Then $s(\mathbf{x})$ obeys a linear equation,

$$(-\nabla^2 + 2b(T - T_C))s(\mathbf{x}) = H(\mathbf{x}). \quad (8.10)$$

To study correlations of spins, we will set

$$H(\mathbf{x}) = H_0\delta^{(3)}(\mathbf{x}). \quad (8.11)$$

The resulting configuration $s(\mathbf{x})$ is then the Green's function of the differential operator in Eq. (8.10), so we call it $D(\mathbf{x})$:

$$(-\nabla^2 + 2b(T - T_C))D(\mathbf{x}) = H_0\delta^{(3)}(\mathbf{x}). \quad (8.12)$$

This Green's function tells us the response at \mathbf{x} when the spin at $\mathbf{x} = 0$ is forced into alignment with H . In Sections 9.2 and 9.3 we will see that $D(\mathbf{x})$ is also proportional to the zero-field spin-spin correlation function in the thermal ensemble,

$$D(\mathbf{x}) \propto \langle s(\mathbf{x})s(0) \rangle \equiv \sum_{\text{all } s(\mathbf{x})} s(\mathbf{x})s(0)e^{-\mathbf{H}/kT}, \quad (8.13)$$

where \mathbf{H} is the Hamiltonian of the magnetic system.

The solution to Eq. (8.12) can be found by Fourier transformation:

$$D(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} \frac{H_0 e^{i\mathbf{k}\cdot\mathbf{x}}}{|\mathbf{k}|^2 + 2b(T - T_C)}. \quad (8.14)$$

This is just the integral we encountered in our discussion of the Yukawa potential, Eq. (4.126). Evaluating it in the same way, we find

$$D(\mathbf{x}) = \frac{H_0}{4\pi} \frac{1}{r} e^{-r/\xi}, \quad (8.15)$$

where

$$\xi = [2b(T - T_C)]^{-1/2} \quad (8.16)$$

is the *correlation length*, the range of correlated spin fluctuations. Notice that this length diverges as $T \rightarrow T_C$.

The main results of this analysis, Eqs. (8.5) and (8.16), involve unknown constants b, c that depend on physics at the atomic scale. On the other hand, the power-law dependence in these formulae on $(T - T_C)$ follows simply from the structure of the Landau equations and is independent of any details of the microscopic physics. In fact, our derivation of this dependence did not even use the fact that G describes a ferromagnet; we assumed only that G can be expanded in powers of an order parameter and that G respects the reflection symmetry $M \rightarrow -M$. These assumptions apply equally well to many other types of systems: binary alloys, superfluids, and even (though the reflection symmetry is less obvious here) the liquid-gas transition. Landau theory predicts that, near the critical point, these systems show a universal behavior in the dependence of M, ξ , and other thermodynamic quantities on $(T - T_C)$.

Critical Exponents

The preceding treatment of the Landau theory of phase transitions emphasizes its similarity to classical field theory. We set up an appropriate free energy and found the thermodynamically preferred configuration by solving a classical variational equation. This gives only an approximation to the full statistical problem, analogous to the approximation of replacing quantum by classical

dynamics in field theory. In Chapter 13, we will use methods of quantum field theory to account properly for the fluctuations about the preferred Landau thermodynamic state. These modifications turn out to be profound, and rather counterintuitive.

To describe the form of these modifications, let us write Eq. (8.15) more generally as

$$\langle s(\mathbf{x})s(0) \rangle = A \frac{1}{r^{1+\eta}} f(r/\xi), \quad (8.17)$$

where A is a constant and $f(y)$ is a function that satisfies $f(0) = 1$ and $f(y) \rightarrow 0$ as $y \rightarrow \infty$. Landau theory predicts that $\eta = 0$ and $f(y)$ is a simple exponential. This expression has a form strongly analogous to that of a Green's function in quantum field theory. The constant A can be absorbed into the field-strength renormalization of the field $s(\mathbf{x})$. The correlation length ξ is, in general, a complicated function of the atomic parameters, but in the continuum description we can simply trade these parameters for ξ . It is appropriate to consider ξ as a cutoff-independent, physical parameter, since it controls the large-distance behavior of a physical correlation. In fact, the analogy between Eq. (8.15) and the Yukawa potential suggests that we should identify ξ^{-1} with the physical mass in the associated quantum field theory. Then Eq. (8.17) gives a cutoff-independent, continuum representation of the statistical system.

If we were working in quantum field theory, we would derive corrections to Eq. (8.17) as a perturbation series in the parameter c multiplying the nonlinear term in (8.9). This would generalize the Landau result to

$$\langle s(\mathbf{x})s(0) \rangle = \frac{1}{r} F(r/\xi, c). \quad (8.18)$$

The perturbative corrections would depend on the properties of the continuum field theory. For example, $F(y, c)$ would depend on the number of components of the field $s(\mathbf{x})$, and its series expansion would differ depending on whether the magnetization formed along a preferred axis, in a preferred plane, or isotropically. For order parameters with many components, the expansion would also depend on higher discrete symmetries of the problem. However, we expect that systems described by the same Landau free energy (for example, a single-axis ferromagnet and a liquid-gas system) should have the same perturbation expansion when this expansion is written in terms of the physical mass and coupling. The complete universality of Landau theory then becomes a more limited concept, in which systems have the same large-distance correlations if their order parameters have the same symmetry. We might say that statistical systems divide into distinct *universality classes*, each with a characteristic large-scale behavior.

If this were the true behavior of systems near second-order phase transitions, it would already be a wonderful confirmation of the ideas required to formulate cutoff-independent quantum field theories. However, the true behavior of statistical systems is still another level more subtle. What one finds experimentally is a dependence of the form of Eq. (8.17), where the function

$F(y)$ is the same within each universality class. There is no need for an auxiliary parameter c . On the other hand, the exponent η takes a specific nonzero value in each universality class. Other power-law relations of Landau theory are also modified, in a specific manner for each universality class. For example, Eq. (8.5) is changed, for $T < T_C$, to

$$M \propto (T_C - T)^\beta, \quad (8.19)$$

where the exponent β takes a fixed value for all systems in a given universality class. For three-dimensional single-axis magnets and for fluids, $\beta = 0.313$. The powers in these nontrivial scaling relations are called *critical exponents*.

The modification from Eq. (8.18) to Eq. (8.17) does not imperil the idea that a condensed matter system, in the vicinity of a second-order phase transition, has a well-defined, cutoff-independent, continuum behavior. However, we would like to understand why Eq. (8.17) should be expected as the correct representation. The answer to this question will come from a thorough analysis of the ultraviolet divergences of the corresponding quantum field theory. In Chapter 12, when we finally conclude our explication of the ultraviolet divergences, we will find that we have in hand the tools not only to justify Eq. (8.17), but also to calculate the values of the critical exponents using Feynman diagrams. In this way, we will uncover a beautiful application of quantum field theory to the domain of atomic physics. The success of this application will guide us, in Part III, to even more powerful tools, which we will need in the relativistic domain of elementary particles.

Functional Methods

Feynman once said that* “every theoretical physicist who is any good knows six or seven different theoretical representations for exactly the same physics.” Following his advice, we introduce in this chapter an alternative method of deriving the Feynman rules for an interacting quantum field theory: the method of *functional integration*.

Aside from Feynman’s general principle, we have several specific reasons for introducing this formalism. It will provide us with a relatively easy derivation of our expression for the photon propagator, completing the proof of the Feynman rules for QED given in Section 4.8. The functional method generalizes more readily to other interacting theories, such as scalar QED (Problem 9.1), and especially the non-Abelian gauge theories (Part III). Since it uses the Lagrangian, rather than the Hamiltonian, as its fundamental quantity, the functional formalism explicitly preserves all symmetries of a theory. Finally, the functional approach reveals the close analogy between quantum field theory and statistical mechanics. Exploiting this analogy, we will turn Feynman’s advice upside down and apply the same theoretical representation to two completely different areas of physics.

9.1 Path Integrals in Quantum Mechanics

We begin by applying the functional integral (or *path integral*) method to the simplest imaginable system: a nonrelativistic quantum-mechanical particle moving in one dimension. The Hamiltonian for this system is

$$H = \frac{p^2}{2m} + V(x).$$

Suppose that we wish to compute the amplitude for this particle to travel from one point (x_a) to another (x_b) in a given time (T). We will call this amplitude $U(x_a, x_b; T)$; it is the position representation of the Schrödinger time-evolution operator. In the canonical Hamiltonian formalism, U is given by

$$U(x_a, x_b; T) = \langle x_b | e^{-iHT/\hbar} | x_a \rangle. \quad (9.1)$$

**The Character of Physical Law* (MIT Press, 1965), p. 168.

(For the next few pages we will display all factors of \hbar explicitly.)

In the path-integral formalism, U is given by a very different-looking expression. We will first try to motivate that expression, then prove that it is equivalent to (9.1).

Recall that in quantum mechanics there is a *superposition principle*: When a process can take place in more than one way, its total amplitude is the coherent sum of the amplitudes for each way. A simple but nontrivial example is the famous double-slit experiment, shown in Fig. 9.1. The total amplitude for an electron to arrive at the detector is the sum of the amplitudes for the two paths shown. Since the paths differ in length, these two amplitudes generally differ, causing interference.

For a general system, we might therefore write the total amplitude for traveling from x_a to x_b as

$$U(x_a, x_b; T) = \sum_{\text{all paths}} e^{i \cdot (\text{phase})} = \int \mathcal{D}x(t) e^{i \cdot (\text{phase})}. \quad (9.2)$$

To be democratic, we have written the amplitude for each particular path as a pure phase, so that no path is inherently more important than any other. The symbol $\int \mathcal{D}x(t)$ is simply another way of writing “sum over all paths”; since there is one path for every function $x(t)$ that begins at x_a and ends at x_b , the sum is actually an integral over this continuous space of functions.

We can define this integral as part of a natural generalization of the calculus to spaces of functions. A function that maps functions to numbers is called a *functional*. The integrand in (9.2) is a functional, since it associates a complex amplitude with any function $x(t)$. The argument of a functional $F[x(t)]$ is conventionally written in square brackets rather than parentheses. Just as an ordinary function $y(x)$ can be integrated over a set of points x , a functional $F[x(t)]$ can be integrated over a set of functions $x(t)$; the measure of such a *functional integral* is conventionally written with a script capital \mathcal{D} , as in (9.2). A functional can also be differentiated with respect to its argument (a function), and this *functional derivative* is denoted by $\delta F / \delta x(t)$. We will develop more precise definitions of this new integral and derivative in the course of this section and the next.

What should we use for the “phase” in Eq. (9.2)? In the classical limit, we should find that only one path, the classical path, contributes to the total amplitude. We might therefore hope to evaluate the integral in (9.2) by the method of stationary phase, identifying the classical path $x_{\text{cl}}(t)$ by the stationary condition,

$$\frac{\delta}{\delta x(t)} \left(\text{phase}[x(t)] \right) \Big|_{x_{\text{cl}}} = 0.$$

But the classical path is the one that satisfies the principle of least action,

$$\frac{\delta}{\delta x(t)} \left(S[x(t)] \right) \Big|_{x_{\text{cl}}} = 0,$$

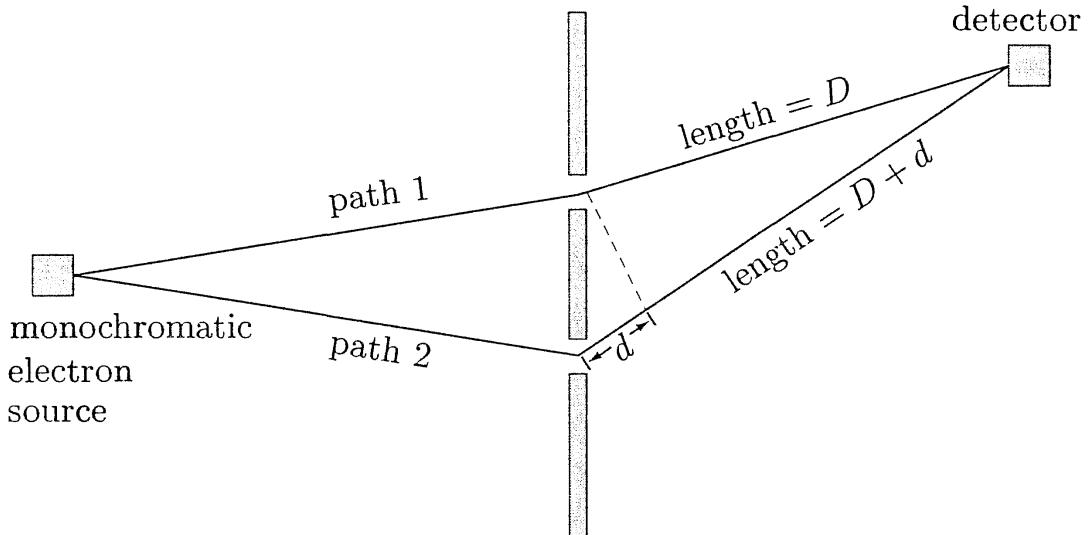


Figure 9.1. The double-slit experiment. Path 2 is longer than path 1 by an amount d , and therefore has a phase that is larger by $2\pi d/\lambda$, where $\lambda = 2\pi\hbar/p$ is the particle's de Broglie wavelength. Constructive interference occurs when $d = 0, \lambda, \dots$, while destructive interference occurs when $d = \lambda/2, 3\lambda/2, \dots$

where $S = \int L dt$ is the classical action. It is tempting, therefore, to identify the phase with S , up to a constant. Since the stationary-phase approximation should be valid in the classical limit—that is, when $S \gg \hbar$ —we will use S/\hbar for the phase. Our final formula for the propagation amplitude is thus

$$\langle x_b | e^{-iHT/\hbar} | x_a \rangle = U(x_a, x_b; T) = \int \mathcal{D}x(t) e^{iS[x(t)]/\hbar}. \quad (9.3)$$

We can easily verify that this formula gives the correct interference pattern in the double-slit experiment. The action for either path shown in Fig. 9.1 is just $(1/2)mv^2t$, the kinetic energy times the time. For path 1 the velocity is $v_1 = D/t$, so the phase is $mD^2/2\hbar t$. For path 2 we have $v_2 = (D+d)/t$, so the phase is $m(D+d)^2/2\hbar t$. We must assume that $d \ll D$, so that $v_1 \approx v_2$ (i.e., the electrons have a well-defined velocity). The excess phase for path 2 is then $mDd/\hbar t \approx pd/\hbar$, where p is the momentum. This is exactly what we would expect from the de Broglie relation $p = h/\lambda$, so we must be doing something right.

To evaluate the functional integral more generally, we must define the symbol $\int \mathcal{D}x(t)$ in the case where the number of paths $x(t)$ is more than two (and, in fact, continuously infinite). We will use a brute-force definition, by discretization. Break up the time interval from 0 to T into many small pieces of duration ϵ , as shown in Fig. 9.2. Approximate a path $x(t)$ as a sequence of straight lines, one in each time slice. The action for this discretized path is

$$S = \int_0^T dt \left(\frac{m}{2} \dot{x}^2 - V(x) \right) \longrightarrow \sum_k \left[\frac{m}{2} \frac{(x_{k+1} - x_k)^2}{\epsilon} - \epsilon V\left(\frac{x_{k+1} + x_k}{2}\right) \right].$$

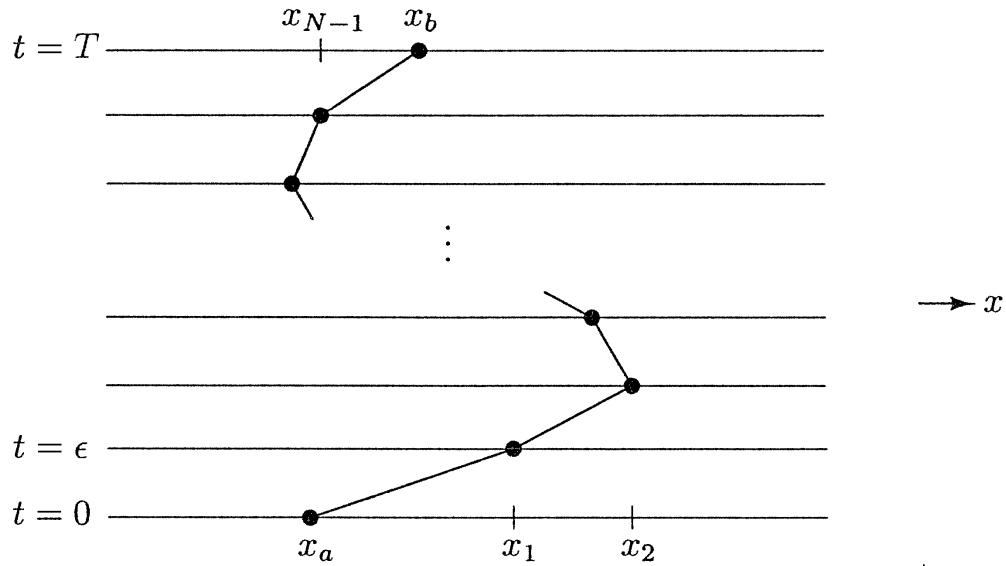


Figure 9.2. We define the path integral by dividing the time interval into small slices of duration ϵ , then integrating over the coordinate x_k of each slice.

We then define the path integral by

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

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 $\epsilon \rightarrow 0$. (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 $\epsilon \rightarrow 0$, 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

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

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

(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

$$e^{-iHT} = e^{-iH\epsilon} e^{-iH\epsilon} e^{-iH\epsilon} \dots e^{-iH\epsilon} \quad (N \text{ factors}).$$

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

$$1 = \left(\prod_i \int dq_k^i \right) |q_k\rangle \langle q_k|.$$

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

$$\langle q_{k+1} | e^{-iH\epsilon} | q_k \rangle \xrightarrow{\epsilon \rightarrow 0} \langle q_{k+1} | (1 - iH\epsilon + \dots) | q_k \rangle. \quad (9.9)$$

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

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

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

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

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

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

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

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

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

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

(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 :

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

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

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

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:

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

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 :

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

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 $\phi(x)$, and the Hamiltonian is

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

Thus our formula becomes

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

where the functions $\phi(x)$ over which we integrate are constrained to the specific configurations $\phi_a(x)$ at $x^0 = 0$ and $\phi_b(x)$ 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

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

where

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

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

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

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:

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

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 $\phi_1(x_1)$ and $\phi_2(x_2)$, 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

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

We can turn the field $\phi_1(x_1)$ into a Schrödinger operator using $\phi_s(x_1)|\phi_1\rangle = \phi_1(x_1)|\phi_1\rangle$. The completeness relation $\int \mathcal{D}\phi_1 |\phi_1\rangle \langle \phi_1| = 1$ then allows us to eliminate the intermediate state $|\phi_1\rangle$. Similar manipulations work for ϕ_2 , yielding the expression

$$\langle \phi_b | e^{-iH(T-x_2^0)} \phi_s(x_2) e^{-iH(x_2^0-x_1^0)} \phi_s(x_1) e^{-iH(x_1^0+T)} |\phi_a\rangle.$$

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

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

This expression is almost equal to the two-point correlation function. To make it more nearly equal, we take the limit $T \rightarrow \infty(1 - i\epsilon)$. 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

$$e^{-iHT} |\phi_a\rangle = \sum_n e^{-iE_n T} |n\rangle \langle n | \phi_a \rangle \xrightarrow[T \rightarrow \infty(1-i\epsilon)]{} \langle \Omega | \phi_a \rangle e^{-iE_0 \cdot \infty(1-i\epsilon)} |\Omega\rangle.$$

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

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

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:

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

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

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

up to an irrelevant overall constant.

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

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

where $k_n^\mu = 2\pi n^\mu / L$, with n^μ an integer, $|n^\mu| < \pi/\epsilon$, 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

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

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

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

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

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

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:

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

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: $t \rightarrow t(1 - i\epsilon)$. This means that we should change $k^0 \rightarrow k^0(1 + i\epsilon)$ in (9.21) and all subsequent equations; in particular, we should replace $(k^2 - m^2) \rightarrow (k^2 - m^2 + i\epsilon)$. 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

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

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

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

$$\begin{aligned}
&= \prod_i \sqrt{\frac{\pi}{b_i}} \\
&= \text{const} \times [\det B]^{-1/2}.
\end{aligned} \tag{9.24}$$

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

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

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

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

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

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

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 $(\text{Re } \phi_m)^2$ is nonzero, but is exactly canceled by the term involving $(\text{Im } \phi_m)^2$. If $k_l = -k_m$, however, the relation $\phi(-k) = \phi^*(k)$ gives an extra minus sign on the $(\text{Im } \phi_m)^2$ term, so the two terms add. When $k_m^0 < 0$ we obtain the same expression, so the numerator is

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

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

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

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

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

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

$$\begin{aligned} & \frac{1}{V^4} \sum_{m,p} e^{-ik_m \cdot (x_1 - x_2)} e^{-ik_p \cdot (x_3 - x_4)} \left(\prod_{n|k_n^0 > 0} \frac{-i\pi V}{m^2 - k_n^2} \right) \frac{-iV}{m^2 - k_n^2 - i\epsilon} \frac{-iV}{m^2 - k_n^2 - i\epsilon} \\ & \xrightarrow[V \rightarrow \infty]{} \left(\prod_{n|k_n^0 > 0} \frac{-i\pi V}{m^2 - k_n^2} \right) D_F(x_1 - x_2) D_F(x_3 - x_4). \end{aligned}$$

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

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

Then the four-point function is simply

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

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.

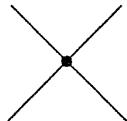
We are now ready to move from the free Klein-Gordon theory to ϕ^4 theory. Add to \mathcal{L}_0 a ϕ^4 interaction:

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

Assuming that λ is small, we can expand

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

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 $i \int d^3x \mathcal{L}_{\text{int}} = -iH_{\text{int}}$, 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



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

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)

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

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

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

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

$$\frac{\delta}{\delta J(x)} \exp \left[i \int d^4 y J(y) \phi(y) \right] = i \phi(x) \exp \left[i \int d^4 y J(y) \phi(y) \right]. \quad (9.32)$$

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

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

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

$$Z[J] \equiv \int \mathcal{D}\phi \exp \left[i \int d^4 x [\mathcal{L} + J(x) \phi(x)] \right]. \quad (9.34)$$

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

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

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

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

(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,

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

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

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

More symbolically, we could write the change of variables as

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

and the result

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

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

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

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

$$Z[J] = Z_0 \exp \left[-\frac{1}{2} \int d^4x d^4y J(x) D_F(x-y) J(y) \right]. \quad (9.39)$$

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

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

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

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

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

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

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

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

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 $t \rightarrow -ix^0$ produces a Euclidean 4-vector product:

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

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

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

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

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

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

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

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

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

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,

$$\overbrace{k \rightarrow}^{\sim\sim\sim\sim} = \frac{-ig_{\mu\nu}}{k^2 + i\epsilon}. \quad (9.49)$$

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

$$\int \mathcal{D}A e^{iS[A]}, \quad (9.50)$$

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

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

This expression vanishes when $\tilde{A}_\mu(k) = k_\mu \alpha(k)$, 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

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

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

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

The troublesome modes are those for which $A_\mu(x) = \partial_\mu \alpha(x)$, 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, $\delta(G(A))$. (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:

$$1 = \int \mathcal{D}\alpha(x) \delta(G(A^\alpha)) \det\left(\frac{\delta G(A^\alpha)}{\delta \alpha}\right), \quad (9.53)$$

where A^α denotes the gauge-transformed field,

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

Equation (9.53) is the continuum generalization of the identity

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

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 $\det(\delta G(A^\alpha)/\delta \alpha)$ 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

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

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

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

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

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

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, $\det(\delta G(A^\alpha)/\delta \alpha) = \det(\partial^2/e)$. Thus the functional integral becomes

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

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

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

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,

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

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

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

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

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

which has the solution

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

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

$$\xi = 0 \quad \text{Landau gauge;}$$

$$\xi = 1 \quad \text{Feynman gauge.}$$

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

$$S = P_0 S_{\text{FP}} P_0. \quad (9.59)$$

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

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

where the sum on the left-hand side runs only over transverse polarizations. The same argument applies if \mathcal{M}^μ 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

$$SS^\dagger = P_0 S_{\text{FP}} P_0 S_{\text{FP}}^\dagger P_0 = P_0 S_{\text{FP}} S_{\text{FP}}^\dagger P_0. \quad (9.61)$$

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 ,

$$\theta\eta = -\eta\theta. \quad (9.62)$$

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 :

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

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 $\theta \rightarrow \theta + \eta$ yields the condition

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

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*

$$\int d\theta (A + B\theta) = B. \quad (9.63)$$

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

$$\int d\theta \int d\eta \ \eta\theta = +1, \quad (9.64)$$

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:

$$(\theta\eta)^* \equiv \eta^*\theta^* = -\theta^*\eta^*. \quad (9.65)$$

To integrate over complex Grassmann numbers, let us define

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

so that $d\theta_1 d\theta_2 = d\theta d\theta^*$. We can now treat θ and θ^* as independent Grassmann numbers.

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

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

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

$$\int d\theta^* d\theta \theta\theta^* e^{-\theta^* b\theta} = 1 = \frac{1}{b} \cdot b. \quad (9.67)$$

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

$$\prod_i \theta'_i = \frac{1}{n!} \epsilon^{ij\cdots l} \theta'_i \theta'_j \cdots \theta'_l$$

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

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

In a general integral

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

the only term of $f(\theta)$ that survives has exactly one factor of each θ_i and θ_i^* ; it is proportional to $(\prod_i \theta_i)(\prod_i \theta_i^*)$. 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 :

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

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

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

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,

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

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

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

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

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

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

$$\frac{d}{d\eta} \theta\eta = -\frac{d}{d\eta} \eta\theta = -\theta. \quad (9.75)$$

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

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

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

$$\begin{aligned}\mathcal{L}_{\text{QED}} &= \bar{\psi}(i\cancel{D} - m)\psi - \frac{1}{4}(F_{\mu\nu})^2 \\ &= \bar{\psi}(i\cancel{\partial} - m)\psi - \frac{1}{4}(F_{\mu\nu})^2 - e\bar{\psi}\gamma^\mu\psi A_\mu \\ &= \mathcal{L}_0 - e\bar{\psi}\gamma^\mu\psi A_\mu,\end{aligned}$$

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

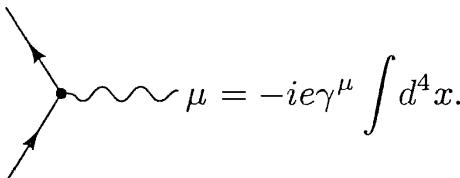
To evaluate correlation functions, we expand the exponential of the interaction term:

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

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

$$\begin{aligned}\overrightarrow{p} &= \int \frac{d^4p}{(2\pi)^4} \frac{i e^{-ip \cdot (x-y)}}{\cancel{p} - m + i\epsilon}; \\ \overbrace{q} &= \int \frac{d^4q}{(2\pi)^4} \frac{-i g_{\mu\nu} e^{-iq \cdot (x-y)}}{q^2 + i\epsilon} \quad (\text{Feynman gauge}).\end{aligned}$$

The interaction term gives the QED vertex,



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

$$\int \mathcal{D}\bar{\psi} \mathcal{D}\psi \exp \left[i \int d^4x \bar{\psi}(i\cancel{D} - m)\psi \right], \quad (9.76)$$

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

$$\begin{aligned} &= \det(i\cancel{D} - m) = \det(i\cancel{\partial} - m - e\cancel{A}) \\ &= \det(i\cancel{\partial} - m) \cdot \det \left(1 - \frac{i}{i\cancel{\partial} - m} (-ie\cancel{A}) \right). \end{aligned}$$

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.

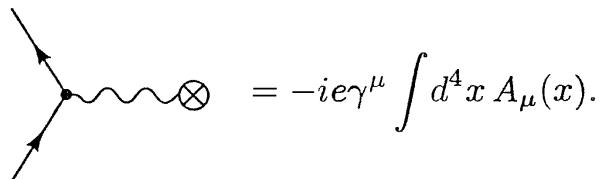
To demonstrate this, we need only apply standard identities from linear algebra. First notice that, if a matrix B has eigenvalues b_i , we can write its determinant as

$$\det B = \prod_i b_i = \exp \left[\sum_i \log b_i \right] = \exp \left[\text{Tr}(\log B) \right], \quad (9.77)$$

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[†]

$$\det \left(1 - \frac{i}{i\cancel{\partial} - m} (-ie\cancel{A}) \right) = \exp \left[\sum_{n=1}^{\infty} -\frac{1}{n} \text{Tr} \left[\left(\frac{i}{i\cancel{\partial} - m} (-ie\cancel{A}) \right)^n \right] \right]. \quad (9.78)$$

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 $\text{Tr}()$ to denote operator traces, and $\text{tr}()$ to denote Dirac traces.

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

$$\det\left(1 - \frac{i(-ie\mathcal{A})}{i\partial - m}\right) = 1 + \text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} + \text{Diagram 4} + \dots$$

$$= \exp \left[\text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} + \dots \right]. \quad (9.79)$$

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

$$\text{Diagram 1} \text{ and } \text{Diagram 2} = \frac{1}{2} \left(\text{Diagram 1} \right)^2.$$

Now let us evaluate the n th 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

$$\text{Diagram 1} = -\frac{1}{n} \int dx_1 \dots dx_n \text{tr} \left[(-ie\mathcal{A}(x_1)) S_F(x_2 - x_1) \dots (-ie\mathcal{A}(x_n)) S_F(x_1 - x_n) \right]$$

$$= -\frac{1}{n} \text{Tr} \left[\left(\frac{i}{i\partial - m} (-ie\mathcal{A}) \right)^n \right], \quad (9.80)$$

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:

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

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

$$\phi(x) \rightarrow \phi'(x) = \phi(x) + \epsilon(x). \quad (9.82)$$

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

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

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

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

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

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

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:

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

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

$$0 = \int \mathcal{D}\phi e^{i\int d^4x \mathcal{L}} \left\{ i \int d^4x \epsilon(x) \frac{\delta}{\delta \varphi(x)} \left(\int d^4x' \mathcal{L} \right) \cdot \varphi(x_1) \varphi(x_2) \right. \\ \left. + \epsilon(x_1) \varphi(x_2) + \varphi(x_1) \epsilon(x_2) \right\}, \quad (9.87)$$

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

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

this the is 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

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

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

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

This Lagrangian is invariant under the transformation $\phi \rightarrow e^{i\alpha}\phi$. 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

$$\phi(x) \rightarrow \phi'(x) = \phi(x) + i\alpha(x)\phi(x). \quad (9.90)$$

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,

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

Expanding this equation to first order in α , we find

$$0 = \int \mathcal{D}\phi e^{i\int d^4x \mathcal{L}} \left\{ i \int d^4x \left[(\partial_\mu \alpha) \cdot i(\phi \partial^\mu \phi^* - \phi^* \partial^\mu \phi) \right] \phi(x_1)\phi^*(x_2) \right. \\ \left. + [i\alpha(x_1)\phi(x_1)]\phi^*(x_2) + \phi(x_1)[-i\alpha(x_2)\phi^*(x_2)] \right\}.$$

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

$$\langle \partial_\mu j^\mu(x) \phi(x_1)\phi^*(x_2) \rangle = (-i) \left\langle (i\phi(x_1)\delta(x-x_1))\phi^*(x_2) \right. \\ \left. + \phi(x_1)(-i\phi^*(x_2)\delta(x-x_2)) \right\rangle, \quad (9.91)$$

where

$$j^\mu = i(\phi \partial^\mu \phi^* - \phi^* \partial^\mu \phi) \quad (9.92)$$

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

$$\varphi_a(x) \rightarrow \varphi_a(x) + \epsilon \Delta \varphi_a(x). \quad (9.93)$$

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:

$$\mathcal{L}[\varphi] \rightarrow \mathcal{L}[\varphi] + \epsilon \partial_\mu \mathcal{J}^\mu. \quad (9.94)$$

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:

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

Summation over the index a is understood. Then

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

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

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

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

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

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

$$\varphi_a \rightarrow \varphi_a + a^\mu(x) \partial_\mu \varphi_a \quad (9.98)$$

the Lagrangian transforms as

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

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

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

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

$$\psi(x) \rightarrow (1 + ie\alpha(x))\psi(x), \quad (9.100)$$

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

$$\mathcal{L} \rightarrow \mathcal{L} - e\partial_\mu\alpha\bar{\psi}\gamma^\mu\psi. \quad (9.101)$$

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

$$0 = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}A e^{i\int d^4x \mathcal{L}} \left\{ -i \int d^4x \partial_\mu\alpha(x) \left[j^\mu(x)\psi(x_1)\bar{\psi}(x_2) \right] \right. \\ \left. + (ie\alpha(x_1)\psi(x_1))\bar{\psi}(x_2) + \psi(x_1)(-ie\alpha(x_2)\bar{\psi}(x_2)) \right\}, \quad (9.102)$$

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

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

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

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

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

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

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

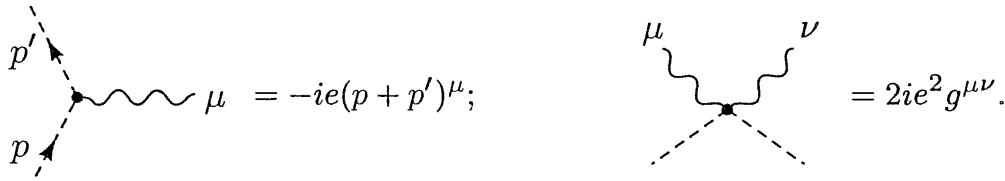
$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + (D_\mu\phi)^*(D^\mu\phi) - m^2\phi^*\phi,$$

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:

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

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



- (b) Compute, to lowest order, the differential cross section for $e^+e^- \rightarrow \phi\phi^*$. 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 $e^+e^- \rightarrow \mu^+\mu^-$.
- (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

$$Z = \text{tr}[e^{-\beta H}]$$

(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,

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

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

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

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).

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

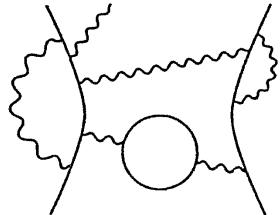
In this section we will use elementary arguments to determine, tentatively, when a Feynman diagram contains an ultraviolet divergence. We begin by analyzing quantum electrodynamics.

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

- N_e = number of external electron lines;
- N_γ = number of external photon lines;
- P_e = number of electron propagators;
- P_γ = number of photon propagators;
- V = number of vertices;
- L = number of loops.

(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:



$$\sim \int \frac{d^4 k_1 d^4 k_2 \cdots d^4 k_L}{(k_i - m) \cdots (k_j^2) \cdots (k_n^2)}.$$

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

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

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

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

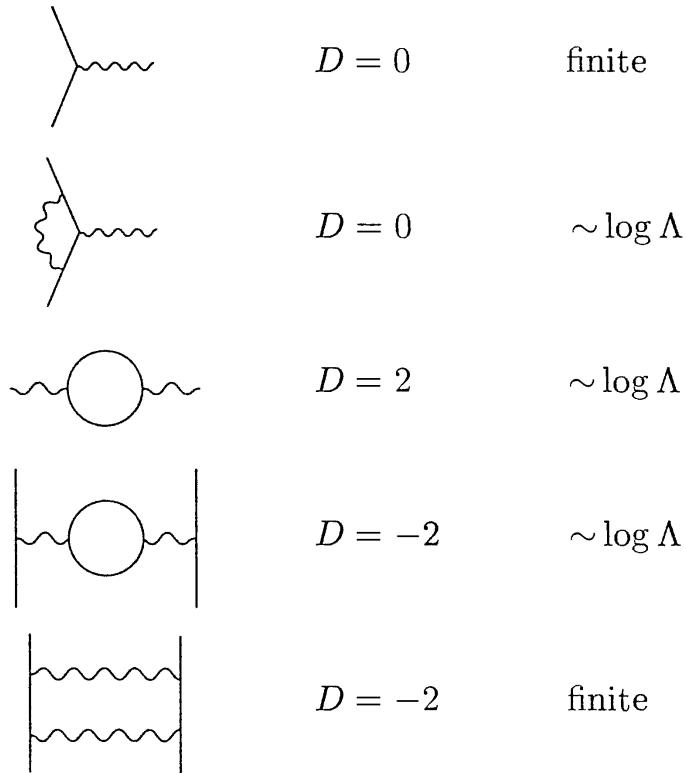


Figure 10.1. Some simple QED diagrams that illustrate the superficial degree of divergence. The first diagram is finite, even though $D = 0$. The third diagram has $D = 2$ but only a logarithmic divergence, due to the Ward identity (see Section 7.5). The fourth diagram diverges, even though $D < 0$, 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

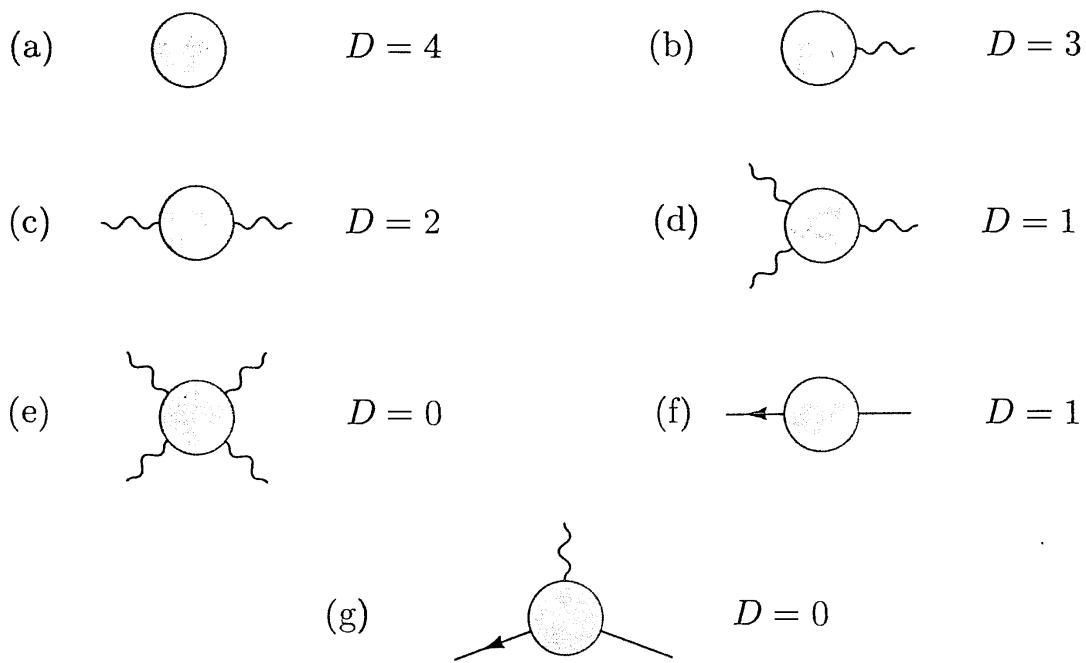


Figure 10.2. The seven QED amplitudes whose superficial degree of divergence (D) is ≥ 0 . (Each circle represents the sum of all possible QED diagrams.) As explained in the text, amplitude (a) is irrelevant to scattering processes, while amplitudes (b) and (d) vanish because of symmetries. Amplitude (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

$$\text{Diagram with a circle and a wavy line entering from the left labeled } q = -ie \int d^4x e^{-iq \cdot x} \langle \Omega | T j_\mu(x) | \Omega \rangle, \quad (10.5)$$

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:

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

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 $p = 0$:

$$\text{Diagram with arrow } p \rightarrow = A_0 + A_1 p + A_2 p^2 + \dots,$$

where each coefficient is independent of p :

$$A_n = \frac{1}{n!} \frac{d^n}{dp^n} \left(\text{Diagram with arrow } p \rightarrow \right) \Big|_{p=0}.$$

(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

$$\frac{d}{dp} \left(\frac{1}{k + p - m} \right) = -\frac{1}{(k + p - m)^2}.$$

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 $m \log \Lambda$. We therefore expect the entire self-energy amplitude to have the form

$$\text{Diagram with arrow } p \rightarrow = a_0 m \log \Lambda + a_1 p \log \Lambda + (\text{finite terms}), \quad (10.6)$$

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:

$$\text{Diagram with arrow } \mu \rightarrow \propto -ie\gamma^\mu \log \Lambda + \text{finite terms.} \quad (10.7)$$

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

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

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:

$$k^\mu \left(\text{---} \text{---} \text{---} \text{---} \right) = 0. \quad (10.9)$$

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

$$D \equiv dL - P_e - 2P_\gamma, \quad (10.10)$$

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

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

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:

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

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

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

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 $S = \int d^d x \mathcal{L}$ must be dimensionless, since we work in units where $\hbar = 1$. In this system of units, the integral $d^d x$ has units (mass) $^{-d}$, and so the Lagrangian has units (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 $\Lambda \rightarrow \infty$.

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

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

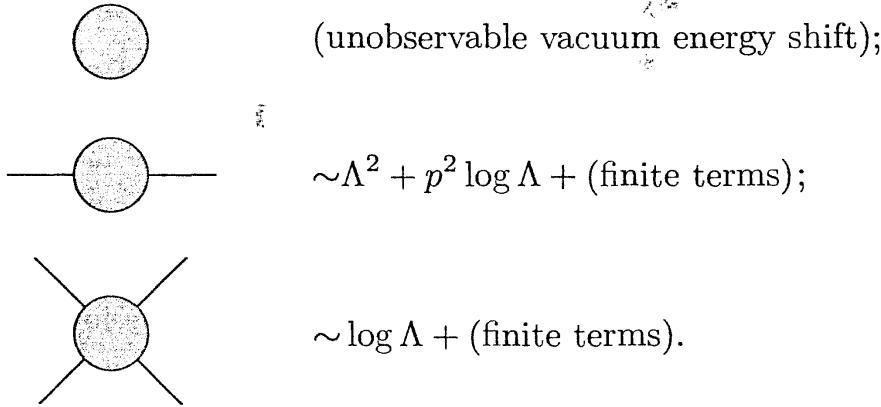
We now write m_0 and λ_0 , to emphasize that these are the bare values of the mass and coupling constant, not the values measured in experiments.

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 $\phi \rightarrow -\phi$, all amplitudes with an odd

number of external legs vanish. The only divergent amplitudes are therefore



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

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

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

$$\phi = Z^{1/2} \phi_r. \quad (10.15)$$

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 we originally guessed in Eq. (4.103).

The Lagrangian is much uglier after the rescaling:

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

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

$$\delta_Z = Z - 1, \quad \delta_m = m_0^2 Z - m^2, \quad \delta_\lambda = \lambda_0 Z^2 - \lambda, \quad (10.17)$$

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

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

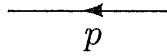
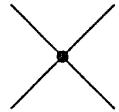
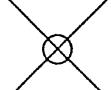
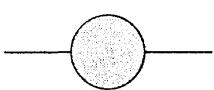
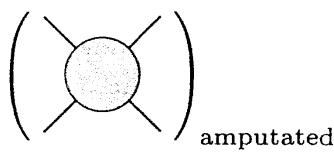
	$= \frac{i}{p^2 - m^2 + i\epsilon}$
	$= -i\lambda$
	$= i(p^2\delta_Z - \delta_m)$
	$= -i\delta_\lambda$

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,

	$= \frac{i}{p^2 - m^2} + (\text{terms regular at } p^2 = m^2);$
	$= -i\lambda \quad \text{at } s = 4m^2, t = u = 0. \quad (10.19)$

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,

$$\begin{aligned}
 i\mathcal{M}(p_1 p_2 \rightarrow p_3 p_4) &= \text{Diagram of a four-point vertex with momenta } p_1, p_2, p_3, p_4 \\
 &= \text{Cross diagram} + \left(\text{Diagram with a loop} + \text{Diagram with a loop} + \text{Diagram with a loop} \right) + \dots
 \end{aligned}$$

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

$$\begin{aligned}
 \text{Diagram with a loop} &= \frac{(-i\lambda)^2}{2} \int \frac{d^4 k}{(2\pi)^4} \frac{i}{k^2 - m^2} \frac{i}{(k + p)^2 - m^2} \\
 &\equiv (-i\lambda)^2 \cdot iV(p^2). \tag{10.20}
 \end{aligned}$$

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

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

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

$$\delta_\lambda = -\lambda^2 [V(4m^2) + 2V(0)]. \quad (10.22)$$

(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

$$\begin{aligned} V(p^2) &= \frac{i}{2} \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{1}{[k^2 + 2xk \cdot p + xp^2 - m^2]^2} \\ &= \frac{i}{2} \int_0^1 dx \int \frac{d^d \ell}{(2\pi)^d} \frac{1}{[\ell^2 + x(1-x)p^2 - m^2]^2} \quad (\ell = k + xp) \\ &= -\frac{1}{2} \int_0^1 dx \int \frac{d^d \ell_E}{(2\pi)^d} \frac{1}{[\ell_E^2 - x(1-x)p^2 + m^2]^2} \quad (\ell_E^0 = -i\ell^0) \\ &= -\frac{1}{2} \int_0^1 dx \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \frac{1}{[m^2 - x(1-x)p^2]^{2-d/2}} \\ &\xrightarrow[d \rightarrow 4]{} -\frac{1}{32\pi^2} \int_0^1 dx \left(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log[m^2 - x(1-x)p^2] \right), \quad (10.23) \end{aligned}$$

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

$$\begin{aligned} \delta_\lambda &= \frac{\lambda^2}{2} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \int_0^1 dx \left(\frac{1}{[m^2 - x(1-x)4m^2]^{2-d/2}} + \frac{2}{[m^2]^{2-d/2}} \right) \\ &\xrightarrow[d \rightarrow 4]{} \frac{\lambda^2}{32\pi^2} \int_0^1 dx \left(\frac{6}{\epsilon} - 3\gamma + 3\log(4\pi) - \log[m^2 - x(1-x)4m^2] - 2\log[m^2] \right). \quad (10.24) \end{aligned}$$

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

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

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

$$\text{---} \circlearrowleft \text{---} = -iM(p^2). \quad (10.26)$$

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

$$\begin{aligned} \text{---} \circlearrowleft \text{---} &= \text{---} \text{---} + \text{---} \circlearrowleft \text{---} + \text{---} \circlearrowleft \text{---} \circlearrowleft \text{---} + \dots \\ &= \frac{i}{p^2 - m^2 - M(p^2)}. \end{aligned} \quad (10.27)$$

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

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

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

Explicitly, to one-loop order,

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

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

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

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

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

$$\text{---} \circlearrowleft \text{---} + \text{---} \circlearrowleft \text{---} \otimes \text{---} + \text{---} \otimes \text{---} \quad (10.31)$$

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. For example, 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

$$\begin{aligned}
 -iM(p^2) &= -\frac{p}{\text{---}} \text{---} \circlearrowleft \text{---} \frac{k+p}{k} \text{---} + \text{---} \otimes \text{---} \\
 &= -(-ig)^2 \int \frac{d^d k}{(2\pi)^d} \text{tr} \left[\frac{i(k+p+m_f)}{(k+p)^2 - m_f^2} \frac{i(k+m_f)}{k^2 - m_f^2} \right] + i(p^2 \delta_Z - \delta_m) \\
 &= -4g^2 \int \frac{d^d k}{(2\pi)^d} \frac{k \cdot (p+k) + m_f^2}{((p+k)^2 - m_f^2)(k^2 - m_f^2)} + i(p^2 \delta_Z - \delta_m), \quad (10.32)
 \end{aligned}$$

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

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

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

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

To determine δ_Z , we cancel also the first derivative with respect to p^2 of the

loop integral (10.33). This gives

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

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

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

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

$$\text{Diagram with incoming electron line and outgoing photon line} = \frac{iZ_2}{p - m} + \dots; \quad \text{Diagram with incoming photon line and outgoing electron line} = \frac{-iZ_3 g_{\mu\nu}}{q^2} + \dots.$$

(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

$$\mathcal{L} = -\frac{1}{4}Z_3(F_r^{\mu\nu})^2 + Z_2\bar{\psi}_r(i\partial - m_0)\psi_r - e_0Z_2Z_3^{1/2}\bar{\psi}_r\gamma^\mu\psi_r A_{r\mu}. \quad (10.36)$$

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

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

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:

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

where

$$\begin{aligned} \delta_3 &= Z_3 - 1, & \delta_2 &= Z_2 - 1, \\ \delta_m &= Z_2 m_0 - m, & \text{and} & \delta_1 = Z_1 - 1 = (e_0/e)Z_2 Z_3^{1/2} - 1. \end{aligned}$$

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:

$$\begin{aligned} \mu \sim \text{1PI} \sim \nu &= i\Pi^{\mu\nu}(q) = i(g^{\mu\nu}q^2 - q^\mu q^\nu)\Pi(q^2), \\ \text{1PI} &= -i\Sigma(\not{p}), \\ \left(\text{amputated} \right) &= -ie\Gamma^\mu(p', p). \end{aligned} \quad (10.39)$$

[†]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).

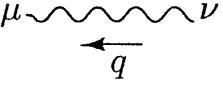
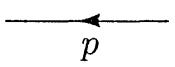
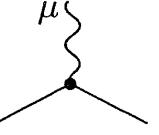
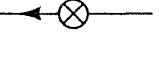
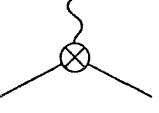
	$= \frac{-ig_{\mu\nu}}{q^2 + i\epsilon}$ (Feynman gauge)
	$= \frac{i}{p - m + i\epsilon}$
	$= -ie\gamma^\mu$
	$= -i(g^{\mu\nu}q^2 - q^\mu q^\nu)\delta_3$
	$= i(\not{p}\delta_2 - \delta_m)$
	$= -ie\gamma^\mu\delta_1$

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 counterterms 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 diagrams 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 ($\epsilon = 4 - d$)

$$-i\Sigma_2(p) = -i \frac{e^2}{(4\pi)^{d/2}} \int_0^1 dx \frac{\Gamma(2-\frac{d}{2})}{((1-x)m^2 + x\mu^2 - x(1-x)p^2)^{2-d/2}} \times ((4-\epsilon)m - (2-\epsilon)x\cancel{p}). \quad (10.41)$$

The counterterm δ_m , according to the first of conditions (10.40), is therefore

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

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

$$\begin{aligned} \delta_2 &= \frac{d}{d\cancel{p}} \Sigma_2(m) \\ &= -\frac{e^2}{(4\pi)^{d/2}} \int_0^1 dx \frac{\Gamma(2-\frac{d}{2})}{((1-x)^2 m^2 + x\mu^2)^{2-d/2}} \\ &\quad \times \left[(2-\epsilon)x - \frac{\epsilon}{2} \frac{2x(1-x)m^2}{(1-x)^2 m^2 + x\mu^2} (4 - 2x - \epsilon(1-x)) \right]. \end{aligned} \quad (10.43)$$

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

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

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

Then

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

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

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

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

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

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 $\delta F_1(0) = d\Sigma_2/d\mathbf{p}|_m$. 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

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

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:

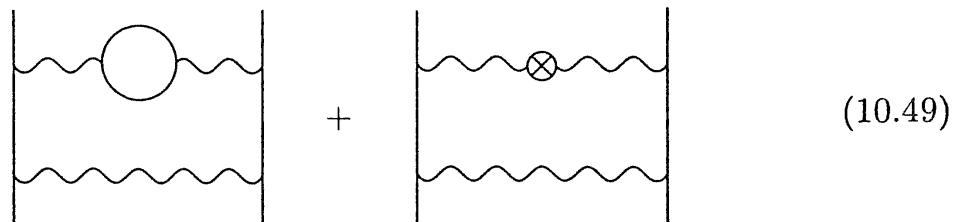
$$e_0 Z_2'^{-1} Z_3^{-1/2} = e Z_1'^{-1}, \quad (10.48)$$

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.

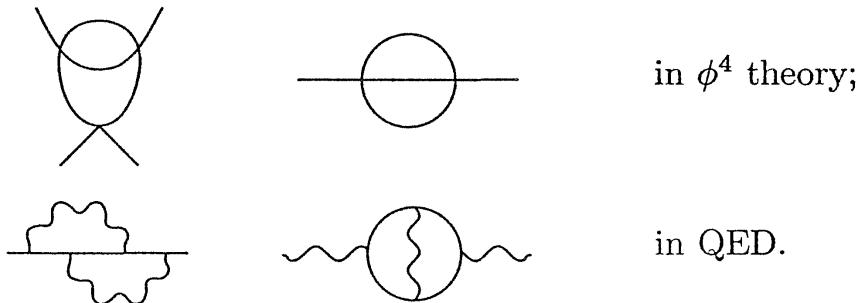
When an otherwise finite diagram contains a divergent subdiagram, the treatment of the divergence is relatively straightforward. For example, the sum of diagrams



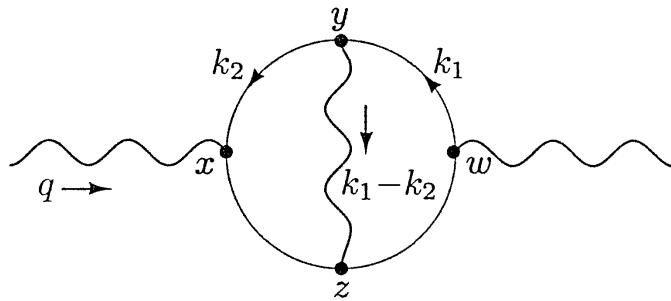
is finite: The divergence in the photon propagator cancels just as when this propagator occurs in a tree diagram. The finite sum of the two propagator

diagrams gives an integrand for the outer loop that falls off fast enough that this integral still converges.

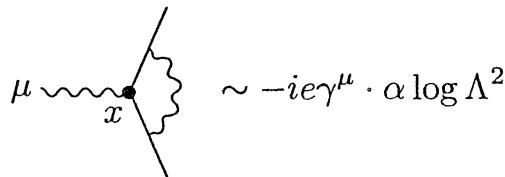
A more difficult situation occurs when we have *nested* or *overlapping* divergences, that is, when two divergent loops share a propagator. Some examples of diagrams with overlapping divergences are



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



in the limit $\Lambda \rightarrow \infty$. 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:

(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 .

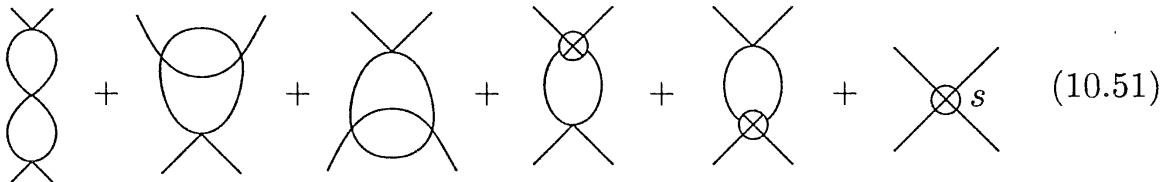
We can extend the lessons of this example to a general picture of the divergences of higher-loop Feynman diagrams and their cancellation. A given diagram may contain local divergences, as predicted by the analysis of Section 10.1. It may also contain nonlocal divergences due to divergent subgraphs embedded 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,



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,

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

[‡]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).

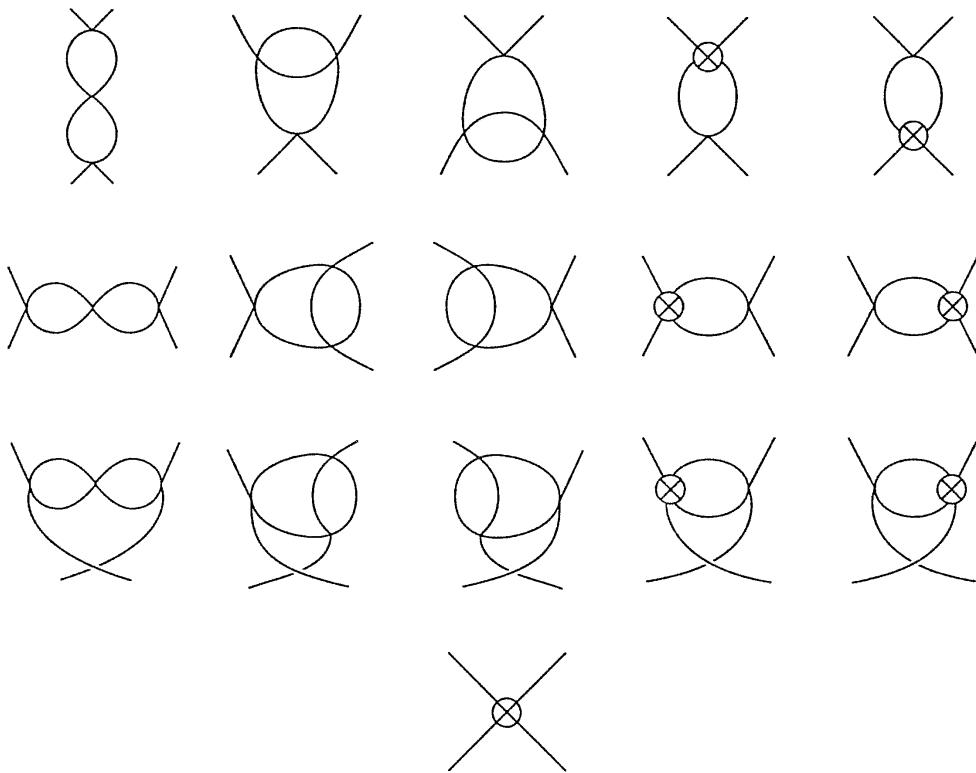


Figure 10.5. The two-loop contributions to the four-point function in ϕ^4 theory. Note that the diagrams in the first three lines are related to each other by crossing, being in the s -, t -, and u -channels, respectively. The last two diagrams in each of these lines involve the $\mathcal{O}(\lambda^2)$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

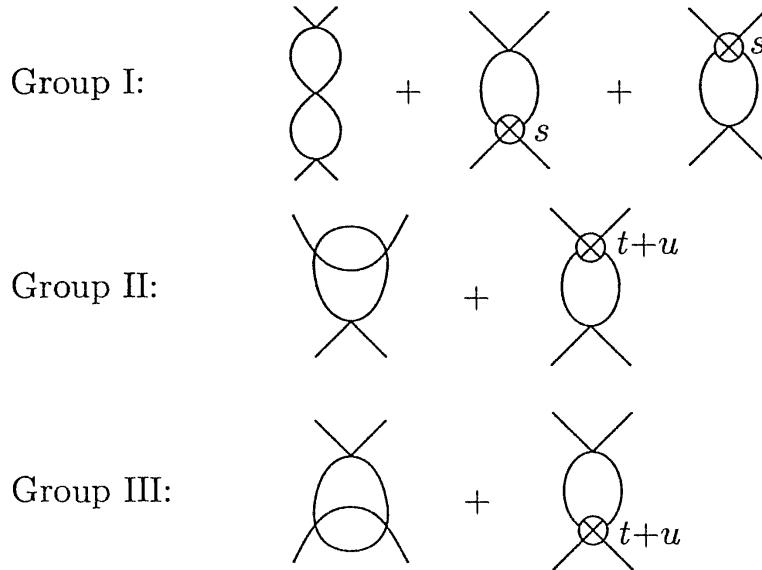
$$\text{Diagram} = -i\delta_\lambda = (-i\lambda)^2 [-iV(4m^2) - 2iV(0)].$$

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

$$\text{Diagram}_s = (-i\lambda)^2 \cdot -iV(4m^2); \quad \text{Diagram}_{t+u} = (-i\lambda)^2 \cdot -2iV(0).$$

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

follows:



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

$$\begin{aligned}
 & \text{Diagram 1:} & & = (-i\lambda)^3 \cdot [iV(p^2)]^2; \\
 & \text{Diagram 2:} & & = \text{Diagram 3:} & & = (-i\lambda)^3 \cdot iV(p^2) \cdot -iV(4m^2).
 \end{aligned}$$

The sum of all three diagrams is therefore

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

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 $\mathcal{O}(\lambda^2)$ vertex counterterm) is proportional to

$$[V(4m^2)]^2 \propto [\Gamma(2 - \frac{d}{2})]^2 \xrightarrow{d \rightarrow 4} \left(\frac{2}{\epsilon}\right)^2 \quad \text{for } d = 4 - \epsilon.$$

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,

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

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

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

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:

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

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

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

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

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

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

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

so that

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

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

$$\begin{aligned} &= -\frac{\lambda^3}{2} \frac{\Gamma(4-\frac{d}{2})}{(4\pi)^{d/2}} \int_0^1 dx \int_0^1 dy \int_0^1 dw \int \frac{d^d k}{(2\pi)^d} \\ &\quad \times \frac{w^{1-\frac{d}{2}}(1-w)}{(w[m^2 - x(1-x)(k+p_3)^2] + (1-w)[m^2 - k^2 + 2yk \cdot p + yp^2])^{4-\frac{d}{2}}}. \end{aligned} \quad (10.57)$$

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

$$[(1-w) + wx(1-x)]\ell^2 + P^2 - m^2, \quad (10.58)$$

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 $w \rightarrow 0$,

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

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

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

This expression has one obvious pole as $d \rightarrow 4$, 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 $f(0)$:

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

The second piece is

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

This term has only a simple pole as $d \rightarrow 4$; 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 $\mathcal{O}(\lambda^2)$ 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

$$\begin{aligned} & -\frac{i\lambda^3}{2(4\pi)^d} \left(\frac{2}{\epsilon} \right) \int_0^1 dy \frac{\Gamma(\epsilon)}{[m^2 - y(1-y)p^2]^\epsilon} \\ & \xrightarrow{d \rightarrow 4} -\frac{i\lambda^3}{2(4\pi)^4} \left(\frac{2}{\epsilon} \right) \int_0^1 dy \left(\frac{1}{\epsilon} - \gamma - \log[m^2 - y(1-y)p^2] \right), \end{aligned} \quad (10.62)$$

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:

$$\begin{aligned}
 \text{Diagram: } & t+u \\
 & = (-i\lambda)^3 \cdot -2iV(0) \cdot iV(p^2) \\
 & = \frac{i\lambda^3}{2(4\pi)^d} \int_0^1 dy \frac{\Gamma(2-\frac{d}{2})}{[m^2]^{2-d/2}} \frac{\Gamma(2-\frac{d}{2})}{[m^2 - y(1-y)p^2]^{2-d/2}} \\
 & \xrightarrow{d \rightarrow 4} \frac{i\lambda^3}{2(4\pi)^4} \int_0^1 dy \left(\frac{2}{\epsilon} - \gamma - \log m^2 \right) \\
 & \quad \times \left(\frac{2}{\epsilon} - \gamma - \log [m^2 - y(1-y)p^2] \right). \quad (10.63)
 \end{aligned}$$

(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 $p \rightarrow \infty$.

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,

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

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

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 $d \rightarrow 4$) of each counterterm, to the one-loop order of perturbation theory, implementing the renormalization conditions that you specified in part (a). 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:

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

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 $s \rightarrow \infty$, t fixed. Ignore all masses on internal lines, and keep external masses nonzero only as infrared regulators where these are needed. Show that

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

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

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, $-\mu^2$:

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

This Lagrangian has a discrete symmetry: It is invariant under the operation $\phi \rightarrow -\phi$. 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

$$\phi_0 = \pm v = \pm \sqrt{\frac{6}{\lambda}}\mu. \quad (11.2)$$

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

$$\phi(x) = v + \sigma(x), \quad (11.3)$$

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

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

This Lagrangian describes a simple scalar field of mass $\sqrt{2}\mu$, with σ^3 and σ^4 interactions. The symmetry $\phi \rightarrow -\phi$ 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.

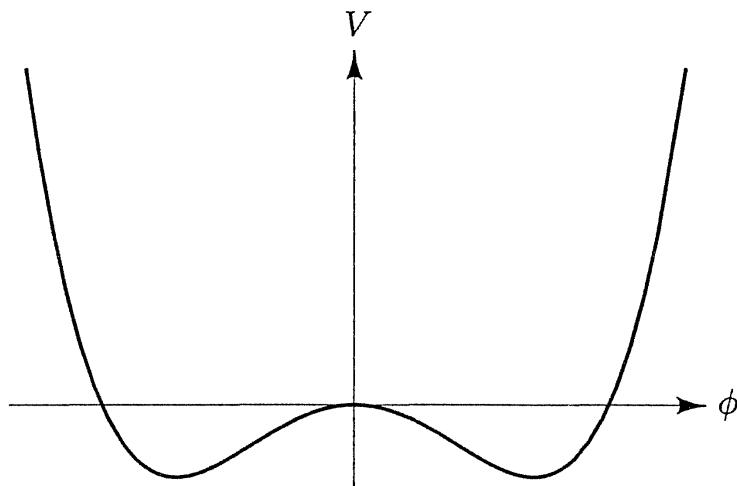


Figure 11.1. Potential for spontaneous symmetry breaking in the discrete case.

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 $\phi^i(x)$:

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

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

$$\phi^i \longrightarrow R^{ij} \phi^j \quad (11.6)$$

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

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

(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

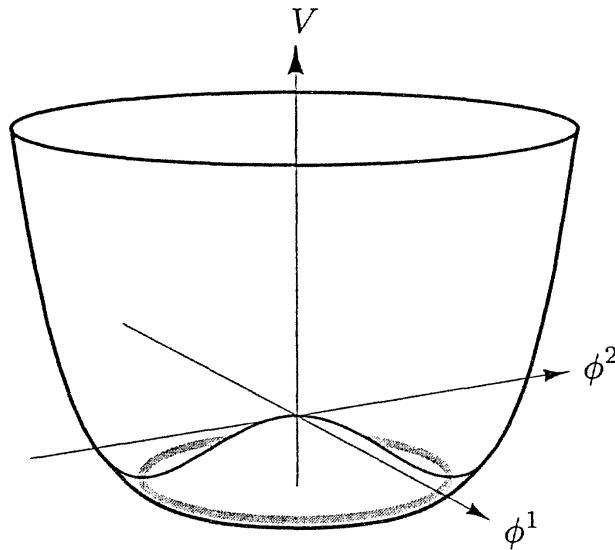


Figure 11.2. Potential for spontaneous breaking of a continuous $O(N)$ symmetry, drawn for the case $N = 2$. Oscillations along the trough in the potential correspond to the massless π fields.

N th direction:

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

We can now define a set of shifted fields by writing

$$\phi^i(x) = (\pi^k(x), v + \sigma(x)), \quad k = 1, \dots, N-1. \quad (11.8)$$

(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

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

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

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

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

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

Expanding V about this minimum, we find

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

The coefficient of the quadratic term,

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

[†]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

$$\phi^a \longrightarrow \phi^a + \alpha \Delta^a(\phi), \quad (11.12)$$

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

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

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

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

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

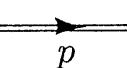
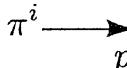
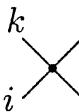
σ  $= \frac{i}{p^2 - 2\mu^2}$	π^i  $\pi^j = \frac{i\delta^{ij}}{p^2}$
 $= -6i\lambda v$	 $= -2i\delta^{ij}\lambda v$
 $= -6i\lambda$	 $= -2i\lambda\delta^{ij}$
 $= -2i\delta_\lambda [\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}]$	

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:

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

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

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

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.

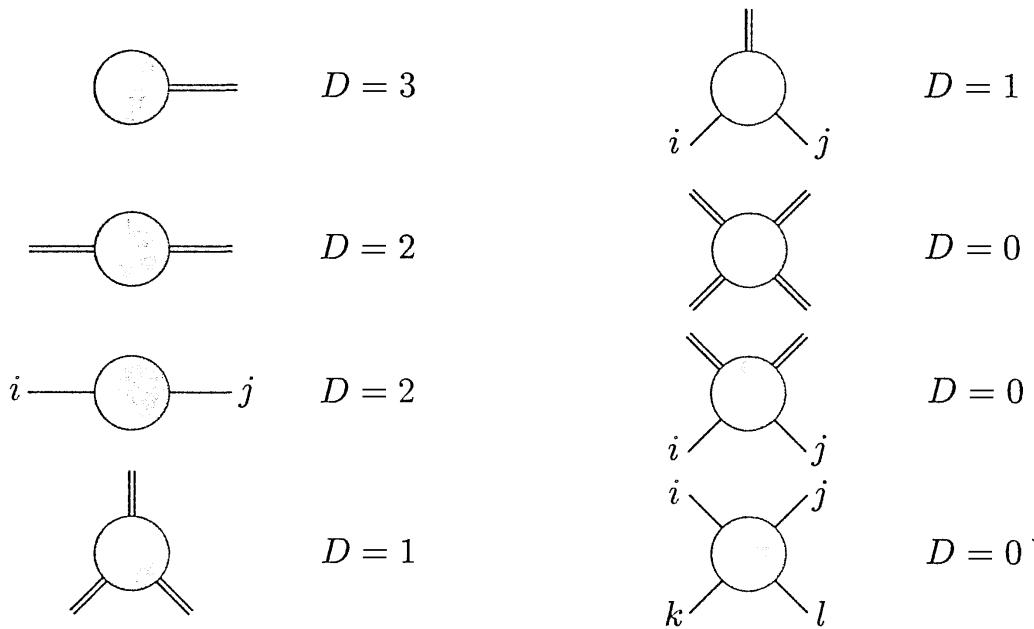


Figure 11.4. Divergent amplitudes in the linear sigma model.

$\otimes \text{---}$	$= -i(\delta_\mu v + \delta_\lambda v^3)$	
$\text{---} \otimes \text{---}$	$= -i(\delta_Z p^2 + \delta_\mu + 3\delta_\lambda v^2)$	$\times \times = -6i\delta_\lambda$
$i \text{---} \otimes \text{---} j$	$= -i\delta^{ij}(\delta_Z p^2 + \delta_\mu + \delta_\lambda v^2)$	$i \times \otimes j = -2i\delta^{ij}\delta_\lambda$
$\text{---} \otimes \text{---}$	$= -6i\delta_\lambda v$	
$i \text{---} \otimes \text{---} j$	$= -2i\delta^{ij}\delta_\lambda v$	$i \text{---} \otimes \text{---} j = -2i\delta_\lambda [\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}]$

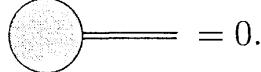
Figure 11.5. Feynman rules for counterterm vertices in the linear sigma model.

Fortunately, the miracle does occur. We will see below that the counter-terms 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

$$\langle \phi^N \rangle = \frac{\mu}{\sqrt{\lambda}} \quad (11.16)$$

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:

$$\begin{aligned} \text{1PI} &= 0; \\ \frac{d}{dp^2} \left(\text{1PI} \right) &= 0 \quad \text{at } p^2 = m^2; \\ \text{4-point} &= -6i\lambda \quad \text{at } s = 4m^2, t = u = 0. \end{aligned} \quad (11.17)$$

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:

$$\begin{array}{c} \text{Diagram 1} \\ + \end{array} \quad \begin{array}{c} \text{Diagram 2} \\ + \end{array} \quad + \text{(crosses)} \\
 + \quad \begin{array}{c} \text{Diagram 3} \\ + \end{array} \quad \begin{array}{c} \text{Diagram 4} \\ + \end{array} \quad + \text{(crosses)} \\
 + \quad \begin{array}{c} \text{Diagram 5} \\ + \end{array} \quad \begin{array}{c} \text{Diagram 6} \\ + \end{array} \quad + \text{(crosses)} \quad + \quad \text{Diagram 7}
 \end{array} \tag{11.18}$$

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,

$$\text{Diagram 8} \sim \int \frac{d^4 k}{(2\pi)^4} \frac{1}{k^2} \frac{1}{k^2} \frac{1}{k^2}.$$

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:

$$\begin{aligned}
 \text{Diagram 1} &= \frac{1}{2} \cdot (-6i\lambda)^2 \cdot \int \frac{d^d k}{(2\pi)^d} \frac{i}{k^2 - 2\mu^2} \frac{i}{(k+p)^2 - 2\mu^2} \\
 &= 18\lambda^2 \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{1}{[k^2 - \Delta]^2}
 \end{aligned}$$

$$\begin{aligned}
&= 18\lambda^2 \int_0^1 dx \frac{i}{(4\pi)^{d/2}} \Gamma(2-\frac{d}{2}) \left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}} \\
&= 18i\lambda^2 \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} + (\text{finite terms}). \tag{11.19}
\end{aligned}$$

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, \dots 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):

$$+ \text{ (crosses)} \sim 6i\lambda^2(N+8) \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \tag{11.21}$$

(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

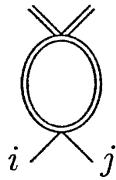
$$\delta_\lambda \sim \lambda^2(N+8) \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \tag{11.22}$$

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

$$+ \text{ (crosses)} + \text{ (diagram with two internal lines and a loop)} + \text{ (diagram with two internal lines and a loop)} . \tag{11.23}$$

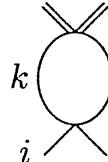
and from several diagrams with three-point vertices which, as argued earlier, are manifestly finite. Each of the diagrams in (11.23) contains a loop integral analogous to that in (11.19), whose infinite part is always $-i\Gamma(2-\frac{d}{2})/(4\pi)^2$.

The only differences are in the vertices and symmetry factors. For example, the infinite part of the first diagram of (11.23) is



$$\sim \frac{1}{2} \cdot (-6i\lambda)(-2i\lambda\delta^{ij}) \cdot \frac{-i}{(4\pi)^2} \Gamma(2-\frac{d}{2}) = 6i\lambda^2\delta^{ij} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

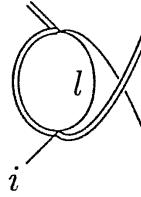
The second diagram is a bit more complicated:



$$\sim \frac{1}{2} \cdot (-2i\lambda\delta^{kl})(-2i\lambda(\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk})) \cdot \frac{-i}{(4\pi)^2} \Gamma(2-\frac{d}{2})$$

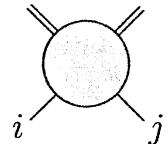
$$= 2i\lambda^2(N+1)\delta^{ij} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

In the third diagram there is no symmetry factor:



$$\sim (-2i\lambda\delta^{il})(-2i\lambda\delta^{jl}) \cdot \frac{-i}{(4\pi)^2} \Gamma(2-\frac{d}{2}) = 4i\lambda^2\delta^{ij} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

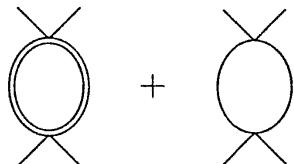
The fourth diagram of (11.23) gives an identical expression, since it is the same as the third but with i and j interchanged. The sum of the four diagrams therefore gives, for the infinite part of the $\sigma\sigma\pi\pi$ vertex,



$$\sim 2i\lambda^2\delta^{ij}(N+8) \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \quad (11.24)$$

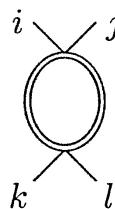
This divergent term is indeed canceled by the $\sigma\sigma\pi\pi$ counterterm, with the value of δ_λ given in (11.22).

The remaining four-point amplitude has four external π fields. The divergent one-loop diagrams are:



$$+ \quad + \quad (\text{crosses}) \quad (11.25)$$

These diagrams all have the same familiar form. The first is



$$\sim \frac{1}{2} \cdot (-2i\lambda\delta^{ij})(-2i\lambda\delta^{kl}) \cdot \frac{-i}{(4\pi)^2} \Gamma(2-\frac{d}{2}) = 2i\lambda^2\delta^{ij}\delta^{kl} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

The second diagram is more complicated:

$$\begin{array}{c} i \\ \diagup \\ \text{circle} \\ \diagdown \\ m \end{array} \begin{array}{c} j \\ \diagup \\ \text{circle} \\ \diagdown \\ n \end{array} \sim \frac{1}{2} \cdot (-2i\lambda(\delta^{ij}\delta^{mn} + \delta^{im}\delta^{jn} + \delta^{in}\delta^{jm})) \\
 \begin{array}{c} k \\ \diagup \\ \text{circle} \\ \diagdown \\ l \end{array} \cdot (-2i\lambda(\delta^{kl}\delta^{mn} + \delta^{km}\delta^{ln} + \delta^{kn}\delta^{lm})) \cdot \frac{-i}{(4\pi)^2} \Gamma(2 - \frac{d}{2}) \\
 = 2i\lambda^2 ((N+3)\delta^{ij}\delta^{kl} + 2\delta^{ik}\delta^{jl} + 2\delta^{il}\delta^{jk}) \frac{\Gamma(2 - \frac{d}{2})}{(4\pi)^2}.
 \end{array}$$

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

$$\begin{array}{c} i \\ \diagup \\ \text{circle} \\ \diagdown \\ k \end{array} \begin{array}{c} j \\ \diagup \\ \text{circle} \\ \diagdown \\ l \end{array} \sim 2i\lambda^2 (\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}) (N+8) \frac{\Gamma(2 - \frac{d}{2})}{(4\pi)^2}. \quad (11.26)$$

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

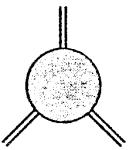
$$\left(\begin{array}{c} \text{circle} \\ \diagup \\ \text{circle} \\ \diagdown \\ \text{circle} \end{array} + \begin{array}{c} \text{circle} \\ \diagup \\ \text{circle} \\ \diagdown \\ \text{circle} \end{array} + \text{crosses} \right) + \begin{array}{c} \text{circle} \\ \diagup \\ \text{circle} \\ \diagdown \\ \text{circle} \end{array} + \begin{array}{c} \text{circle} \\ \diagup \\ \text{circle} \\ \diagdown \\ \text{circle} \end{array}. \quad (11.27)$$

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

$$\begin{array}{c} \text{circle} \\ \diagup \\ \text{circle} \\ \diagdown \\ \text{circle} \end{array} = \frac{1}{2} \cdot (-6i\lambda)(-6i\lambda v) \cdot \int \frac{d^d k}{(2\pi)^d} \frac{i}{k^2 - 2\mu^2} \frac{i}{(k+p)^2 - 2\mu^2} \\
 \sim 18i\lambda^2 v \frac{\Gamma(2 - \frac{d}{2})}{(4\pi)^2}.$$

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,

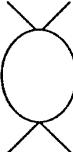


$$\sim 6i\lambda^2 v(N+8) \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}. \quad (11.28)$$

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



$$= \quad (11.29)$$

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

$$0 = \text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} \quad (11.30)$$

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

$$\text{Diagram 1} = \frac{1}{2}(-6\lambda v) \int \frac{d^d k}{(2\pi)^d} \frac{i}{k^2 - 2\mu^2} = -3i\lambda v \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{1}{2\mu^2}\right)^{1-\frac{d}{2}}. \quad (11.31)$$

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

$$\begin{aligned} \text{Diagram} &= \frac{1}{2}(-2i\lambda v)\delta^{ij} \int \frac{d^d k}{(2\pi)^d} \frac{i\delta^{ij}}{k^2 - \zeta^2} \\ &= -i(N-1)\lambda v \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{1}{\zeta^2}\right)^{1-\frac{d}{2}}. \end{aligned} \quad (11.32)$$

Notice that, for $d > 2$, the diagram vanishes in the limit as $\zeta \rightarrow 0$; 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

$$(\delta_\mu + v^2 \delta_\lambda) = -\lambda \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{3}{(2\mu^2)^{1-d/2}} + \frac{N-1}{(\zeta^2)^{1-d/2}} \right). \quad (11.33)$$

Now consider the 2σ amplitude. The one-particle-irreducible amplitude receives contributions from four one-loop diagrams and a counterterm:

$$\text{Diagram}_1 + \text{Diagram}_2 + \text{Diagram}_3 + \text{Diagram}_4 + \text{Diagram}_5. \quad (11.34)$$

It is convenient to write the counterterm vertex as

$$-i(2v^2 \delta_\lambda) - i(\delta_\mu + v^2 \delta_\lambda) + ip^2 \delta_Z. \quad (11.35)$$

In a general renormalization scheme, the σ mass will also be shifted by the tadpole diagrams (and their counterterm):

$$\text{Diagram}_1 + \text{Diagram}_2 + \text{Diagram}_5. \quad (11.36)$$

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

$$\text{Diagram}_1 \sim 18i\lambda^2 v^2 \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2},$$

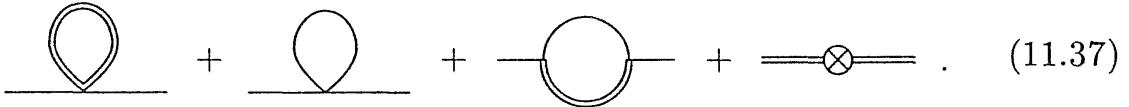
and for the second diagram

$$\text{Diagram}_2 \sim 2i\lambda^2 v^2 (N-1) \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2}.$$

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 finite at the one-loop level, just as in ordinary ϕ^4 theory.

There remains only one potentially divergent amplitude—the $\pi\pi$ amplitude:



In analogy with (11.31), the first diagram is

$$\frac{\text{Diagram 1}}{i} = \frac{1}{2}(-2i\lambda\delta^{ij}) \int \frac{d^d k}{(2\pi)^d} \frac{i}{k^2 - 2\mu^2} = -i\lambda\delta^{ij} \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{1}{2\mu^2}\right)^{1-\frac{d}{2}}.$$

The second diagram is quite similar. As in (11.32), it is useful to introduce a small pion mass as an infrared regulator.

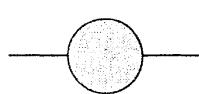
$$\begin{aligned} \frac{\text{Diagram 2}}{i} &= \frac{1}{2}(-2i\lambda(\delta^{ij}\delta^{kk} + \delta^{ik}\delta^{jk} + \delta^{ik}\delta^{jk})) \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 - \zeta^2} \\ &= -i\lambda(N+1)\delta^{ij} \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{1}{\zeta^2}\right)^{1-\frac{d}{2}}. \end{aligned}$$

The third diagram is given by

$$\begin{aligned} \frac{\text{Diagram 3}}{i} &= (-2i\lambda v\delta^{ik})(-2i\lambda v\delta^{kj}) \int \frac{d^d k}{(2\pi)^d} \frac{i}{k^2 - \zeta^2} \frac{i}{(k+p)^2 - 2\mu^2} \\ &= 4i\lambda^2 v^2 \delta^{ij} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \int_0^2 dx \left(\frac{1}{2\mu^2 x + (1-x)\zeta^2 - p^2 x(1-x)} \right)^{2-\frac{d}{2}}. \end{aligned}$$

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



$$\begin{aligned}
 (-i\lambda\delta^{ij}) \Big|_{p=0} = & \left\{ \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{1}{(2\mu^2)^{1-d/2}} + \frac{N+1}{(\zeta^2)^{1-d/2}} \right) \right. \\
 & - 4\lambda v^2 \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \int_0^2 dx \left(\frac{1}{2\mu^2 x + \zeta^2(1-x)} \right)^{2-\frac{d}{2}} \\
 & \left. - \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left(\frac{3}{(2\mu^2)^{1-d/2}} + \frac{N-1}{(\zeta^2)^{1-d/2}} \right) \right\}. \tag{11.38}
 \end{aligned}$$

It is not hard to simplify this expression. The first and third lines can be combined to give

$$2\lambda\delta^{ij} \frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}} \left[\frac{1}{(\zeta^2)^{1-d/2}} - \frac{1}{(2\mu^2)^{1-d/2}} \right].$$

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

$$2\lambda\delta^{ij} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} \frac{1}{1-d/2} \left[\frac{\zeta^2}{(\zeta^2)^{2-d/2}} - \frac{2\mu^2}{(2\mu^2)^{2-d/2}} \right]. \tag{11.39}$$

The first term vanishes for $d > 2$ and $\zeta \rightarrow 0$, 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 $\zeta \rightarrow 0$. 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

$$Z(H) = e^{-\beta F(H)} = \int \mathcal{D}s \exp\left[-\beta \int dx (\mathcal{H}[s] - Hs(x))\right], \quad (11.40)$$

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

$F(H)$:

$$\begin{aligned}
 -\frac{\partial F}{\partial H} \bigg|_{\beta \text{ fixed}} &= -\frac{1}{\beta} \frac{\partial}{\partial H} \log Z \\
 &= \frac{1}{Z} \int dx \int \mathcal{D}s s(x) \exp \left[-\beta \int dx (\mathcal{H}[s] - Hs) \right] \\
 &= \int dx \langle s(x) \rangle \equiv M.
 \end{aligned} \tag{11.41}$$

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

$$Z[J] = e^{-iE[J]} = \int \mathcal{D}\phi \exp \left[i \int d^4x (\mathcal{L}[\phi] + J\phi) \right]. \tag{11.43}$$

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 $J(x)$:

$$\frac{\delta}{\delta J(x)} E[J] = i \frac{\delta}{\delta J(x)} \log Z = - \frac{\int \mathcal{D}\phi e^{i \int (\mathcal{L} + J\phi)} \phi(x)}{\int \mathcal{D}\phi e^{i \int (\mathcal{L} + J\phi)}}. \quad (11.44)$$

We abbreviate this relation as

$$\frac{\delta}{\delta J(x)} E[J] = - \langle \Omega | \phi(x) | \Omega \rangle_J; \quad (11.45)$$

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 $\phi_{\text{cl}}(x)$, called the *classical field*, by

$$\phi_{\text{cl}}(x) = \langle \Omega | \phi(x) | \Omega \rangle_J. \quad (11.46)$$

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 $\phi_{\text{cl}}(x)$ 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 $E[J]$:

$$\Gamma[\phi_{\text{cl}}] \equiv -E[J] - \int d^4y J(y) \phi_{\text{cl}}(y). \quad (11.47)$$

This quantity is known as the *effective action*. In analogy with Eq. (11.42), we can now compute

$$\begin{aligned} \frac{\delta}{\delta \phi_{\text{cl}}(x)} \Gamma[\phi_{\text{cl}}] &= - \frac{\delta}{\delta \phi_{\text{cl}}(x)} E[J] - \int d^4y \frac{\delta J(y)}{\delta \phi_{\text{cl}}(x)} \phi_{\text{cl}}(y) - J(x) \\ &= - \int d^4y \frac{\delta J(y)}{\delta \phi_{\text{cl}}(x)} \frac{\delta E[J]}{\delta J(y)} - \int d^4y \frac{\delta J(y)}{\delta \phi_{\text{cl}}(x)} \phi_{\text{cl}}(y) - J(x) \\ &= -J(x). \end{aligned} \quad (11.48)$$

In the last step we have used Eq. (11.44).

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

$$\frac{\delta}{\delta \phi_{\text{cl}}(x)} \Gamma[\phi_{\text{cl}}] = 0. \quad (11.49)$$

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

$$\Gamma[\phi_{\text{cl}}] = -(VT) \cdot V_{\text{eff}}(\phi_{\text{cl}}). \quad (11.50)$$

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

$$\frac{\partial}{\partial \phi_{\text{cl}}} V_{\text{eff}}(\phi_{\text{cl}}) = 0. \quad (11.51)$$

*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 :

$$\phi_{\text{cl}} = x\phi_1 + (1 - x)\phi_3, \quad 0 < x < 1. \quad (11.52)$$

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

$$V_{\text{eff}}(\phi_{\text{cl}}) = xV_{\text{eff}}(\phi_1) + (1 - x)V_{\text{eff}}(\phi_3), \quad (11.53)$$

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).

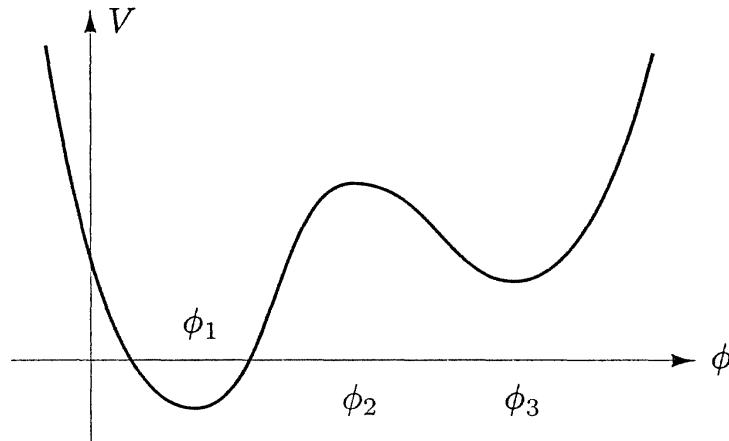


Figure 11.6. A possible form for the effective potential in a scalar field theory. The extrema of the effective potential occur at the points $\phi_{\text{cl}} = \phi_1, \phi_2, \phi_3$. The true vacuum state is the one corresponding to ϕ_1 . The state ϕ_2 is unstable. The state ϕ_3 is metastable, but it can decay to ϕ_1 by quantum-mechanical tunneling.

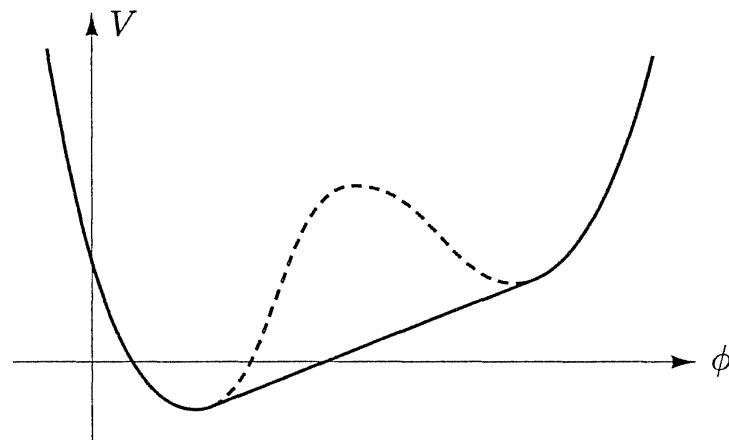


Figure 11.7. Exact convex form of the effective potential for the system of Fig. 11.6.

type shown in Fig. 11.6. The Maxwell construction must be performed by hand to yield the final form of $V_{\text{eff}}(\phi_{\text{cl}})$. 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 $V_{\text{eff}}(\phi_{\text{cl}})$. 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:

$$\mathcal{L} = \mathcal{L}_1 + \delta\mathcal{L}. \quad (11.54)$$

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:

$$\left. \frac{\delta\mathcal{L}}{\delta\phi} \right|_{\phi=\phi_{\text{cl}}} + J(x) = 0 \quad (\text{to lowest order}).$$

Let us define $J_1(x)$ to be whatever function satisfies this equation exactly, when $\mathcal{L} = \mathcal{L}_1$:

$$\left. \frac{\delta\mathcal{L}_1}{\delta\phi} \right|_{\phi=\phi_{\text{cl}}} + J_1(x) = 0 \quad (\text{exactly}). \quad (11.55)$$

We will think of the difference between J and J_1 as a counterterm, analogous to $\delta\mathcal{L}$, so we write

$$J(x) = J_1(x) + \delta J(x), \quad (11.56)$$

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

$$e^{-iE[J]} = \int \mathcal{D}\phi e^{i \int d^4x (\mathcal{L}_1[\phi] + J_1\phi)} e^{i \int d^4x (\delta\mathcal{L}[\phi] + \delta J\phi)}. \quad (11.57)$$

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. D* **9**, 1686 (1974).

$\phi(x) = \phi_{\text{cl}}(x) + \eta(x)$. This exponent takes the form

$$\begin{aligned} \int d^4x(\mathcal{L}_1 + J_1\phi) &= \int d^4x(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) + \int d^4x\eta(x)\left(\frac{\delta\mathcal{L}_1}{\delta\phi} + J_1\right) \\ &\quad + \frac{1}{2}\int d^4x d^4y\eta(x)\eta(y)\frac{\delta^2\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)} \\ &\quad + \frac{1}{3!}\int d^4x d^4y d^4z\eta(x)\eta(y)\eta(z)\frac{\delta^3\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)\delta\phi(z)} + \dots, \end{aligned} \quad (11.58)$$

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:

$$\begin{aligned} &\int \mathcal{D}\eta \exp\left[i\left(\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) + \frac{1}{2}\int\eta\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\eta\right)\right] \\ &= \exp\left[i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}})\right] \cdot \left(\det\left[-\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right]\right)^{-1/2}. \end{aligned} \quad (11.59)$$

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

$$-i\left(\frac{\delta^2\mathcal{L}}{\delta\phi\delta\phi}\right)^{-1} \quad (11.60)$$

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

$$(\delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}}) + (\delta\mathcal{L}[\phi_{\text{cl}} + \eta] - \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\eta). \quad (11.61)$$

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 $E[J]$:

$$\begin{aligned} -iE[J] = & i \int d^4x (\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) - \frac{1}{2} \log \det \left[-\frac{\delta^2 \mathcal{L}_1}{\delta \phi \delta \phi} \right] \\ & + (\text{connected diagrams}) + i \int d^4x (\delta \mathcal{L}[\phi_{\text{cl}}] + \delta J \phi_{\text{cl}}). \end{aligned} \quad (11.62)$$

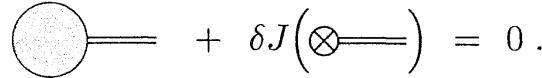
From this equation, Γ follows directly: Using $J_1 + \delta J = J$ and the Legendre transform (11.47), we find

$$\begin{aligned} \Gamma[\phi_{\text{cl}}] = & \int d^4x \mathcal{L}_1[\phi_{\text{cl}}] + \frac{i}{2} \log \det \left[-\frac{\delta^2 \mathcal{L}_1}{\delta \phi \delta \phi} \right] \\ & - i \cdot (\text{connected diagrams}) + \int d^4x \delta \mathcal{L}[\phi_{\text{cl}}]. \end{aligned} \quad (11.63)$$

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



$$+ \delta J \left(\text{circle with cross inside} \right) = 0. \quad (11.64)$$

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 $\Gamma[\phi_{\text{cl}}]$. Let us now clarify the meaning of this equation, and also put it to some good use, by computing $\Gamma[\phi_{\text{cl}}]$ 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):

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi^i)^2 + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}[(\phi^i)^2]^2. \quad (11.65)$$

Expand about the classical field: $\phi^i = \phi_{\text{cl}}^i + \eta^i$. 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 integration. 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

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\mu^2(\phi_{\text{cl}}^i)^2 - \frac{\lambda}{4}[(\phi_{\text{cl}}^i)^2]^2 + (\mu^2 - \lambda(\phi_{\text{cl}}^i)^2)\eta \\ & + \frac{1}{2}(\partial_\mu \eta^i)^2 + \frac{1}{2}\mu^2(\eta^i)^2 - \frac{\lambda}{2}[(\phi_{\text{cl}}^i)^2(\eta^i)^2 + 2(\phi_{\text{cl}}^i \eta^i)^2] + \dots \end{aligned} \quad (11.66)$$

According to Eq. (11.63), we should drop the term linear in η .

From the terms quadratic in η , we can read off

$$\frac{\delta^2 \mathcal{L}}{\delta \phi^i \delta \phi^j} = -\partial^2 \delta^{ij} + \mu^2 \delta^{ij} - \lambda[(\phi_{\text{cl}}^k)^2 \delta^{ij} + 2\phi_{\text{cl}}^i \phi_{\text{cl}}^j]. \quad (11.67)$$

Notice that this object has the general form of a Klein-Gordon operator. To clarify this relation, let us orient the coordinates so that ϕ_{cl}^i points in the N th direction,

$$\phi_{\text{cl}}^i = (0, 0, \dots, 0, \phi_{\text{cl}}), \quad (11.68)$$

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

$$m_i^2 = \begin{cases} \lambda\phi_{\text{cl}}^2 - \mu^2 & \text{acting on } \eta^1, \dots, \eta^{N-1}; \\ 3\lambda\phi_{\text{cl}}^2 - \mu^2 & \text{acting on } \eta^N. \end{cases} \quad (11.69)$$

The functional determinant in Eq. (11.63) is the product of the determinants of these Klein-Gordon operators:

$$\det \frac{\delta\mathcal{L}}{\delta\phi\delta\phi} = [\det(\partial^2 + (\lambda\phi_{\text{cl}}^2 - \mu^2))]^{N-1} [\det(\partial^2 + (3\lambda\phi_{\text{cl}}^2 - \mu^2))]. \quad (11.70)$$

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

$$\log \det(\partial^2 + m^2) = \text{Tr} \log(\partial^2 + m^2).$$

Now evaluate the trace of the operator as the sum of its eigenvalues:

$$\begin{aligned} \text{Tr} \log(\partial^2 + m^2) &= \sum_k \log(-k^2 + m^2) \\ &= (VT) \int \frac{d^4 k}{(2\pi)^4} \log(-k^2 + m^2). \end{aligned} \quad (11.71)$$

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 $\Gamma[\phi_{\text{cl}}]$. This manipulation gives an integral that can be evaluated in dimensional regularization after a Wick rotation:

$$\begin{aligned} \int \frac{d^4 k}{(2\pi)^4} \log(-k^2 + m^2) &= i \int \frac{d^4 k_E}{(2\pi)^4} \log(k_E^2 + m^2) \\ &= -i \frac{\partial}{\partial \alpha} \int \frac{d^4 k_E}{(2\pi)^4} \frac{1}{(k_E^2 + m^2)^\alpha} \Big|_{\alpha=0} \\ &= -i \frac{\partial}{\partial \alpha} \left(\frac{1}{(4\pi)^{d/2}} \frac{\Gamma(\alpha - \frac{d}{2})}{\Gamma(\alpha)} \frac{1}{(m^2)^{\alpha - d/2}} \right) \Big|_{\alpha=0} \\ &= -i \frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}} \frac{1}{(m^2)^{-d/2}}. \end{aligned} \quad (11.72)$$

In the last line, we have used $\Gamma(\alpha) \rightarrow 1/\alpha$ as $\alpha \rightarrow 0$. Thus,

$$\frac{1}{(VT)} \log \det(\partial^2 + m^2) = -i \frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}} (m^2)^{d/2}. \quad (11.73)$$

Using this result to evaluate the determinant in Eq. (11.63), and choosing

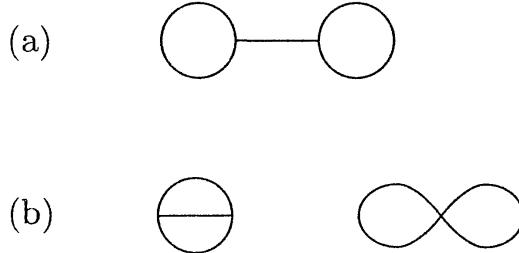


Figure 11.8. Feynman diagrams contributing to the evaluation of the effective potential of the $O(N)$ linear sigma model: (a) a diagram that is removed by (11.64); (b) the first nonzero diagrammatic corrections.

the counterterm Lagrangian as in Eq. (11.14), we find

$$\begin{aligned}
 V_{\text{eff}}(\phi) &= -\frac{1}{(VT)} \Gamma[\phi_{\text{cl}}] \\
 &= -\frac{1}{2} \mu^2 \phi_{\text{cl}}^2 + \frac{\lambda}{4} \phi_{\text{cl}}^4 \\
 &\quad - \frac{1}{2} \frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}} [(N-1)(\lambda\phi_{\text{cl}}^2 - \mu^2)^{d/2} + (3\lambda\phi_{\text{cl}}^2 - \mu^2)^{d/2}] \\
 &\quad + \frac{1}{2} \delta_\mu \phi_{\text{cl}}^2 + \frac{1}{4} \delta_\lambda \phi_{\text{cl}}^4. \tag{11.74}
 \end{aligned}$$

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 $d \rightarrow 2$ if we set

$$\delta_\mu = -\lambda(N+2) \frac{\Gamma(1-\frac{d}{2})}{(4\pi)} + \text{finite.}$$

The expression is finite as $d \rightarrow 4$ if we set

$$\begin{aligned} \delta_\mu &= -\lambda\mu^2(N+2) \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} + \text{finite;} \\ \delta_\lambda &= \lambda^2(N+8) \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^2} + \text{finite.} \end{aligned} \quad (11.76)$$

These expressions agree with our earlier results from Section 11.2, Eqs. (11.33) and (11.22), in the limits $d \rightarrow 2$ and $d \rightarrow 4$ 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

$$\begin{aligned} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}(m^2)^{2-d/2}} &= \frac{1}{(4\pi)^2} \left(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log(m^2) \right) \\ &\rightarrow \frac{1}{(4\pi)^2} (-\log(m^2/M^2)), \end{aligned} \quad (11.77)$$

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):

$$\begin{aligned} \frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}} (m^2)^{d/2} &= \frac{1}{\frac{d}{2}(\frac{d}{2}-1)} \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}} (m^2)^{d/2} \\ &= \frac{m^4}{2(4\pi)^2} \left(\frac{2}{\epsilon} - \gamma + \log(4\pi) - \log(m^2) + \frac{3}{2} \right) \\ &\rightarrow \frac{m^4}{2(4\pi)^2} \left(-\log(m^2/M^2) + \frac{3}{2} \right). \end{aligned} \quad (11.78)$$

Modifying our result (11.74) in this way, we find

$$\begin{aligned} V_{\text{eff}} &= -\frac{1}{2}\mu^2\phi_{\text{cl}}^2 + \frac{\lambda}{4}\phi_{\text{cl}}^4 \\ &+ \frac{1}{4}\frac{1}{(4\pi)^2} \left((N-1)(\lambda\phi_{\text{cl}}^2 - \mu^2)^2 \left(\log[(\lambda\phi_{\text{cl}}^2 - \mu^2)/M^2] - \frac{3}{2} \right) \right. \\ &\quad \left. + (3\lambda\phi_{\text{cl}}^2 - \mu^2)^2 \left(\log[(3\lambda\phi_{\text{cl}}^2 - \mu^2)/M^2] - \frac{3}{2} \right) \right). \end{aligned} \quad (11.79)$$

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

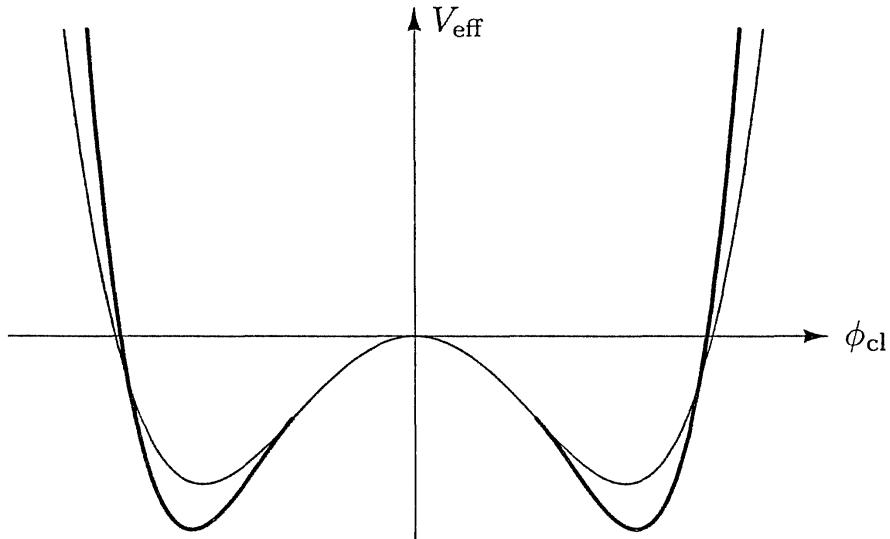


Figure 11.9. The effective potential for ϕ^4 theory ($N = 1$), with quantum corrections included as in Eq. (11.79). The lighter-weight curve shows the classical potential energy, for comparison.

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

$$\begin{aligned} \lambda &\rightarrow \lambda + \frac{\lambda^2}{(4\pi)^2} (N + 8) \cdot \frac{\delta M^2}{M^2}, \\ \mu^2 &\rightarrow \mu^2 - \frac{\lambda \mu^2}{(4\pi)^2} (N + 2) \cdot \frac{\delta M^2}{M^2}. \end{aligned} \quad (11.80)$$

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 $\mu^2 \rightarrow 0$; then Eq. (11.79) takes the form

$$\begin{aligned} V_{\text{eff}} &= \frac{\lambda}{4} \phi_{\text{cl}}^4 + \frac{1}{4} \frac{\lambda^2}{(4\pi)^2} \phi_{\text{cl}}^4 \left(\log(\lambda \phi_{\text{cl}}^2 / M^2) - \frac{3}{2} \right) \\ &= \frac{1}{4} \phi_{\text{cl}}^4 \left(\lambda + \frac{\lambda^2}{(4\pi)^2} (N + 8) \left(\log(\lambda \phi_{\text{cl}}^2 / M^2) - \frac{3}{2} \right) \right). \end{aligned} \quad (11.81)$$

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 the very small value

$$\phi_{\text{cl}}^2 = e^{3/2} \frac{M^2}{\lambda} \cdot \exp \left[-\frac{(4\pi)^2}{(N+8)\lambda} \right],$$

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

$$\begin{aligned} \frac{\delta^2 E[J]}{\delta J(x) \delta J(y)} &= -\frac{i}{Z} \int \mathcal{D}\phi e^{i \int (\mathcal{L} + J\phi)} \phi(x) \phi(y) \\ &\quad + \frac{i}{Z^2} \int \mathcal{D}\phi e^{i \int (\mathcal{L} + J\phi)} \phi(x) \cdot \int \mathcal{D}\phi e^{i \int (\mathcal{L} + J\phi)} \phi(y) \\ &= -i \left[\langle \phi(x) \phi(y) \rangle - \langle \phi(x) \rangle \langle \phi(y) \rangle \right]. \end{aligned} \quad (11.82)$$