CHAPTER 1

Quantum Field Theory Basics

Introduction

This chapter is devoted to basic aspects of quantum field theory, ranging from the foundations to perturbation theory and renormalization, and is limited to the canonical formalism (functional methods are treated in Chapter 2) and to the traditional workflow (Lagrangian $\longrightarrow$ Feynman rules $\longrightarrow$ time-ordered products of fields $\longrightarrow$ scattering amplitudes) for the calculation of scattering amplitudes (the spinor-helicity formalism and on-shell recursion are considered in Chapter 4). The problems of this chapter deal with questions in scalar field theory and quantum electrodynamics, while non-Abelian gauge theories are discussed in Chapter 3.

Non-interacting Field Theory

A non-interacting field theory may be defined by a quadratic Lagrangian. In the simplest case of a scalar field theory, it reads

$$\mathcal{L} \equiv \int d^3\boldsymbol{x} \left\{ \tfrac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \tfrac{1}{2}m^2\phi^2 \right\}. \tag{1.1}$$

Such a Lagrangian defines a dynamical system with infinitely many degrees of freedom, corresponding to the values taken by $\phi(x)$ at every point $\boldsymbol{x}$ of space. The momentum canonically conjugate to $\phi(x)$ is given by

$$\Pi(x) \equiv \frac{\partial \mathcal{L}}{\partial(\partial^0\phi(x))} = \partial^0\phi(x), \tag{1.2}$$

which leads to the Hamiltonian

$$\mathcal{H} \equiv \int d^3\boldsymbol{x}\, \Pi(x)\partial^0\phi(x) - \mathcal{L} = \int d^3\boldsymbol{x} \left\{ \tfrac{1}{2}\Pi^2 + \tfrac{1}{2}(\boldsymbol{\nabla}\phi)^2 + \tfrac{1}{2}m^2\phi^2 \right\}. \tag{1.3}$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

From the Hamiltonian or Lagrangian, one obtains the equation of motion of the field, which in the present example reads

$$\left(\square_x + m^2\right)\phi(x) = 0, \tag{1.4}$$

known as the *Klein–Gordon equation*. Generic real solutions of this linear equation are superpositions of plane waves:

$$\phi(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \left\{\alpha^*_\mathbf{k}\, e^{-ik\cdot x} + \alpha_\mathbf{k}\, e^{ik\cdot x}\right\}, \tag{1.5}$$

where $E_\mathbf{k} \equiv \sqrt{\mathbf{p}^2 + m^2}$ is the dispersion relation associated with the wave equation (1.4), and $\alpha_\mathbf{k}$ is a function of momentum that depends on the boundary conditions imposed on the solution.

Canonical quantization consists in promoting the coefficients $\alpha_\mathbf{k}, \alpha^*_\mathbf{k}$ into annihilation and creation operators $a_\mathbf{k}, a^\dagger_\mathbf{k}$ that obey the following commutation relations:

$$\left[a_\mathbf{p}, a^\dagger_\mathbf{q}\right] \equiv (2\pi)^3 2E_\mathbf{p}\, \delta(\mathbf{p}-\mathbf{q}). \tag{1.6}$$

The normalization in Eqs. (1.5) and (1.6) is chosen so that $\left[\mathcal{H}, a^\dagger_\mathbf{p}\right] = E_\mathbf{p} a^\dagger_\mathbf{p}$ and $\left[\mathcal{H}, a_\mathbf{p}\right] = -E_\mathbf{p} a_\mathbf{p}$, which means that $a^\dagger_\mathbf{p}$ increases the energy of the system by $E_\mathbf{p}$ while $a_\mathbf{p}$ decreases it by the same amount. As a consequence, this setup describes a collection of non-interacting particles. The commutation relation (1.6) implies the following equal-time commutation relation between the field operator and its conjugate momentum:

$$\left[\phi(x), \Pi(y)\right] \underset{x^0=y^0}{=} i\,\delta(\mathbf{x}-\mathbf{y}), \tag{1.7}$$

which one may view as the quantum version of the classical Poisson bracket between a coordinate and its conjugate momentum.

Interacting Field Theory and Interaction Representation

Interactions are introduced via terms of degree higher than two in the Lagrangian:

$$\mathcal{L} \equiv \int d^3\mathbf{x} \left\{ \underbrace{\tfrac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \tfrac{1}{2}m^2\phi^2}_{\mathcal{L}_0,\ \text{non-interacting theory}} \underbrace{-V(\phi)}_{\text{interactions}} \right\}. \tag{1.8}$$

(In order to have a causal theory, the potential $V(\phi)$ must be a local function of the field $\phi(x)$; see Problem 4.) In the presence of interactions, the Klein–Gordon equation of motion becomes

$$\left(\square_x + m^2\right)\phi(x) + V'(\phi(x)) = 0. \tag{1.9}$$

Since the degree of $V(\phi)$ is higher than two, this equation is non-linear, which induces a mixing between the Fourier modes of the field and prevents writing its solutions as superpositions of plane waves.

By assuming that the interactions are turned off at large times, $x^0 \to \pm\infty$, we may define free fields ϕ_{in} and ϕ_{out} that coincide with the interacting field ϕ of the Heisenberg representation, respectively when $x^0 \to -\infty$ and $x^0 \to +\infty$. For instance, ϕ and $\phi_{\rm in}$ are related by

$$\phi(x) = U(-\infty, x^0)\, \phi_{\rm in}(x)\, U(x^0, -\infty),$$
$$U(t_2, t_1) \equiv T \exp\left(-i \int_{t_1}^{t_2} dx^0 d^3\boldsymbol{x}\; V(\phi_{\rm in}(x))\right). \tag{1.10}$$

In this representation, the time dependence of the field $\phi(x)$ is split into a trivial one that comes from the free field $\phi_{\rm in}$ and the time evolution operator U that depends on the interactions. Since they are free fields obeying Eq. (1.4), $\phi_{\rm in}$ and $\phi_{\rm out}$ can be written as superpositions of plane waves, with coefficients $a_{\boldsymbol{p},{\rm in}}, a^\dagger_{\boldsymbol{p},{\rm in}}$ and $a_{\boldsymbol{p},{\rm out}}, a^\dagger_{\boldsymbol{p},{\rm out}}$, respectively. These two sets of creation and annihilation operators define two towers of *Fock states*, i.e., states with a definite particle content at $x^0 = -\infty$ and $x^0 = +\infty$, respectively.

Lehmann–Symanzik–Zimmermann Reduction Formulas

Experimentally measurable quantities, such as cross-sections, may be related to correlation functions of the field operator as follows. An intermediate step involves the transition amplitudes between *in* and *out* states,

$$\left\langle \boldsymbol{q}_1 \cdots \boldsymbol{q}_{n\,{\rm out}} \middle| \boldsymbol{p}_1 \cdots \boldsymbol{p}_{m\,{\rm in}} \right\rangle \equiv (2\pi)^4 \delta\big(\sum_i p_i - \sum_j q_j\big)\; \mathcal{T}(\boldsymbol{q}_{1\cdots n} | \boldsymbol{p}_{1\cdots m}), \tag{1.11}$$

in terms of which a cross-section in the center of momentum frame is given by

$$\sigma_{12\to 1\cdots n}\Big|_{\substack{\text{center of}\\ \text{momentum}}} = \frac{1}{4\sqrt{s}\,|\boldsymbol{p}_1|} \int d\Gamma_n(\boldsymbol{p}_{1,2}) \left|\mathcal{T}(\boldsymbol{q}_{1,\cdots,n}|\boldsymbol{p}_{1,2})\right|^2, \tag{1.12}$$
$$\text{with}\;\; d\Gamma_n(\boldsymbol{p}_{1,2}) \equiv \prod_j \frac{d^3\boldsymbol{q}_j}{(2\pi)^3 2E_{q_j}}\; (2\pi)^4 \delta\big(p_1 + p_2 - \sum_j q_j\big), \quad s \equiv (p_1 + p_2)^2.$$

In turn, the transition amplitudes from *in* to *out* states are expressed in terms of expectation values of time-ordered products of field operators by the Lehmann–Symanzik–Zimmermann (LSZ) reduction formulas:

$$\left\langle \boldsymbol{q}_1 \cdots \boldsymbol{q}_{n\;{\rm out}} \middle| \boldsymbol{p}_1 \cdots \boldsymbol{p}_{m\;{\rm in}} \right\rangle = \frac{i^{m+n}}{Z^{\frac{m+n}{2}}} \int \prod_{i=1}^{m} d^4x_j\; e^{-ip_i\cdot x_i}\; (\Box_{x_i} + m^2)$$
$$\times \int \prod_{j=1}^{n} d^4y_j\; e^{iq_j\cdot y_j}\; (\Box_{y_j} + m^2)\, \left\langle 0_{\rm out} \middle| T\, \phi(x_1)\cdots\phi(x_m)\phi(y_1)\cdots\phi(y_n) \middle| 0_{\rm in} \right\rangle, \tag{1.13}$$

where Z is the wavefunction renormalization factor.

Generating Functional, Feynman Propagator

The calculation of expectation values of time-ordered products of field operators is usually organized by encapsulating them in a generating functional

$$\left\langle 0_{\rm out} \middle| T\, \phi(x_1)\cdots\phi(x_n) \middle| 0_{\rm in} \right\rangle = \frac{\delta^n Z[j]}{i\delta j(x_1)\cdots i\delta j(x_n)}\bigg|_{j\equiv 0}, \tag{1.14}$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

$$\text{with } Z[j] \equiv \big\langle 0_{\text{out}} \big| T \exp i \int d^4x\, j(x)\phi(x) \big| 0_{\text{in}} \big\rangle \tag{1.15}$$

$$= \exp\left(-i \int d^4x\, V\left(\frac{\delta}{i\delta j(x)} \right) \right) \underbrace{\big\langle 0_{\text{in}} \big| T \exp i \int d^4x\, j(x)\phi_{\text{in}}(x) \big| 0_{\text{in}} \big\rangle}_{Z_0[j],\quad \text{non-interacting theory}} . \tag{1.16}$$

The last factor, the generating functional of the non-interacting theory, is a Gaussian in the auxiliary source j:

$$Z_0[j] = \exp\left(-\frac{1}{2} \int d^4x d^4y\, j(x) j(y) G_F^0(x,y) \right), \tag{1.17}$$

where $G_F^0(x,y)$ is the free *Feynman propagator*, which can be expressed in various equivalent ways:

$$G_F^0(x,y) = \big\langle 0_{\text{in}} \big| T\, \phi_{\text{in}}(x)\phi_{\text{in}}(y) \big| 0_{\text{in}} \big\rangle \tag{1.18}$$

$$= \int \frac{d^3p}{(2\pi)^3 2E_p} \left(\theta(x^0 - y^0)\, e^{-ip\cdot(x-y)} + \theta(y^0 - x^0)\, e^{+ip\cdot(x-y)} \right), \tag{1.19}$$

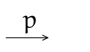
$$G_F^0(p) = \frac{i}{p^2 - m^2 + i0^+}. \tag{1.20}$$

Feynman Rules of Scalar Field Theory

The effect of interactions can be calculated order-by-order by expanding the first exponential in Eq. (1.16). The successive terms of this expansion are obtained from a diagrammatic expansion, where each diagram is converted into a formula by means of *Feynman rules*. Below we list these rules in momentum space, for a scalar field theory:

1. Draw all the graphs with as many external lines as field operators in the correlation function, and a number of vertices equal to the desired order. The vertices allowed in these graphs must have valences equal to the degrees of the terms in $V(\phi)$. Graphs with multiple connected components need not be considered in the calculation of scattering amplitudes.

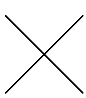
2. A 4-momentum k is assigned to each internal line of the graph, and the associated Feynman rule is a free propagator $G_F^0(k)$:



$$= \frac{i}{p^2 - m^2 + i0^+}.$$

No propagator should be assigned to the external lines of a graph when calculating a scattering amplitude (because of the factors $\square + m^2$ in the reduction formulas).

3. For an interaction $\frac{\lambda}{n!}\phi^n$, each vertex of valence n brings a factor $-i\lambda(2\pi)^4\delta\big(\sum_i k_i\big)$, where the k_i are the momenta incoming into this vertex:



$$= -i\lambda.$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

3. All the internal momenta that are not constrained by the delta functions at the vertices should be integrated over with a measure $d^4k/(2\pi)^4$. In a connected graph with n_I internal lines and n_V vertices, there are $n_L = n_I - n_V + 1$ of them, which is also the number of *loops* in the graph.

4. Each graph must be weighted by a *symmetry factor*, defined as the inverse of the order of the discrete symmetry group of the graph (assuming interaction terms properly symmetrized, as in $V(\phi) = \phi^n/n!$).

Dimensional Regularization

The momentum integrals that correspond to loops in Feynman diagrams may be divergent at large momentum. Convergence may be assessed from the *superficial degree of divergence* of a graph, $\omega(\mathcal{G}) \equiv 4n_L - 2n_I$ for a graph with n_L loops and n_I internal lines in a scalar field theory with quartic coupling in four spacetime dimensions: the graph $\mathcal{G}$ is convergent if $\omega(\mathcal{G}) < 0$ and the superficial degree of divergence of all its subgraphs is negative as well. In order to safely manipulate possibly divergent loop integrals, the first step is to introduce a *regularization*, i.e., a modification of the Feynman rules such that all loop integrals become well defined. Many regularization methods are possible: Pauli–Villars subtraction, lattice discretization, momentum cutoff, dimensional regularization.

Dimensional regularization, based on the observation that loop integrals calculated in an arbitrary number D of dimensions have an analytical continuation which is well defined at all D's except a discrete set of values, is particularly adapted to analytical calculations. With this regularization scheme, some common (Euclidean) loop integrals are given by

$$\begin{aligned}
\int \frac{d^D k_E}{(2\pi)^D} \frac{1}{(k_E^2 + \Delta)^n} &= \frac{\Delta^{\frac{D}{2}-n}}{(4\pi)^{\frac{D}{2}}} \frac{\Gamma(n - \frac{D}{2})}{\Gamma(n)}, \\
\int \frac{d^D k_E}{(2\pi)^D} \frac{k_E^\mu k_E^\nu}{(k_E^2 + \Delta)^n} &= \frac{g^{\mu\nu}}{2} \frac{\Delta^{\frac{D}{2}+1-n}}{(4\pi)^{\frac{D}{2}}} \frac{\Gamma(n - 1 - \frac{D}{2})}{\Gamma(n)}, \\
\int \frac{d^D k_E}{(2\pi)^D} \frac{k_E^\mu k_E^\nu k_E^\rho k_E^\sigma}{(k_E^2 + \Delta)^n} &= \frac{g^{\mu\nu}g^{\rho\sigma} + g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}}{4} \frac{\Delta^{\frac{D}{2}+2-n}}{(4\pi)^{\frac{D}{2}}} \frac{\Gamma(n - 2 - \frac{D}{2})}{\Gamma(n)}, \\
\int \frac{d^D k_E}{(2\pi)^D} \frac{k_E^{\mu_1} \cdots k_E^{\mu_{2n+1}}}{(k_E^2 + \Delta)^n} &= 0.
\end{aligned} \tag{1.21}$$

The first of these equations is obtained by integration in D-dimensional spherical coordinates, and the subsequent equations follow from Lorentz invariance.

Renormalization

The list of correlation functions that exhibit ultraviolet divergences can be obtained from the superficial degree of divergence $\omega(\mathcal{G})$ (except in situations where a symmetry produces cancellations that cannot be anticipated by power counting). For a scalar field theory with a quartic coupling, one has $\omega(\mathcal{G}) = 4 - n_E + (D-4)n_L$ in D spacetime dimensions, where

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

n_E is the number of external points and n_L the number of loops. The *Weinberg convergence theorem* states that a graph is ultraviolet convergent if and only if the superficial degree of divergence of the graph, and of any of its subgraphs, is negative.

In $D = 4$ spacetime dimensions, $\omega(\mathcal{G})$ is negative for all correlation functions with $n_E > 4$ points, implying that only a finite number of correlation functions have intrinsic divergences. Moreover, these divergent correlation functions are the expectation values of the operators already present in the Lagrangian, $(\partial_\mu\phi)^2, \phi^2, \phi^4$. The divergences that appear in these functions can be subtracted order-by-order via a redefinition of their coefficients in the Lagrangian, i.e., Z (this one is usually not explicit in the bare Lagrangian because it is set to 1), m^2 and λ, respectively. Such a quantum field theory is called *renormalizable*.

In $D > 4$ dimensions, $\omega(\mathcal{G})$ increases with the number of loops at fixed n_E. This implies that any correlation function exhibits intrinsic ultraviolet divergences beyond a certain loop order. Removing these divergences would require that one adds arbitrarily many new terms in the Lagrangian, reducing considerably the predictive power of such a theory (but it may nevertheless be of some use in an effective sense, at low loop orders). It is called *non-renormalizable*.

When $D < 4$, the superficial degree of divergence of any correlation function eventually becomes negative after a certain loop order. These theories have a finite number of ultraviolet divergent Feynman graphs, whose calculation is sufficient to determined the renormalized Lagrangian once and for all. These theories are called *super-renormalizable*.

For general interactions in arbitrary dimensions, the above criteria can be expressed in terms of the mass dimension of the prefactor that accompanies the operator in the Lagrangian. The corresponding operator is renormalizable if the mass dimension of its coupling constant is zero, non-renormalizable if this dimension is negative, super-renormalizable if it is positive.

Renormalization Group

In a renormalized quantum field theory, one may still freely choose the *renormalization scale* μ at which the conditions that define the parameters of the renormalized Lagrangian (masses, couplings, etc.) are imposed. Physical results should not depend on this scale. The dependence of various renormalized quantities with respect to μ is controlled by the *Callan–Symanzik equations*, also known as *renormalization group* equations. For the renormalized n-point correlation function G_n, this equation reads

$$\Big(\underbrace{\mu\partial_\mu + \beta\partial_\lambda + \gamma_m\, m\partial_m}_{\equiv \mathcal{D}_\mu} + n\gamma\Big)\, G_n = 0, \tag{1.22}$$

$$\text{with} \;\; \gamma \equiv \frac{1}{2}\frac{\partial \ln(Z)}{\partial \ln(\mu)}, \quad \beta \equiv \frac{\partial \lambda}{\partial \ln(\mu)}, \quad \gamma_m \equiv \frac{\partial \ln(m)}{\partial \ln(\mu)} \tag{1.23}$$

(γ is called an *anomalous dimension*, and β is the β *function*). Physical quantities are invariant under the action of $\mathcal{D}_\mu$, i.e., under the simultaneous change of the scale μ and of the parameters Z, λ, m as prescribed by the above differential equations (the solutions $\lambda(\mu)$ and $m(\mu)$ are called the running coupling and running mass, respectively). The curves $(Z(\mu), \lambda(\mu), m(\mu))$ in the parameter space of the renormalized theory, along which physical quantities are invariant, define a vector field called the *renormalization flow*.

From the Callan–Symanzik equation satisfied by the propagator, $(\mathcal{D}_\mu + 2\gamma)G_2 = 0$, one obtains the corresponding flow equations for the pole mass m_P (defined from the value of p^2

at the pole of the propagator) and for the residue Z at the pole:

$$\mathcal{D}_\mu m_p = 0, \quad (\mathcal{D}_\mu + 2\gamma) Z = 0. \tag{1.24}$$

Thus, a n-point scattering amplitude $\mathcal{A}_n \sim Z^{-n/2} G_n$ also satisfies $\mathcal{D}_\mu \mathcal{A}_n = 0$. Amputated correlation functions $\Gamma_n \equiv (G_2)^{-n} G_n$ obey $(\mathcal{D}_\mu - n\gamma)\Gamma_n = 0$.

Spin-1/2 Fields

The representation of the Lorentz algebra of lowest even dimension is defined by the generators $M^{\mu\nu}_{1/2} \equiv \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right]$, where the γ^μ are the *Dirac* 4×4 *matrices*, which obey $\left\{\gamma^\mu, \gamma^\nu\right\} = 2\, g^{\mu\nu}$. Under a Lorentz transformation $\Lambda \equiv \exp\left(-\frac{i}{2}\omega_{\mu\nu} M^{\mu\nu}\right)$, a *Dirac spinor* is a four-component field that transforms as

$$\psi(x) \quad \to \quad \exp\left(-\tfrac{i}{2}\omega_{\mu\nu} M^{\mu\nu}_{1/2}\right) \psi(\Lambda^{-1} x). \tag{1.25}$$

In the absence of interactions, such a field obeys the – Lorentz invariant – *Dirac equation*,

$$\left(i\gamma^\mu \partial_\mu - m\right) \psi = 0, \tag{1.26}$$

which can be obtained as the equation of motion that results from the following Lagrangian:

$$\mathcal{L} = \overline{\psi} \left(i\gamma^\mu \partial_\mu - m\right) \psi, \quad \text{with } \overline{\psi} \equiv \psi^\dagger \gamma^0. \tag{1.27}$$

The canonical quantization of a free spinor (i.e., a solution of the Dirac equation (1.26)) consists in replacing its Fourier coefficients by creation and annihilation operators:

$$\psi(x) \equiv \sum_{s=\pm} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \left\{ d^\dagger_{s\mathbf{p}} v_s(\mathbf{p}) e^{+ip\cdot x} + b_{s\mathbf{p}} u_s(\mathbf{p}) e^{-ip\cdot x} \right\}. \tag{1.28}$$

Since ψ is not Hermitian, the two operators in this decomposition need not be mutual conjugates (except in the special case of Majorana fermions). The spinors u_s, v_s are a basis of free spinors in momentum space defined by

$$\left(\gamma^\mu p_\mu - m\right) u_s(\mathbf{p}) = 0, \quad \left(\gamma^\mu p_\mu + m\right) v_s(\mathbf{p}) = 0, \tag{1.29}$$

$$u^\dagger_r(\mathbf{p}) u_s(\mathbf{p}) = 2E_\mathbf{p} \delta_{rs}, \quad v^\dagger_r(\mathbf{p}) v_s(\mathbf{p}) = 2E_\mathbf{p} \delta_{rs}. \tag{1.30}$$

For the Hamiltonian of this system to have a well-defined ground state, these creation and annihilation operators must obey anti-commutation relations. The non-zero ones read

$$\{d_{s\mathbf{p}}, d^\dagger_{s'\mathbf{p}'}\} = \{b_{s\mathbf{p}}, b^\dagger_{s'\mathbf{p}'}\} = (2\pi)^3 2E_\mathbf{p} \delta_{ss'} \delta(\mathbf{p} - \mathbf{p}'), \tag{1.31}$$

or, equivalently,

$$\{\psi_\alpha(x), \psi^\dagger_\beta(y)\} \underset{x^0=y^0}{=} \delta_{\alpha\beta} \delta(\mathbf{x} - \mathbf{y}). \tag{1.32}$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

The Dirac Lagrangian has a $U(1)$ symmetry, $\psi \to e^{-i\alpha}\psi$, which by Noether's theorem leads to a conserved current $J^\mu \equiv \overline{\psi}\gamma^\mu\psi$ and conserved charge

$$Q \equiv \int d^3\mathbf{x}\; J^0 = \sum_{s=\pm}\int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}}\left(b^\dagger_{s\mathbf{p}} b_{s\mathbf{p}} - d^\dagger_{s\mathbf{p}} d_{s\mathbf{p}}\right). \tag{1.33}$$

Spin-1 Fields

A *vector field* $A^\mu(x)$ is a four-component field that transforms as $A^\mu(x) \to \Lambda^\mu{}_\nu\, A^\nu(\Lambda^{-1}x)$ under a Lorentz transformation Λ. The simplest such (massless) field is the electromagnetic field, whose Lagrangian reads

$$\mathcal{L} = -\tfrac{1}{4}\, F_{\mu\nu}F^{\mu\nu}, \quad \text{with } F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{1.34}$$

The corresponding equation of motion, $\partial_\mu F^{\mu\nu} = 0$, has several remarkable properties:

- *Gauge invariance*: for any function θ, $A^\mu - \partial^\mu\theta$ is a solution if A^μ is a solution.
- The field A^0 is not dynamical, but given by a constraint from the spatial components A^i.
- Only the transverse (i.e., transverse to the momentum k^i in Fourier space) components of A^i are constrained by the equation of motion.

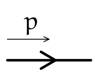
The unphysical redundancy due to gauge invariance is removed by imposing a *gauge condition* – e.g., $\partial_\mu A^\mu = 0$ (Lorenz gauge), $\partial_i A^i = 0$ (Coulomb gauge), $A^0 = 0$ (temporal gauge) – leaving only two independent dynamical solutions per Fourier mode. The quantization of the vector field A^μ amounts to replacing the coefficients in its Fourier decomposition by creation and annihilation operators:

$$A^\mu(x) \equiv \sum_{\lambda=1,2}\int \frac{d^3\mathbf{p}}{(2\pi)^3 2|\mathbf{p}|}\left\{a^\dagger_{\lambda\mathbf{p}}\,\epsilon^{\mu *}_\lambda(\mathbf{p})\, e^{ip\cdot x} + a_{\lambda\mathbf{p}}\,\epsilon^\mu_\lambda(\mathbf{p})\, e^{-ip\cdot x}\right\}, \tag{1.35}$$

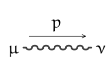
with the canonical commutation relation $\left[a_{\lambda\mathbf{p}}, a^\dagger_{\lambda'\mathbf{p}'}\right] = (2\pi)^3 2|\mathbf{p}|\delta_{\lambda\lambda'}\delta(\mathbf{p}-\mathbf{p}')$ and where the objects $\epsilon^\mu_\lambda(\mathbf{p})$ are *polarization vectors* that encode the Lorentz indices of a vector field of polarization λ and momentum $\mathbf{p}$. The polarization vectors may depend on the choice of gauge condition, but always satisfy $p_\mu \epsilon^\mu_\lambda(\mathbf{p}) = 0$.

Quantum Electrodynamics

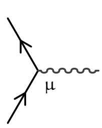
The conserved charge of the Dirac fermions can be interpreted as an electrical charge. Interactions between these fermions and photons are introduced by *minimal coupling*, i.e., by requesting that the modified Dirac Lagrangian is invariant under spacetime-dependent $U(1)$ transformations, $\psi(x) \to e^{-ie\theta(x)}\psi(x)$. This is achieved by replacing the ordinary derivative by a *covariant derivative*, $D_\mu \equiv \partial_\mu - ieA_\mu$. Perturbation theory in QED has the following Feynman rules:



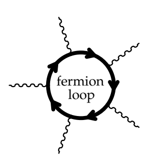
$= \dfrac{i(\not{p} + m)}{p^2 - m^2 + i0^+}$



$= \dfrac{i\,C^{\mu\nu}(p)}{p^2 + i0^+}$



$= -i\,e\,\gamma^\mu$



= (minus sign)

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

The numerator $C^{\mu\nu}$ in the photon propagator depends on the gauge fixing (for instance, $C^{\mu\nu} = -g^{\mu\nu}$ in Feynman gauge).

Ward–Takahashi Identity

A crucial property of QED amplitudes with external photons is the *Ward–Takahashi identity*, namely

$$p_\mu \, \Gamma^{\mu\cdots}(p, \dots) = 0, \tag{1.36}$$

where $\Gamma^{\mu\cdots}(p, \dots)$ is an amplitude amputated of its external propagators, containing a photon of momentum p with Lorentz index μ. The dots represent the other external lines, either photons or charged particles. The conditions of validity of this identity, which follows from the conservation of the electrical current, are the following:

- All the external lines corresponding to charged particles must be on-shell, and contracted in the appropriate spinors if they are fermions.
- The gauge fixing condition must be linear in the gauge potential, in order not to have three- and four-photon vertices.

The Ward–Takahashi identity plays a crucial role in ensuring that QED scattering amplitudes are gauge invariant, and that they fulfill the requirements of unitarity despite the presence of non-physical photon polarization in certain gauges.

Unitarity, the Optical Theorem and Cutkosky's Cutting Rules

The time evolution operator from $x^0 = -\infty$ to $x^0 = +\infty$ (also called the S-matrix) is unitary, $SS^\dagger = 1$. Writing it as $S \equiv 1 + iT$ to separate the interactions, this implies the *optical theorem*:

$$\mathrm{Im}\,\big\langle \alpha_{\rm in} \big| T \big| \alpha_{\rm in} \big\rangle = \frac{1}{2} \sum_{\text{states } \beta} \big| \big\langle \alpha_{\rm in} \big| T \big| \beta_{\rm in} \big\rangle \big|^2 .$$

This relation implies that the total probability of scattering from the state α to any state β (with at least one interaction) equals twice the imaginary part of the forward scattering amplitude $\alpha \to \alpha$. In perturbation theory, the imaginary part of a transition amplitude Γ can be obtained by means of *Cutkosky's cutting rules*:

$$\mathrm{Im}\,\Gamma = \frac{1}{2} \sum_{\text{cuts } \gamma} \big[\Gamma\big]_\gamma ,$$

where a cut is a fictitious line that divides the graph into two subgraphs, with at least one external leg on each side of the cut. A cut graph $\big[\Gamma\big]_\gamma$ is calculated with the following rules:

- Left of the cut: use the propagator $G^0_{++}(p) = \frac{i}{p^2 - m^2 + i0^+}$ and the vertex $-i\lambda$,
- Right of the cut: use the propagator $G^0_{--}(p) = \frac{-i}{p^2 - m^2 - i0^+}$ and the vertex $+i\lambda$,
- The propagators traversing the cut should be $G^0_{+-}(p) = 2\pi\,\theta(-p^0)\delta(p^2)$.

About the Problems of this Chapter

- **Problem 1** establishes a crucial relationship between the field operators ϕ (Heisenberg representation) and $\phi_{\rm in}$ (interaction representation), namely that the former obeys the interacting equation of motion if the latter obeys the free Klein–Gordon equation.
- In **Problem 2**, we derive an explicit form of the elements of the *little group* for massless particles. This is then used in **Problem 9** in order to show that, in a theory with massless spin-1 bosons, the Lorentz invariance of scattering amplitudes implies a property that may be viewed as a weak form of the Ward–Takahashi identity. This observation, due to Weinberg, is extended to gravity in **Problem 10**.
- **Problem 3** establishes some formal relationships between various expressions for the time evolution operator and the S-matrix. Then, **Problem 4** shows that the expression for the S-matrix as the time-ordered exponential of a local interaction term is to a large extent a consequence of causality.
- In **Problem 5**, we derive a set of conditions, known as the *Landau equations*, for a given loop integral to have infrared or collinear singularities. An explicit multi-loop integration is studied in **Problem 6**, which provides another point of view on these conditions.
- **Problem 7** establishes *Weinberg's convergence theorem* in the simple case of scalar field theory, a crucial result in the discussion of renormalization since it clarifies the role of the superficial degree of divergence in assessing whether a particular diagram is ultraviolet divergent.
- The electron anomalous magnetic moment is calculated at one loop in **Problem 8**. This is a classic QED calculation of great historical importance, which has now been pushed to five loops and provides one of the most precise agreements between theory and experiment in all of physics.
- **Problem 11** derives the *Lee–Nauenberg theorem*, an important result about soft and collinear singularities which states that such divergences are removed by summing transition probabilities over degenerate states, thereby providing a link between the finiteness of a quantity and its practical measurability.
- In **Problem 12**, we discuss the external classical field approximation, thanks to which a heavy charged object may be replaced by its classical Coulomb field.
- **Problems 13** and **14** are devoted to a derivation of the *Low–Burnett–Kroll theorem*, a result that states that the emission probability of a soft photon is proportional to the probability of the underlying hard process, at the first two orders in the energy of the emitted photon.
- *Coherent states* are introduced in **Problem 15** and their main properties established. They will be discussed further in **Problems 20, 21** and **22**.
- **Problems 16** and **17** study the running coupling in a scalar field theory with two fields, and in a QCD-like theory at two-loop order.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

1. Relationship between the Equations of Motion of ϕ and $\phi_{\rm in}$ Recall that the field operators in the Heisenberg representation (ϕ) and in the interaction representation ($\phi_{\rm in}$) are related by

$$\phi(x) = U(-\infty, x^0)\,\phi_{\rm in}(x)\,U(x^0, -\infty). \tag{1.37}$$

The goal of the problem is to show that, if $V(\phi)$ is the interaction potential, this implies the following relationship between the left-hand sides of their respective equations of motion:

$$(\square + m^2)\phi(x) + V'(\phi(x)) = U(-\infty, x^0)\Big[(\square + m^2)\phi_{\rm in}(x)\Big]U(x^0, -\infty),$$

provided that $U(t_2, t_1) \equiv T\,\exp -i\int_{t_1}^{t_2} d^4x\,V(\phi_{\rm in}(x))$. This is an important consistency check, since it implies that $\phi_{\rm in}$ is a free field while ϕ evolves as prescribed by the self-interaction term in the Lagrangian.

1.a Apply a derivative ∂_μ to (1.37). Note that spatial derivatives do not act on the U's. In particular, show that $\partial_0\Big[U(-\infty, x^0)\phi_{\rm in}(x)U(x^0, -\infty)\Big] = U(-\infty, x^0)\Pi_{\rm in}(x)U(x^0, -\infty)$. How did the terms coming from the time derivative of the U's cancel?

1.b Apply a second time derivative to this result, to obtain

$$\begin{aligned}&\partial_0^2\Big[U(-\infty, x^0)\phi_{\rm in}(x)U(x^0, -\infty)\Big]\\ &\quad = U(-\infty, x^0)\Big[\partial_0^2\phi_{\rm in}(x) - i\int d^3\boldsymbol{y}\;\big[\Pi_{\rm in}(x), V(\phi_{\rm in}(x^0, \boldsymbol{y}))\big]\Big]U(x^0, -\infty).\end{aligned}$$

1.c Calculate the commutator on the right-hand side (one may prove that if $[A, B]$ is an object that commutes with all other operators, then $[A, f(B)] = f'(B)\,[A, B]$).

1.a Let us start from

$$\phi(x) = U(-\infty, x^0)\phi_{\rm in}(x)U(x^0, -\infty).$$

Since the evolution operators depend only on time, we have trivially

$$(-\boldsymbol{\nabla}^2 + m^2)\phi(x) = U(-\infty, x^0)\Big[(-\boldsymbol{\nabla}^2 + m^2)\phi_{\rm in}(x)\Big]U(x^0, -\infty),$$

and the main difficulty is to deal with the time derivatives. The first time derivative reads

$$\begin{aligned}\partial_0\phi(x) &= \Big[\partial_0 U(-\infty, x^0)\Big]\phi_{\rm in}(x)U(x^0, -\infty) + U(-\infty, x^0)\phi_{\rm in}(x)\Big[\partial_0 U(x^0, -\infty)\Big]\\ &\quad + U(-\infty, x^0)\Big[\partial_0\phi_{\rm in}(x)\Big]U(x^0, -\infty)\\ &= i\,U(-\infty, x^0)\big[\phi_{\rm in}(x), \mathcal{I}(x^0)\big]U(x^0, -\infty)\\ &\quad + U(-\infty, x^0)\Big[\partial_0\phi_{\rm in}(x)\Big]U(x^0, -\infty),\end{aligned}$$

where we denote $\mathcal{I}(x^0) \equiv -\int d^3\boldsymbol{x}\;V(\phi_{\rm in}(x^0, \boldsymbol{x}))$. The first line contains an equal-time commutator of $\phi_{\rm in}$ with some functional of $\phi_{\rm in}$, which is zero, leaving only the non-vanishing term of the second line.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

1.b A second differentiation with respect to time gives

$$\partial_0^2\phi(x) = i\,U(-\infty,x^0)\big[\partial_0\phi_{in}(x),\mathcal{I}(x^0)\big]U(x^0,-\infty) \\ +U(-\infty,x^0)\Big[\partial_0^2\phi_{in}(x)\Big]U(x^0,-\infty).$$

1.c The commutator in the first line is an equal-time commutator between the canonical momentum $\partial_0\phi_{in}$ and a functional of the field ϕ_{in}. In order to evaluate it, we need the following result:

$$[A,f(B)] = [A,B]\,f'(B)\,,\quad \text{valid when}\ \ [[A,B],B]=0.$$

This can be shown by using the Taylor series of $f(B)$, by first showing by recursion that $[A,B^n]=n[A,B]B^{n-1}$. Then, we can write

$$\begin{aligned} i\big[\partial_0\phi_{in}(x),\mathcal{I}(x^0)\big] &= -i\int d^3\boldsymbol{y}\ \big[\partial_0\phi_{in}(x),V(\phi_{in}(x^0,\boldsymbol{y}))\big] \\ &= -i\int d^3\boldsymbol{y}\ \underbrace{\big[\partial_0\phi_{in}(x),\phi_{in}(x^0,\boldsymbol{y})\big]}_{-i\delta(\boldsymbol{x}-\boldsymbol{y})}\,V'(\phi_{in}(x^0,\boldsymbol{y})) \\ &= -V'(\phi_{in}(x)). \end{aligned}$$

Now, using $U(-\infty,x^0)V'(\phi_{in}(x))U(x^0,-\infty)=V'(\phi(x))$, we get

$$(\square+m^2)\phi(x)-\mathcal{L}'(\phi(x)) = U(-\infty,x^0)\Big[(\square+m^2)\phi_{in}(x)\Big]U(x^0,-\infty).$$

Therefore, the left-hand sides of the equations of motion for the Heisenberg representation (interacting) field ϕ and for the interaction representation (free) field ϕ_{in} are related by a unitary transformation identical to the formula that relates the field operators themselves.

2. Little-Group Elements for Massless Particles The *little group* is the subgroup of the Lorentz group that leaves a fixed reference vector q^μ invariant. In this problem, we derive a particularly convenient explicit form of the elements of the little group in the case where q^μ is the light-like vector $q^\mu \equiv (\omega,0,0,\omega)$.

2.a First, show that an infinitesimal little-group transformation of this kind can be written as follows:

$$R \approx 1 - i\theta\,J^3 + i\alpha_1\,B^1 + i\alpha_2\,B^2,$$

with three generators J^3, B^1, B^2 (the first one being the generator of rotations about the third direction of space) that one should determine explicitly. Check also that they satisfy the following commutation relations (after an appropriate normalization):

$$\big[J^3,B^i\big] = i\epsilon_{ij}\,B^j,\qquad \big[B^1,B^2\big]=0.$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

2.b Show that any finite element R of the massless little group can be written as

$$R^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1+\frac{\beta_1^2+\beta_2^2}{2} & \beta_1 & \beta_2 & -\frac{\beta_1^2+\beta_2^2}{2} \\ \beta_1 & 1 & 0 & -\beta_1 \\ \beta_2 & 0 & 1 & -\beta_2 \\ \frac{\beta_1^2+\beta_2^2}{2} & \beta_1 & \beta_2 & 1-\frac{\beta_1^2+\beta_2^2}{2} \end{pmatrix}.$$

Hint: use the exact Baker–Campbell–Hausdorff formula:

$$\ln\left(e^{iX}e^{iY}\right) = i\,X + i\int_0^1 dt\; F\left(e^{t\,\mathrm{ad}_Y}\, e^{\mathrm{ad}_X}\right)\; Y,$$

$$\textit{where}\quad \mathrm{ad}_X(Y) \equiv -i\big[X,Y\big] \quad \textit{and} \quad F(z) \equiv \frac{\ln(z)}{z-1},$$

in order to exponentiate the infinitesimal form.

2.a This question is not difficult to solve by "brute force," i.e., by looking for the most general little-group transformation

$$R^\mu{}_\nu \equiv \delta^\mu{}_\nu + \omega^\mu{}_\nu$$

such that $\omega^\mu{}_\nu q^\nu = 0$ for all values of the index μ, with the additional constraint that $\omega_{\mu\nu} = 0$ is antisymmetric (so we have a legitimate infinitesimal Lorentz transformation). If we parameterize

$$\omega_{\mu\nu} \equiv \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix},$$

the condition that $q^\nu \equiv (\omega, 0, 0, \omega)$ is invariant is equivalent to

$$c = 0, \quad a = e, \quad b = f,$$

implying that there is a three-parameter family of $\omega^\mu{}_\nu$'s that fulfill all the requirements:

$$\begin{aligned} \omega^\mu{}_\nu &= \theta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \alpha_1 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= -i\theta\,(\underbrace{M^{12}}_{J^3})^\mu{}_\nu + i\,\alpha_1(\underbrace{M^{10} - M^{13}}_{\equiv B^1})^\mu{}_\nu + i\,\alpha_2(\underbrace{M^{20} - M^{23}}_{\equiv B^2})^\mu{}_\nu. \end{aligned}$$

(The identification in the second line follows from $\omega_{\mu\nu} = -\frac{i}{2}\omega_{\alpha\beta}\left(M^{\alpha\beta}\right)_{\mu\nu}$, valid for the spin-1 representation of the Lorentz algebra, where $(M^{\alpha\beta})_{\mu\nu} = i(\delta^\alpha{}_\mu\delta^\beta{}_\nu - \delta^\alpha{}_\nu\delta^\beta{}_\mu)$.) Note that the first term corresponds to rotations about the x^3 axis, which trivially leaves invariant any vector whose only non-zero spatial component is along the third direction.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

The announced commutation relations can be checked by an explicit evaluation of the corresponding matrix products. Alternatively, they can also be obtained from

$$J^3 = M^{12}, \quad B^1 = M^{10} - M^{13}, \quad B^2 = M^{20} - M^{23},$$

and by using the defining commutation relation of the Lorentz algebra,

$$\left[M^{\mu\nu}, M^{\rho\sigma}\right] = i(g^{\nu\rho}M^{\mu\sigma} - g^{\mu\rho}M^{\nu\sigma}) - i(g^{\nu\sigma}M^{\mu\rho} - g^{\mu\sigma}M^{\nu\rho}).$$

2.b Given the infinitesimal form of little-group transformations for massless particles derived above, any finite little-group transformation R can be obtained by exponentiating the inifinitesimal ones:

$$R \equiv e^{i(-\theta J^3 + \alpha_1 B^1 + \alpha_2 B^2)}.$$

Note first that the factor on the left of the proposed formula is nothing but the rotation $e^{-i\theta J^3}$, i.e., a rotation by an angle θ about the x^3 axis, which affects only the coordinates $1, 2$:

$$\left(e^{-i\theta J^3}\right)^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus, the form proposed in the statement of the problem suggests that R may also be written as

$$R = e^{-i\theta J^3}\, e^{i(\beta_1 B^1 + \beta_2 B^2)}.$$

Our ansatz for the form of the second factor is based on the case $\theta = 0$ (in this case, we should of course have $\beta_i = \alpha_i$, but this is not necessarily true for $\theta \neq 0$). Verifying that the two expressions for R are equivalent could in principle be performed by calculating both expressions and equating them in order to find the relationship between the coefficients $\alpha_{1,2}$ and $\beta_{1,2}$, but this is rather challenging.

A much more efficient method is to use the exact form of the Baker–Campbell–Hausdorff formula, recalled in the statement of the problem. In the present case, the commutation relations among J^3, B^1, B^2 lead to

$$\mathrm{ad}_{-\theta J^3}(\beta_i B^i) = -\theta\, \beta_i \epsilon_{ij} B^j, \quad e^{\mathrm{ad}_{-\theta J^3}}\, \beta_i B^i = \beta_i \left(e^{-\theta\epsilon}\right)_{ij} B^j, \quad \mathrm{ad}_{\beta_i B^i}(B^j) = 0,$$

and the Baker–Campbell–Hausdorff formula gives

$$\ln\left(e^{-i\theta J^3} e^{i\beta_i B^i}\right) = -i\theta J^3 - i\beta_i \left[\frac{\theta\epsilon}{e^{-\theta\epsilon} - 1}\right]_{ij} B^j.$$

Using the fact that $\epsilon_{ij}\epsilon_{jk} = -\delta_{ik}$, we have

$$e^{-\theta\epsilon} - 1 = \cos\theta - 1 - \epsilon\sin\theta,$$

$$\ln\left(e^{-i\theta J^3} e^{i\beta_i B^i}\right) = -i\theta J^3 - i\,\underbrace{\theta\beta_i\left[\epsilon(\cos\theta - 1 - \epsilon\sin\theta)^{-1}\right]_{ij}}_{-\alpha_j} B^j.$$

This formula shows that $e^{i(-\theta J^3 + \alpha_i B^i)} = e^{-i\theta J^3} e^{i\beta_i B^i}$ with the following relationship between the coefficients $\alpha_{1,2}$ and $\beta_{1,2}$:

$$\beta_i = \theta^{-1}\left(\sin\theta + \epsilon(1 - \cos\theta)\right)_{ij} \alpha_j.$$

(Note that $\beta_i = \alpha_i$ when $\theta \to 0$, as expected trivially in this limit.)

In order to calculate the factor $e^{i\beta_i B^i}$, one should first note that $e^{i\beta_i B^i} = e^{i\beta_1 B^1} e^{i\beta_2 B^2}$ since B^1 and B^2 commute. Using the explicit representations of $B^{1,2}$, simple algebra shows that

$$\big((iB^1)^2\big)^\mu{}_\nu = \big((iB^2)^2\big)^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix},$$

and

$$\big((iB^1)^n\big)^\mu{}_\nu = \big((iB^2)^n\big)^\mu{}_\nu = 0 \quad \text{for } n \geq 3.$$

Therefore, we have

$$\big(e^{i\beta_1 B^1}\big)^\mu{}_\nu = \Big(1 + i\beta_1 B^1 + \tfrac{1}{2}\beta_1^2 (iB^1)^2\Big)^\mu{}_\nu = \begin{pmatrix} 1+\frac{\beta_1^2}{2} & \beta_1 & 0 & -\frac{\beta_1^2}{2} \\ \beta_1 & 1 & 0 & -\beta_1 \\ 0 & 0 & 1 & 0 \\ \frac{\beta_1^2}{2} & \beta_1 & 0 & 1-\frac{\beta_1^2}{2} \end{pmatrix},$$

$$\big(e^{i\beta_2 B^2}\big)^\mu{}_\nu = \Big(1 + i\beta_2 B^2 + \tfrac{1}{2}\beta_2^2 (iB^2)^2\Big)^\mu{}_\nu = \begin{pmatrix} 1+\frac{\beta_2^2}{2} & 0 & \beta_2 & -\frac{\beta_2^2}{2} \\ 0 & 1 & 0 & 0 \\ \beta_2 & 0 & 1 & -\beta_2 \\ \frac{\beta_2^2}{2} & 0 & \beta_2 & 1-\frac{\beta_2^2}{2} \end{pmatrix},$$

and finally

$$\big(e^{i\beta_i B^i}\big)^\mu{}_\nu = \begin{pmatrix} 1+\frac{\beta_1^2+\beta_2^2}{2} & \beta_1 & \beta_2 & -\frac{\beta_1^2+\beta_2^2}{2} \\ \beta_1 & 1 & 0 & -\beta_1 \\ \beta_2 & 0 & 1 & -\beta_2 \\ \frac{\beta_1^2+\beta_2^2}{2} & \beta_1 & \beta_2 & 1-\frac{\beta_1^2+\beta_2^2}{2} \end{pmatrix},$$

which establishes the announced result.

3. S-matrix in Terms of ϕ_{in} and ϕ_{out} Given an interaction Lagrangian $\mathcal{L}_I$, the field operators in the Heisenberg representation (ϕ) and in the interaction representation (ϕ_{in}) are related by means of a time evolution operator

$$U_{in}(t_2, t_1) \equiv T\,\exp\left(i\int_{t_1}^{t_2} d^4x\,\mathcal{L}_I\big(\phi_{in}(x)\big)\right)$$

expressed in terms of the free field ϕ_{in} (we have added a subscript *in* to this evolution operator, in order to recall that it is defined in terms of ϕ_{in}, since we are about to introduce its counterpart defined in terms of ϕ_{out}). Likewise, we define a similar evolution operator in terms of ϕ_{out}, the field operator in the interaction picture that coincides with the Heisenberg picture at $x^0 = +\infty$:

$$U_{out}(t_2, t_1) \equiv T\,\exp\left(i\int_{t_1}^{t_2} d^4x\,\mathcal{L}_I\big(\phi_{out}(x)\big)\right).$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

3.a Show that $U_{in}(+\infty, -\infty) = U_{out}(+\infty, -\infty)$. In other words, the S-matrix (i.e., the time evolution operator over the entire time range) does not depend on whether it is defined in terms ϕ_{in} or ϕ_{out}.

3.b Are U_{in} and U_{out} identical in general? Find the relationship between the two.

3.c Show that the S-matrix is also given by $S = U_{in}(x^0, -\infty)U_{out}(+\infty, x^0)$, for any intermediate time x^0. Note that, on the surface, this expression does not seem to be properly time-ordered. Why is it nevertheless a correct formula?

3.a Recall the relationship between the interacting field ϕ and the free field ϕ_{in} of the interaction representation:

$$\phi(x) = U_{in}(-\infty, x^0)\, \phi_{in}(x)\, U_{in}(x^0, -\infty),$$

$$U_{in}(t_2, t_1) \equiv T \exp\left(i \int_{t_1}^{t_2} d^4x\, \mathcal{L}_I(\phi_{in}(x))\right). \tag{1.38}$$

(Since we shall shortly write the analogous relationship with ϕ_{out}, it is important to have a subscript *in* on the evolution operator to avoid confusion, since it depends on ϕ_{in}.) By taking the limit $x^0 \to +\infty$ in this equation, and using the fact that in this limit the interacting field becomes identical to the free field ϕ_{out}, we obtain a first relationship between ϕ_{in} and ϕ_{out}:

$$\phi_{out}(x) \underset{x^0 \to +\infty}{=} U_{in}(-\infty, +\infty)\, \phi_{in}(x)\, U_{in}(+\infty, -\infty).$$

Strictly speaking, this limiting procedure gives a relationship between the two fields only for large x^0. Then, we use the fact that two fields obeying the same equation of motion (here, ϕ_{in} and ϕ_{out} both obey the Klein–Gordon equation) and identical in some region of time are equal at all times (this argument relies on the uniqueness of the solutions of the Klein–Gordon equation, if their value and that of their first time derivative are prescribed at some time). Therefore, the above equation is in fact valid at all times.

Note that the right-hand side of this equation depends only on ϕ_{in}, but in a completely non-linear and non-local fashion because of the evolution operators. Another noteworthy aspect of this equation is that, despite the fact that both ϕ_{in} and ϕ_{out} are free fields, the relationship between the two involves the interactions.

The easiest way to invert the relationship between ϕ_{in} and ϕ_{out} is to write the analogue of (1.38) for the free field ϕ_{out}:

$$\phi(x) = U_{out}(+\infty, x^0)\, \phi_{out}(x)\, U_{out}(x^0, +\infty),$$

$$U_{out}(t_2, t_1) \equiv T \exp\left(i \int_{t_1}^{t_2} d^4x\, \mathcal{L}_I(\phi_{out}(x))\right). \tag{1.39}$$

Taking the limit $x^0 \to -\infty$ in this equation leads to a second form of the formula that relates

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

ϕ_{in} and ϕ_{out}:

$$\phi_{in}(x) = U_{out}(+\infty, -\infty)\, \phi_{out}(x)\, U_{out}(-\infty, +\infty),$$

or, equivalently,

$$\phi_{out}(x) = U_{out}(-\infty, +\infty)\, \phi_{in}(x)\, U_{out}(+\infty, -\infty).$$

In order for the two relations we have obtained to be consistent, we must have

$$U_{in}(+\infty, -\infty) = U_{out}(+\infty, -\infty).$$

Therefore, the evolution operators *over the entire time range* are identical, regardless of whether they are constructed with the fields ϕ_{in} or ϕ_{out}. For this combination of time arguments, we may drop the subscripts *in/out* on the evolution operators.

3.b But it is important to realize that this property is not true for arbitrary time intervals. By requesting that (1.38) and (1.39) give the same interacting field ϕ, we must have

$$\begin{aligned} &U_{in}(-\infty, x^0)\, \phi_{in}(x)\, U_{in}(x^0, -\infty) \\ &\quad = U_{out}(+\infty, x^0)\, \phi_{out}(x)\, U_{out}(x^0, +\infty) \\ &\quad = U_{out}(+\infty, x^0) U(-\infty, +\infty)\, \phi_{in}(x)\, U(+\infty, -\infty) U_{out}(x^0, +\infty), \end{aligned}$$

implying that in general we have

$$U_{in}(x^0, -\infty) = U(+\infty, -\infty) U_{out}(x^0, +\infty). \tag{1.40}$$

Writing $U_{in}(x^0, -\infty) = U_{in}(x^0, y^0) U_{in}(y^0, -\infty)$ and using the same identity with x^0 replaced by y^0, we obtain

$$U_{out}(x^0, y^0) = U(-\infty, +\infty) U_{in}(x^0, y^0) U(+\infty, -\infty).$$

(This relation could also have been obtained from the definition of U_{out}, by performing its Taylor expansion in powers of ϕ_{out}, replacing every occurrence of ϕ_{out} by its expression in terms of ϕ_{in}, and at the end repackaging the series to obtain a U_{in}.)

3.c By multiplying (1.40) on the right by the inverse of $U_{out}(x^0, +\infty)$, we obtain another formula for the full evolution operator,

$$U(+\infty, -\infty) = U_{in}(x^0, -\infty) U_{out}(+\infty, x^0),$$

which is rather counterintuitive since the order of the operators on the right-hand side may (wrongly) suggest that it is inconsistent with the time ordering. The resolution of this paradox is that $U(+\infty, -\infty)$ is time-ordered when expressed entirely in terms of ϕ_{in} or entirely in terms of ϕ_{out}; but the right-hand side of the above formula mixes ϕ_{in} and ϕ_{out}, and the relationship between ϕ_{in} and ϕ_{out} is non-local in time, which obscures the actual time ordering of the operators.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

4. Constraints on the S-matrix from Causality The goal of this problem is to derive general constraints on the S-matrix from causality. To that end, let us assume that the coupling constant λ that controls the interactions is a function of spacetime, $\lambda(x)$. With this modification, the S-matrix becomes a functional $S[\lambda]$.

4.a Consider two regions of spacetime, Ω_1 and Ω_2, such that Ω_2 lies *in the future light-cone* of Ω_1, and denote by $\lambda_{1,2}(x)$ the coupling function restricted to these domains (we assume it is zero outside of $\Omega_1 \cup \Omega_2$). Show that $S[\lambda_1 + \lambda_2] = S[\lambda_2]S[\lambda_1]$.

4.b Generalize this result to the more general situation where Ω_1 and Ω_2 are simply separated by a locally space-like surface.

4.c Consider now the case where the separation between any pair of points of Ω_1 and Ω_2 is space-like. Show that $\big[S[\lambda_1], S[\lambda_2]\big] = 0$.

4.d If two coupling functions λ and λ' coincide for $x^0 \leq y^0$, show that $S[\lambda']S^\dagger[\lambda]$ does not depend on the behavior of the coupling at times $\leq y^0$. By considering an infinitesimal variation of the coupling function, show that

$$\frac{\delta}{\delta\lambda(y)}\left(\frac{\delta S[\lambda]}{\delta\lambda(x)}\, S^\dagger[\lambda]\right) = 0 \quad \text{if } y \text{ is not in the future light-cone of } x.$$

4.e Solve the constraints of causality and unitarity to obtain the form of $S[\lambda]$ up to $\mathcal{O}(\lambda^2)$.

4.a Let us consider two regions Ω_1 and Ω_2 of spacetime, as shown in Figure 1.1. Thus, for any pair of points $x_1 \in \Omega_1, x_2 \in \Omega_2$, we have $x_2^0 > x_1^0$ and $(x_1 - x_2)^2 > 0$. We assume that interactions exist only in $\Omega_1 \cup \Omega_2$ and are zero elsewhere, and we denote by $\lambda_{1,2}$ the coupling functions in these two domains. The coupling function over the entire spacetime is thus

$$\lambda(x) = \lambda_1(x) + \lambda_2(x),$$

and the full S-matrix is $S[\lambda_1 + \lambda_2]$. Recall that the S-matrix is the operator that connects the *in* states at $x^0 = -\infty$ and the *out* states at $x^0 = +\infty$:

$$\big\langle \alpha_{\text{out}}\big| = \big\langle \alpha_{\text{in}}\big|\, S_{\text{in}}[\lambda_1 + \lambda_2].$$

For the time being, it is safer to add a subscript *in* on the S-matrix in order to indicate that it is expressed in terms of the field operator ϕ_{in}. Given the relative configuration of the domains Ω_1 and Ω_2, we could also construct another version of the interaction representation, where the fields coincide with the Heisenberg representation ones at some intermediate time located between Ω_1 and Ω_2. Let us call *intermediate* this representation, and ϕ_{inter} the corresponding free field operator. We have

$$\big\langle \alpha_{\text{inter}}\big| = \big\langle \alpha_{\text{in}}\big|\, S_{\text{in}}[\lambda_1], \quad \big\langle \alpha_{\text{out}}\big| = \big\langle \alpha_{\text{inter}}\big|\, S_{\text{inter}}[\lambda_2].$$

Thus, we obtain

$$S_{\text{in}}[\lambda_1 + \lambda_2] = S_{\text{in}}[\lambda_1]\, S_{\text{inter}}[\lambda_2].$$

Note that this equation is subject to the same paradox regarding the ordering of the operators as in the last equation in Problem 3: the seemingly unnatural ordering between $S_{\text{in}}[\lambda_1]$ and

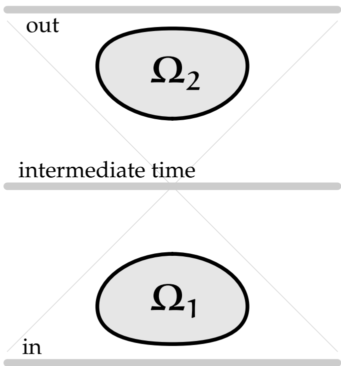


Figure 1.1 Domains Ω_1 and Ω_2, with Ω_2 under the causal influence of Ω_1.

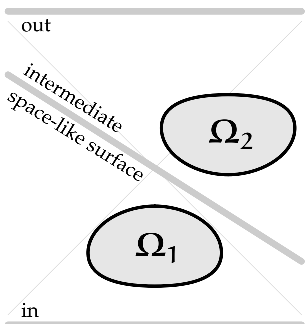


Figure 1.2 Domains $\Omega_{1,2}$, such that Ω_2 does not causally influence Ω_1.

$S_{inter}[\lambda_2]$ is due to the fact that the latter is implicitly expressed in terms of the field ϕ_{inter} while the former depends on the field ϕ_{in}. Using exactly the same manipulations as in Problem 3, we can rewrite this expression solely in terms of ϕ_{in}, which gives

$$S_{inter}[\lambda_2] = S_{out}[\lambda_2],$$
$$S_{in}[\lambda_1 + \lambda_2] = S_{in}[\lambda_1]S_{out}[\lambda_2],$$
$$S_{in}[\lambda_1 + \lambda_2] = S_{in}[\lambda_2]\, S_{in}[\lambda_1].$$

From now on, we implicitly assume that all S-matrix operators are expressed in terms of ϕ_{in}, and we suppress the subscript *in*.

4.b In words, the previous setup could be described by saying that Ω_2 *is under the influence of* Ω_1. A much more general situation would be to simply request that Ω_2 *does not influence* Ω_1. This is achieved by dividing spacetime with a hyper-surface Σ located between the domains Ω_1 and Ω_2, provided that this surface is locally space-like (i.e., no signal can travel from the surface towards the domain below it). This setup is shown in Figure 1.2. This ordering of the domains Ω_1 and Ω_2 is sufficient to reproduce the preceding arguments, leading again to $S[\lambda_1 + \lambda_2] = S[\lambda_2]\, S[\lambda_1]$.

4.c Consider now the situation where the interval between any point in Ω_1 and any point in Ω_2 is space-like, as illustrated in Figure 1.3 (left panel). Since it is possible to find an

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

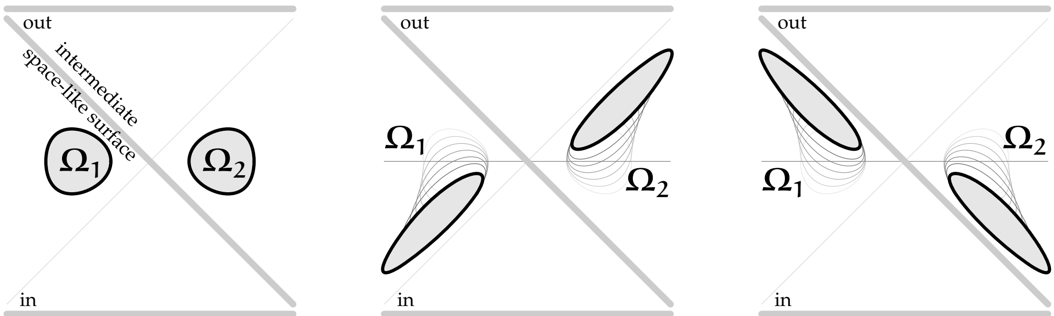


Figure 1.3 Left: domains Ω_1 and Ω_2 with mutual space-like separations. Center and right: displacements of the domains under a Lorentz boost.

appropriate space-like surface separating the two domains, the previous result applies. But now there is an ambiguity regarding which domain should be considered as being "before" the other one. In fact, what is special about this kinematical configuration is that the time ordering between the two domains can be altered by applying a Lorentz boost, as illustrated by the second ($\beta < 0$) and third panels ($\beta > 0$) of Figure 1.3. In this case, because of this lack of absolute time ordering between the two domains, the following two equations are both true:

$$S[\lambda_1 + \lambda_2] = S[\lambda_1]\, S[\lambda_2] \quad \text{and} \quad S[\lambda_1 + \lambda_2] = S[\lambda_2]\, S[\lambda_1],$$

which implies that $\big[S[\lambda_1], S[\lambda_2]\big] = 0$.

4.d Consider now two distinct coupling functions in the domain Ω_2, namely λ_2 and λ'_2. Thus, we have two realizations of the coupling function:

$$\lambda \equiv \lambda_1 + \lambda_2, \quad \lambda' \equiv \lambda_1 + \lambda'_2.$$

For these two coupling functions, the S-matrix is given by

$$S[\lambda] = S[\lambda_2]\, S[\lambda_1], \quad S[\lambda'] = S[\lambda'_2]\, S[\lambda_1],$$

and we therefore have

$$S[\lambda']\, S^\dagger[\lambda] = S[\lambda'_2]\, S[\lambda_1]\, S^\dagger[\lambda_1]\, S^\dagger[\lambda_2] = S[\lambda'_2]\, S^\dagger[\lambda_2].$$

This combination is independent of the function λ_1, i.e., independent of the behavior of the coupling function in the portion of spacetime that cannot receive any causal influence from Ω_2. Let us now assume that the difference $\delta\lambda \equiv \lambda'_2 - \lambda_2$ is infinitesimal. To first order in $\delta\lambda$, we

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

have

$$S[\lambda'] = S[\lambda] + \int d^4x\, \delta\lambda(x)\, \frac{\delta S[\lambda]}{\delta\lambda(x)} + \mathcal{O}(\delta\lambda^2),$$

and

$$S[\lambda']\, S^\dagger[\lambda] = 1 + \int d^4x\, \delta\lambda(x)\, \frac{\delta S[\lambda]}{\delta\lambda(x)}\, S^\dagger[\lambda] + \mathcal{O}(\delta\lambda^2).$$

Since this should be independent of λ_1 for any variation $\delta\lambda$, we must have

$$\frac{\delta S[\lambda]}{\delta\lambda(x)}\, S^\dagger[\lambda] \quad \text{is independent of } \lambda_1 \text{ if } \Omega_1 \text{ is not under the causal influence of } x.$$

This condition can also be phrased as

$$\frac{\delta}{\delta\lambda(y)}\left(\frac{\delta S[\lambda]}{\delta\lambda(x)}\, S^\dagger[\lambda]\right) = 0 \quad \text{if } y \text{ is not in the future light-cone of } x. \tag{1.41}$$

Note that the same identity would be true in a theory where the fields are coupled to some external source J_{ext}, if we replace λ by J_{ext}, again thanks to causality.

4.e In order to see the consequences of this constraint on the S-matrix, let us write a formal Taylor expansion of the functional $S[\lambda]$:

$$S[\lambda] = 1 + \int d^4x\, S_1(x)\, \lambda(x) + \frac{1}{2}\int d^4x d^4y\, S_2(x,y)\, \lambda(x)\lambda(y) + \cdots.$$

In this expansion, the objects $S(x_1, \ldots, x_n)$ are operator-valued symmetric functions of their arguments. In addition to the constraint (1.41), the S-matrix must be unitary, and also satisfy $\big[S[\lambda_1], S[\lambda_2]\big] = 0$ when the supports of λ_1 and λ_2 have purely space-like separations. The last constraint implies that the coefficients S_n are *multi-local* operators (i.e., $S_n(x_1, \ldots, x_n)$ depends only on the field operator and its derivatives at the points $x_1, \ldots, x_n$) constructed with the field operator and its derivatives (non-locality would lead to violations of this commutation relation). In the first two orders, the unitarity of S implies that

$$S_1(x) + S_1^\dagger(x) = 0, \quad S_2(x,y) + S_2^\dagger(x,y) + S_1^\dagger(x)S_1(y) + S_1^\dagger(y)S_1(x) = 0.$$

(In deriving the second equation, we must be careful to symmetrize the coefficient that multiplies $\lambda(x)\lambda(y)$ in the integrand of the second-order term.) These equations can be rewritten as

$$S_1^\dagger(x) = -S_1(x), \quad S_2(x,y) + S_2^\dagger(x,y) = S_1(x)S_1(y) + S_1(y)S_1(x).$$

Note that unitarity can only constrain the Hermitian part of the coefficients S_n, and does not say anything about their anti-Hermitian part. To put constrains on the latter, we need to make

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

use of (1.41). It is straightforward to check that

$$\frac{\delta}{\delta\lambda(y)}\left(\frac{\delta S[\lambda]}{\delta\lambda(x)}\,S^\dagger[\lambda]\right) = S_2(x,y) + S_1(x)S_1^\dagger(y) + \mathcal{O}(\lambda).$$

Therefore, at lowest order, (1.41) tells us that

$$S_2(x,y) = -S_1(x)S_1^\dagger(y) = S_1(x)S_1(y) \quad \text{if } y \text{ is not in the future light-cone of } x.$$

Using the fact that $S_2(x,y)$ should be symmetric, we also have

$$S_2(x,y) = S_2(y,x) = S_1(y)S_1(x) \quad \text{if } x \text{ is not in the future light-cone of } y.$$

(Note that these two conditions on x, y are both satisfied if their separation is space-like. For such a space-like separation, we could thus use either of the two formulas. This does not lead to any contradiction, provided that $S_1(x)$ is a local function of the field operator and its derivatives.) Therefore, the answer valid for any x, y can be written as

$$S_2(x,y) = \mathrm{T}\,(S_1(x), S_1(y)).$$

(One may check a posteriori that the unitarity constraint is satisfied.) Using the same constraints, one could show by induction that the coefficient of order n, $S_n(x_1,\ldots,x_n)$, is the time-ordered product of n factors S_1. Therefore, we see that unitarity and causality provide an almost closed form for the S-matrix,

$$S[\lambda] = \mathrm{T}\,\exp\left(\int d^4x\; S_1(x)\lambda(x)\right),$$

in which the only remaining unknown is the first coefficient $S_1(x)$. The latter can be related to the interaction Lagrangian by considering a $\lambda(x)$ which is non-zero in an infinitesimal region of spacetime. Therefore, regardless of the microscopic details of a given theory – which control the first coefficient S_1 –, the general structure of the S-matrix is governed to a large extent by the constraints provided by unitarity and causality.

5. Landau Equations for Soft and Collinear Singularities The goal of this problem is to study the singularities that may occur in a Feynman integral due to vanishing denominators (not to be confused with ultraviolet divergences, due to an integrand that decreases too slowly at large momentum). Consider a Feynman integral with L loops and m denominators:

$$\mathcal{I}(\{\mathbf{p}_i\}) \equiv \int \prod_{j=1}^{L} \frac{d^D\ell_j}{(2\pi)^D}\; \frac{\mathcal{N}(\{p_i\},\{\ell_j\})}{(q_1^2 - m_1^2 + i0^+)\cdots(q_m^2 - m_m^2 + i0^+)}.$$

In this integral, the p_i are the external momenta, the ℓ_j the loop momenta, and the q_r the momenta of the propagators in the loops, i.e., linear combinations of loop momenta and external momenta of the form $q_r \equiv \sum_{j=1}^{L} \epsilon_{rj}\ell_j + \Delta_r$ (where the coefficients ϵ_{rj} take values in $\{-1, 0, +1\}$ and the Δ_r depend only on the external momenta).

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

5.a Use Feynman parameters x_r to combine the m denominators into a single one, $\mathcal{D}^m$.

5.b By considering the following elementary examples,

$$\int_{-1}^{+1} \frac{dx}{x+i0^+}, \quad \int_{0}^{+1} \frac{dx}{x+i0^+}, \quad \int_{-1}^{+1} \frac{dx}{(x+i0^+)(x-i0^+)},$$

explain why singularities occur only when a pole is located on the boundary of the integration manifold, or when multiple poles "pinch" the integration manifold.

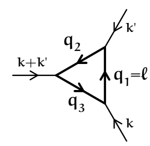
5.c Study the variations of $\mathcal{D}$ as a function of the loop momenta ℓ_j, and explain why the zeroes of $\mathcal{D}$ must also be extrema of $\mathcal{D}$ in order to produce singularities.

5.d Now study the dependence with respect to the Feynman parameters and show that, for each r, one must have either $x_r = 0$ or $q_r^2 = m_r^2$ in order to have a singularity.

5.e Conclude that the conditions for a singularity are given by the *Landau equations*:

$$\text{For each propagator } r\text{: } x_r(q_r^2 - m_r^2) = 0; \quad \text{For each loop } j\text{: } \sum_{r=1}^{m} \epsilon_{rj} x_r q_r^\mu = 0.$$

5.f Determine the singularities of the one-loop Feynman diagram shown on the right (assume $k^2 = k'^2 = 0$, $(k+k')^2 > 0$ and all internal particles are massless)



5.a Consider a Feynman integral with L loops and m denominators. Generically, such an integral may be written as

$$\mathcal{I}(\{p_i\}) \equiv \int \prod_{j=1}^{L} \frac{d^D \ell_j}{(2\pi)^D} \frac{\mathcal{N}(\{p_i\},\{\ell_j\})}{(q_1^2 - m_1^2 + i0^+) \cdots (q_m^2 - m_m^2 + i0^+)},$$

where the p_i are the momenta external to the loops and the ℓ_j are L independent loop momenta. The momenta q_r are the momenta carried by the various propagators along the loops. They are all of the form

$$q_r \equiv \sum_{j=1}^{L} \epsilon_{rj} \ell_j + \Delta_r,$$

where the coefficients ϵ_{rj} take values in $\{-1, 0, +1\}$ (a propagator may belong to a loop or not, and may be oriented in the same way as the loop momentum or in the opposite direction) and where the Δ_r depend only on the external momenta (they are thus constants from the point of view of evaluating the loop integrals). $\mathcal{N}(\{p_i\},\{\ell_j\})$ is a numerator that comprises all the momentum dependence that may arise, e.g., from three-gluon vertices in QCD or from the Dirac traces if there are fermion loops. This factor plays no role in analyzing the singularities of

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

the integral, except in those rare situations where a singularity due to a vanishing denominator may be canceled by an accidental concomitant vanishing of the numerator.

The first step is to combine the m denominators into a single one thanks to Feynman parameterization:

$$\mathcal{I}(\{p_i\}) = \Gamma(m)\int\prod_{j=1}^{L}\frac{d^D\ell_j}{(2\pi)^D}\int_{x_r\geq 0}\prod_{r=1}^{m}dx_r\,\delta\big(1-\sum_r x_r\big)\,\frac{\mathcal{N}(\{p_i\},\{\ell_j\})}{\big(\mathcal{D}(\{q_r\},\{x_r\})+i0^+\big)^m}, \quad (1.42)$$

$$\mathcal{D}(\{q_r\},\{x_r\}) \equiv \sum_{r=1}^{m} x_r(q_r^2 - m_r^2).$$

5.b In this problem, we assume that the ultraviolet divergences have been properly disposed of by some regularization, and we are chiefly interested in the possibility of additional singularities that may arise from a vanishing denominator. Obviously, that the equation

$$\mathcal{D}(\{q_r\},\{x_r\}) = 0$$

has solutions in the integration domain is a necessary condition for having such singularities. But a zero of the denominator of the integrand does not always lead to a singularity in the integral. In order to see this, let us consider the following toy examples:

$$\int_{-1}^{+1}\frac{dx}{x+i0^+} = \int_{-1}^{+1}dx\,\left(P\left(\frac{1}{x}\right) - i\pi\,\delta(x)\right) = -i\pi,$$

$$\int_{0}^{+1}\frac{dx}{x+i0^+} = \infty,$$

$$\int_{-1}^{+1}\frac{dx}{(x+i0^+)(x-i0^+)} = \lim_{\epsilon\to 0^+}\int_{-1}^{+1}\frac{dx}{x^2+\epsilon^2} = \infty.$$

In the first example, the integral is finite despite the fact that denominator vanishes at $x = 0$, because the integration contour is slightly shifted from the singularity by the presence of the $i0^+$. An infinite result would be obtained when it is impossible to shift the contour to avoid the pole. In the second example, we cannot avoid the singularity because the pole occurs on the boundary (here, at an *endpoint*) of the integration domain. In the third example, there are two poles in the interior of the integration range, but these poles *pinch* the integration contour, which prevents moving the contour to avoid the poles and also leads to an infinite result.

5.c Although the conditions of occurrence of a genuine singularity are the same, the situation we have to analyze is arguably more complex than these toy examples because of the multivariate nature of the denominator. For a multi-dimension integral, the above condition is that the poles of the integrand cannot be avoided by deforming the integration manifold. We can make the following observations:

- Note first that the fact that $\mathcal{D}$ is raised to the m-th power in (1.42) is irrelevant for this discussion: every zero of $\mathcal{D}$ leads to a pole of order m of the integrand, but this is equivalent to having m poles all on the same side of the integration domain, so this cannot produce a pinch.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

- For the L loop momenta ℓ_j, the integration domain is $\mathbb{R}^{DL}$. We may add to this domain a point at infinity (assuming an ultraviolet regularization, the integrand goes to zero in all directions at infinity), which leads to an integration domain topologically equivalent to a DL-dimensional sphere S_{DL}. This domain is boundaryless and therefore the only possibility of singularities when integrating over the loop momenta is to have a pinch.

- For the Feynman parameters x_r, the integration domain is a $(m-1)$-dimensional simplex,

$$\left\{(x_1,\ldots,x_m)\middle| x_r \geq 0, \sum_r x_r = 1\right\},$$

i.e., a line segment for $m = 2$, a triangle for $m = 3$, a tetrahedron for $m = 4$, etc. Clearly, this domain has a boundary and therefore endpoint singularities may occur.

Let us now study the behavior of the denominator $\mathcal{D}$. This function is quadratic in the loop momenta, and linear in the Feynman parameters. Note also that, after a Wick rotation of the loop momenta, $\mathcal{D}$ is a negative definite quadratic form in the Euclidean loop momenta, at fixed $\{x_r\}$ (this is obvious from the fact that the second-degree part of $\mathcal{D}$ is a sum of squares weighted by the negative coefficients $-x_r$). Therefore, when varying the loop momenta at fixed $\{x_r\}$, the denominator $\mathcal{D}$ has a maximum. We can distinguish the following cases:

- If $\max_{\{\ell_j\}}(\mathcal{D}) < 0$, the denominator is always non-zero and there is no singularity.

- If $\max_{\{\ell_j\}}(\mathcal{D}) > 0$, the denominator can vanish, but the zeroes are simple zeroes that cannot pinch the integration manifold for the variables $\{\ell_j\}$. The integral is also finite in this case.

- The dangerous situation is when $\max_{\{\ell_j\}}(\mathcal{D}) = 0$, because the location of the maximum is then a double zero (as in the third of the toy examples considered earlier) in all the ℓ_j^μ variables.

Therefore, we are seeking zeroes of $\mathcal{D}$ that are also extrema of its dependence with respect to all the loop momenta:

$$\mathcal{D} = 0, \qquad \frac{\partial \mathcal{D}}{\partial \ell_j^\mu} = 0.$$

Note that these conditions are still not sufficient for a genuine singularity since we have only discussed what happens at fixed $\{x_r\}$, and we have not yet analyzed whether it may be avoided by a deformation of the integration domain for the x_r's. This discussion can be divided into two cases:

- First, note that if $q_r^2 \neq m_r^2$, a small variation of x_r will change the value of $\mathcal{D}$ and move the denominator away from zero. For x_r in the interior of its allowed range, this means that there is no actual singularity. The only exception is at $x_r = 0$, since this is on the boundary of the integration range.

- In contrast, when $q_r^2 = m_r^2$, the denominator $\mathcal{D}$ is independent of x_r, and a zero of $\mathcal{D}$ persists at all values of x_r.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

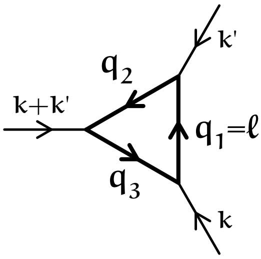


Figure 1.4 One-loop example of application of the Landau equations.

5.d We now have all the information to give the conditions of occurrence of a singularity in the Feynman integral $\mathcal{I}(\{p_i\})$:

$$\mathcal{D} = 0, \quad \frac{\partial \mathcal{D}}{\partial \ell_j^\mu} = 0; \quad \text{for each } r, \text{ either } x_r = 0 \text{ or } q_r^2 = m_r^2.$$

(Note that the last condition can be phrased as $x_r(q_r^2 - m_r^2) = 0$ for each r, which makes the first one, $\mathcal{D} = 0$, redundant.) These conditions are known as the *Landau equations*. By explicitly evaluating the derivative of the denominator with respect to the loop momenta, these conditions can be written as

$$\text{for each propagator } r\text{: } x_r(q_r^2 - m_r^2) = 0; \qquad \text{for each loop } j\text{: } \sum_{r=1}^{m} \epsilon_{rj} x_r q_r^\mu = 0.$$

5.e Consider now the example of Figure 1.4. The corresponding Landau equations read

$$x_1 \ell^2 = x_2(\ell + k')^2 = x_3(\ell - k)^2 = 0, \quad x_1 \ell^\mu + x_2(\ell^\mu + k'^\mu) + x_3(\ell^\mu - k^\mu) = 0.$$

A first solution is obtained for

$$\ell^2 = (\ell - k)^2 = 0, \quad x_2 = 0, \quad x_1 \ell^\mu + x_3(\ell^\mu - k^\mu) = 0,$$

which (since we have assumed that $k^2 = 0$) is equivalent to

$$\ell^2 = k^2 = \ell \cdot k = 0, \quad \ell^\mu = \frac{x_3}{x_1 + x_3} k^\mu.$$

This type of singularity is called a *collinear singularity*, since it occurs when the loop momentum is aligned with one of the external momenta. A similar type of singularity is obtained at $x_3 = 0$, i.e.,

$$\ell^2 = k'^2 = \ell \cdot k' = 0, \quad \ell^\mu = -\frac{x_2}{x_1 + x_2} k'^\mu.$$

Another way to fulfill the Landau equations is to have $\ell^\mu \ll k^\mu, k'^\mu$. In this limit, the first three conditions are automatically satisfied, and the last one becomes $x_2 k'^\mu = x_3 k^\mu$. Since k^μ and k'^μ are a priori not collinear, this implies that $x_2 = x_3 = 0$. This last type of singularity is called a *soft* or *infrared singularity*, since it occurs when all the components of the loop momentum go to zero.

Coleman–Norton Interpretation: A more intuitive physical interpretation of the Landau equations was found by Coleman and Norton (see Coleman, S. and Norton, R. E. (1965), *Nuovo Cimento* 38: 438). They propose to interpret a Feynman graph literally, as if the graph represents a process in spacetime, with the vertices being the locations where some instantaneous interactions happen. They further propose that the spacetime separation dx_r^μ between two vertices is parallel to the momentum q_r^μ, and proportional to the Feynman parameter x_r:

$$dx_r^\mu = x_r q_r^\mu.$$

(The constant of proportionality could be different from 1, but should be the same for all propagators. Its value is not important, as one may freely rescale the entire diagram without affecting the Landau equations.) Based on this identification, one may make the following observations:

- If a singularity happens at $x_r = 0$, the spacetime separation is zero, and we may shrink the corresponding propagator to a point. The resulting graph is called a *reduced graph*. A similar reduction is possible if $q_r^\mu = 0$, in the case of a soft divergence.
- If $x_r \neq 0$, the Landau equations tell us that we should have $q_r^2 = m_r^2$ instead. Thus, in this interpretation, the propagators of a reduced graph represent the on-shell propagation of a particle between two interactions. In this case, we may also write

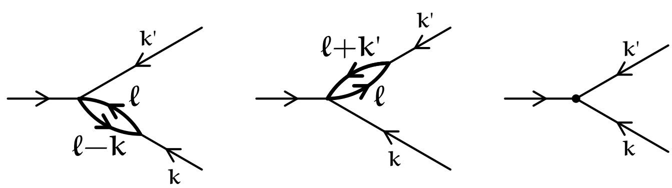
$$x_r = \frac{dx_r^0}{q_r^0}, \quad dx_r^\mu = dx_r^0 \frac{q_r^\mu}{q_r^0}, \quad \text{i.e.,} \quad \frac{dx_r^\mu}{dx_r^0} = \frac{q_r^\mu}{q_r^0} = v_r^\mu,$$

  where v_r^μ is the 4-velocity of an on-shell particle of momentum q_r^μ.
- The last of the Landau equations becomes

$$\sum_r \epsilon_{rj}\, dx_r^\mu = 0, \qquad \text{for every loop } j.$$

  This equation is consistent with dx_r^μ being a separation in spacetime, since adding these separations along a closed loop should obviously give zero. This also implies that the spacetime separation between two vertices does not depend on which path we follow on the graph to connect them, as it should for this interpretation to make sense.

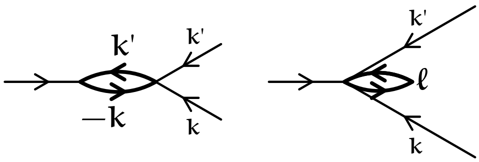
As an illustration, let us show the reduced diagrams for the three singularities we have found for the triangle one-loop graph studied above:



In more complicated cases, where solving the Landau equations may be difficult, the Coleman–Norton interpretation and the associated reduced diagrams can be used as guidance for identifying the possible solutions. This is based on the fact that there is a one-to-one correspondence

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

between the solutions of the Landau equations and the spacetime diagrams proposed by Coleman and Norton. For instance, in the case of the above example, we could use it as a way of checking that there are no other solutions. Indeed, there could in principle be two additional reduced diagrams:



However, neither of them is kinematically allowed; the left graph would have two non-collinear particles propagating along the same spacetime interval, and the right graph would have a physical free particle making a closed loop.

6. Multi-loop Integration in D Dimensions Start from equation (1.42) in Problem 5, in the special case where the numerator is trivial, i.e., $\mathcal{N} \equiv 1$. Note that the denominator $\mathcal{D}$ is a quadratic form in the components of the loop momenta, which we may arrange as follows:

$$\mathcal{D} \equiv L^t_\mu A L^\mu + L^t_\mu B^\mu + B^t_\mu L^\mu + C, \quad L^\mu \equiv (\ell_1^\mu, \ell_2^\mu, \dots, \ell_L^\mu).$$

6.a Perform explicitly the integration over the L loop momenta in D dimensions in order to obtain

$$\mathcal{I}(\{p_i\}) = \frac{i^L \Gamma(m - \frac{DL}{2})}{(-4\pi)^{\frac{DL}{2}}} \int_{x_r \geq 0} \prod_{r=1}^m dx_r \, \frac{\delta(1 - \sum_r x_r)\, (\det A)^{m - \frac{(L+1)D}{2}}}{\big(\underbrace{C \det A - B^t_\mu \, C^t_A \, B^\mu}_{\equiv \Delta} + i0^+\big)^{m - \frac{DL}{2}}},$$

where C_A is the matrix of co-factors of A.

6.b Show that $\Delta = \det \begin{pmatrix} A_{11} & \dots & A_{1L} & B_1^\mu \\ A_{21} & \dots & A_{2L} & B_2^\mu \\ \vdots & & \vdots & \vdots \\ A_{L1} & \dots & A_{LL} & B_L^\mu \\ B_{1\mu} & \dots & B_{L\mu} & C \end{pmatrix}$ and is of degree two in each variable x_r.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

6.a Using the notation of Problem 5, $q_r \equiv \sum_j \epsilon_{rj}\ell_j + \Delta_r$, the denominator $\mathcal{D}$ can be rewritten as

$$\begin{aligned}
\mathcal{D} &= \sum_r x_r(q_r^2 - m_r^2) \\
&= \sum_{j,k} \underbrace{\Big(\sum_r x_r\epsilon_{rj}\epsilon_{rk}\Big)}_{\equiv A_{jk}} \ell_j^\mu \ell_{k\mu} + 2\sum_j \underbrace{\Big(\sum_r x_r\epsilon_{rj}\Delta_r^\mu\Big)}_{\equiv B_j^\mu} \ell_{j\mu} + \underbrace{\sum_r x_r(\Delta_r^2 - m_r^2)}_{\equiv C} \\
&= L_\mu^t\, A\, L^\mu + L_\mu^t B^\mu + B_\mu^t L^\mu + C.
\end{aligned}$$

In this expression, note that:

- A is a real symmetric $L \times L$ matrix, whose coefficients are linear in the Feynman parameters x_r and are independent of the external momenta. This matrix is positive since $z_j A_{jk} z_k = \sum_r x_r(\sum_j \epsilon_{rj} z_j)^2$, and positive definite if all $x_r > 0$ (depending on the details of the Feynman graph under consideration, it can still be positive definite if some – but not all – of the x_r are zero).
- B^μ is a column vector with L components, linear in the Feynman parameters and linear in the external momenta.
- C is linear in the Feynman parameters, and also contains the squared masses and Lorentz invariant scalar products of external momenta.

Let us first rearrange the quadratic form $\mathcal{D}$ as follows:

$$\begin{aligned}
\mathcal{D} &= L_\mu^t\, A\, L^\mu + L_\mu^t B^\mu + B_\mu^t L^\mu + C \\
&= (L_\mu + A^{-1}B_\mu)^t \underbrace{A}_{\Omega^t D \Omega} \underbrace{(L^\mu + A^{-1}B^\mu)}_{K^\mu} + C - B_\mu^t\, A^{-1}\, B^\mu \\
&= (\Omega K_\mu)^t\, D\, (\Omega K^\mu) + C - B_\mu^t\, A^{-1}\, B^\mu \\
&= R_\mu R^\mu + C - B_\mu^t\, A^{-1}\, B^\mu, \quad \text{with } R_j^\mu \equiv \sqrt{D_{jj}}\, (\Omega K)_j^\mu.
\end{aligned}$$

The second line is the standard manipulation that eliminates the linear terms from the quadratic form (the Jacobian is 1, since this is just a translation). In the third line, we used the fact that the symmetric matrix A_{jk} (of size $L \times L$) is diagonalizable by an orthogonal transformation Ω (the Jacobian is also 1). In the last line, we have rescaled the various loop momenta in order to absorb the diagonal elements of D. In the final form, we may view $R_\mu R^\mu$ as the norm of a unique vector with DL components (D being the dimension of spacetime). The overall

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

Jacobian of this sequence of transformations is given by

$$d^{DL}L = \big(\det A\big)^{-D/2} \times d^{DL}R.$$

Therefore, the integration over the loop momenta reads

$$\begin{aligned}\int \frac{d^{DL}L}{(2\pi)^{DL}} \frac{1}{\mathcal{D}^m} &= \frac{1}{(\det A)^{D/2}} \int \frac{d^{DL}R}{(2\pi)^{DL}} \frac{1}{\big(R_\mu R^\mu + C - B^t_\mu A^{-1} B^\mu\big)^m} \\ &= \frac{i^L}{(\det A)^{D/2}} \int \frac{d^{DL}R_E}{(2\pi)^{DL}} \frac{(-1)^m}{\big(R_E^2 + B^t_\mu A^{-1} B^\mu - C\big)^m} \\ &= \frac{(-1)^m\, i^L}{(4\pi)^{\frac{DL}{2}}} \frac{\Gamma(m - \frac{DL}{2})}{\Gamma(m)} \frac{\big(B^t_\mu A^{-1} B^\mu - C\big)^{\frac{DL}{2}-m}}{(\det A)^{D/2}}.\end{aligned}$$

In the second line, we have applied a Wick rotation to the temporal components of the L loop momenta, and the result of the last line is a standard integration in DL dimensions (see the first of Eqs. (1.21)). Therefore, the expression for the L-loop integral is

$$\begin{aligned}\mathcal{I}(\{p_i\}) = \frac{(-1)^m i^L \Gamma(m - \frac{DL}{2})}{(4\pi)^{\frac{DL}{2}}} \int\limits_{x_r \geq 0} \prod_{r=1}^{m} dx_r\, \delta\big(1 - \sum_r x_r\big) \\ \times \frac{\big(B^t_\mu A^{-1} B^\mu - C - i0^+\big)^{\frac{DL}{2}-m}}{(\det A)^{D/2}}.\end{aligned}$$

(We have reinstated the $i0^+$ prescription of Feynman propagators.) Using $A^{-1} = \big(\det A\big)^{-1} C^t_A$, where C_A is the matrix of co-factors of A, we may rewrite this as

$$\mathcal{I}(\{p_i\}) = \frac{i^L \Gamma(m - \frac{DL}{2})}{(-4\pi)^{\frac{DL}{2}}} \int\limits_{x_r \geq 0} \prod_{r=1}^{m} dx_r\, \delta\big(1 - \sum_r x_r\big) \frac{\big(\det A\big)^{m - \frac{(L+1)D}{2}}}{\big(\underbrace{C \det A - B^t_\mu C^t_A B^\mu}_{\equiv \Delta} + i0^+\big)^{m - \frac{DL}{2}}}.$$

Written in this form, the fraction in the integrand is a rational function of the Feynman parameters, with $\Delta \equiv C \det A - B^t_\mu C^t_A B^\mu$ a homogeneous polynomial of degree $L + 1$ and $\det A$ a polynomial of degree L.

6.b This expression for the loop integral can shed extra light on the discussion of singularities in Problem 5, which led to the Landau equations. First, note that the determinant of A is zero only in accidental situations, since in general $\mathcal{D}$ is a negative definite quadratic form of the Euclidean loop momenta. Thus, the singularities in the above expression come from zeroes in Δ. This quantity is precisely the maximum of the denominator $\mathcal{D}$, viewed as a function of the Euclidean loop momenta ℓ^μ_j. Therefore, we recover the fact that the singularities of $\mathcal{I}(\{p_i\})$ may only occur when $\mathcal{D}$ has a vanishing maximum.

Recall now that $C \det A - B^t_\mu C^t_A B^\mu = C \det A - B^t_\mu (\det A)\, A^{-1} B^\mu$. Consider now the linear equation $A_{jk} X^\mu_k = B^\mu_j$ (where μ is treated as a fixed parameter). Its solution may be

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

written as

$$(\det A)\,(A^{-1})_{jk}B^{\mu}_{k} = (\det A)\,X^{\mu}_{j} = \det\left(A_{[1,j-1]}\,B^{\mu}\,A_{[j+1,L]}\right),$$

where the notation $A_{[1,j-1]}\,B^{\mu}\,A_{[j+1,L]}$ stands for the $L \times L$ matrix obtained with columns 1 to $j-1$ of A, the column vector B^{μ}, and columns $j+1$ to L of A. Therefore, we have

$$\begin{aligned}\Delta &= C\det A - B^{t}_{j\mu}\det\left(A_{[1,j-1]}\,B^{\mu}\,A_{[j+1,L]}\right)\\ &= C\det A - B^{t}_{j\mu}\,(-1)^{L-j}\det\left(A_{[1,j-1]}\,A_{[j+1,L]}\,B^{\mu}\right)\\ &= \det\begin{pmatrix} A_{11} & \dots & A_{1L} & B^{\mu}_{1}\\ A_{21} & \dots & A_{2L} & B^{\mu}_{2}\\ \vdots & & \vdots & \vdots\\ A_{L1} & \dots & A_{LL} & B^{\mu}_{L}\\ B_{1\mu} & \dots & B_{L\mu} & C\end{pmatrix}.\end{aligned} \tag{1.43}$$

In the second line, we just need to count the number of column permutations necessary to bring the column B^{μ} to the rightmost position. The last equality can be checked by expanding the determinant on the right-hand side according to the minors of the last line. Observe now that the function of $x_1, x_2, \dots, x_m$ obtained after integrating over the loop momenta does not depend on how we labeled the loop momenta (possibly up to a permutation of the $\{x_r\}$). Therefore, let us assume that, for the internal propagator $r = 1$, we have made the choice

$$q^{\mu}_{1} = \ell^{\mu}_{1}, \quad \text{i.e.,} \;\; \epsilon_{1j} = \delta_{1j}, \;\; \Delta^{\mu}_{1} = 0.$$

With this choice, we have

$$A_{jk} = x_1\delta_{1j}\delta_{1k} + \sum_{r\geq 2} x_r\epsilon_{rj}\epsilon_{rk}, \qquad B^{\mu}_{j} = \sum_{r\geq 2} x_r\epsilon_{rj}\Delta^{\mu}_{r},$$
$$C = -m_1^2x_1 + \sum_{r\geq 2} x_r(\Delta_r^2 - m_r^2).$$

In particular, the only coefficients that depend (linearly) on x_1 are A_{11} and C. By expanding the determinant in (1.43), we see that Δ is a quadratic function of x_1 (with the other x_r fixed). Since our choice of setting $q_1 = \ell_1$ was arbitrary, we conclude that Δ is of degree two in every variable x_r (and overall homogeneous of degree $L+1$).

7. Weinberg Convergence Theorem The goal of this problem is to establish a criterion for ultraviolet convergence for loop integrals in scalar field theories (this restriction simplifies the problem a bit, since the integrands have a trivial numerator equal to one). This result, known as *Weinberg's convergence theorem*, states that a loop integral is ultraviolet convergent if the superficial degree of divergence of the loop integral, and of any of its restrictions to hyperplanes obtained by setting linear combinations of the loop momenta to constants, is negative. To establish this, we consider general Euclidean integrals of the form

$$\mathcal{I}(C,q,m) \equiv \int \frac{d^4\ell_1\cdots d^4\ell_m}{\prod_{j=1}^{n}\left((C_{ji}\ell_i + q_j)^2 + m_j^2\right)},$$

where C_{ji} is an $n \times m$ constant matrix.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

7.a Why can one consider instead the simpler integrals

$$\mathcal{J}(C) \equiv \int_D \frac{d^4\ell_1 \cdots d^4\ell_m}{\prod_{j=1}^{n} (C_{ji}\ell_i)^2}, \quad D \equiv \{(\ell_1^\mu, \ldots, \ell_m^\mu) | 1 \leq (C\ell)^2\},$$

that do not depend on the masses or on the external momenta?

7.b Show that $\mathcal{I}(C, q, m)$ is absolutely convergent if $4m - 2n < 0$ and if the restriction of this integral to any hyperplane of co-dimension 4 is also absolutely convergent.

Hint: define $k_j \equiv C_{ji}\ell_i$ *and write the integral* $\mathcal{J}(C)$ *as a sum of terms* $\mathcal{J}_j(C)$ *in which the squared norm* k_j^2 *is the smallest among all the* k_ℓ^2*'s, then rescale all the* $k_{\ell \neq j}$ *by* $|k_j|$. *Show that the integral* $\mathcal{J}_j(C)$ *factorizes into a one-dimensional integral whose convergence is determined by power counting, and an integral over a subspace of co-dimension* 4 *of the same type as the original integral.*

7.c Show that an equivalent convergence criterion is that the superficial degree of divergence of $\mathcal{I}(C, q, m)$, and of restrictions of $\mathcal{I}$ to any hyperplane defined by setting some linear combinations of the momenta to constants, is negative.

7.a Since we are interested only in the ultraviolet convergence of this integral, the masses m_j and the shifts q_j do not play any role at large momenta (even though they matter for the convergence in the infrared and for the precise value of the integral). As far as ultraviolet convergence is concerned, we may as well set these parameters to zero, but we need to cut out a small region around $\ell_j = 0$ in order to avoid infrared problems. Thus, we may consider instead

$$\mathcal{J}(C) \equiv \int_D \frac{d^4\ell_1 \cdots d^4\ell_m}{\prod_{j=1}^{n} (C_{ji}\ell_i)^2}, \quad D \equiv \{(\ell_1^\mu, \ldots, \ell_m^\mu) | 1 \leq (C\ell)^2\}, \tag{1.44}$$

which has exactly the same ultraviolet behavior as the original integral. (For technical reasons that will become clear later, it turns out to be a bit simpler to remove a ball of radius unity in the space of the variables $k_j \equiv C_{ji}\ell_j$ rather than ℓ_i.)

7.b Note that the first of the two conditions listed in the statement of the problem for the ultraviolet convergence of this integral, namely $4m - 2n < 0$, is nothing but the demand that the superficial degree of divergence of this integral be negative. This is a necessary condition because the integrand cannot decrease faster than ξ^{4m-2n} when all the ℓ_i's are rescaled according to $\ell_i \to \xi\ell_i$. The reason why this condition alone is insufficient to ensure the convergence is that, depending on the matrix C_{ji}, there could be directions in $\mathbb{R}^{4m}$ along which the decrease of the integrand is slower, for instance if one or more of the ℓ_i's do not appear in one of the k_j's. As we shall see, the second condition ensures that this problematic situation does not occur.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

Then, we may write the integral $\mathcal{J}(C)$ as a sum of integrals $\mathcal{J}_j(C)$ in which the squared norm k_j^2 is the smallest among all the k_l^2's:

$$\mathcal{J}(C) = \sum_{j=1}^{n} \mathcal{J}_j(C), \quad \mathcal{J}_j(C) \equiv \int_{D_j} \frac{d^4\ell_1 \cdots d^4\ell_m}{\prod_{l=1}^{n} k_l^2},$$
$$D_j \equiv \big\{(\ell_1^\mu, \dots, \ell_m^\mu)\big| 1 \leq k_j^2 \leq \cdots \big\}.$$

The next step is to perform a linear change of the integration variables, $t \equiv A\ell$, such that

$$\det A = 1, \quad t_1 = k_j.$$

In terms of the t_i^μ, the integral $\mathcal{J}_j(C)$ may be rewritten as

$$\mathcal{J}_j(C) = \int_{t_1^2 \geq 1} d^4t_1 \int_U \frac{d^4t_2 \cdots d^4t_m}{\prod_{l=1}^{n} (C_{li}A_{ik}^{-1}t_k)^2}, \quad U \equiv \big\{(t_2^\mu, \dots, t_n^\mu)\big| t_1^2 \leq (CA^{-1}t)^2\big\}.$$

Then, we rescale $t_2, \dots, t_n$ by the Euclidean norm of t_1,

$$t_i^\mu \equiv |t_1|\, \ell_i'^\mu \quad \big(|t_1| \equiv \sqrt{t_1^2}\big),$$

which leads to the following form of the integral $\mathcal{J}_j(C)$:

$$\mathcal{J}_j(C) = \int_{t_1^2 \geq 1} d^4t_1\; |t_1|^{4m-2n-4} \times \int_{U'} \frac{d^4\ell_2' \cdots d^4\ell_m'}{\prod_{l \neq j} (C_{lk}'\ell_k')^2}, \tag{1.45}$$
$$\text{with } \; C_{lk}' \equiv C_{li}A_{ik}^{-1}, \quad U' \equiv \big\{(\ell_2'^\mu, \dots, \ell_n'^\mu)\big| 1 \leq (C'\ell')^2\big\}.$$

In this form, it is clear that:

- The first integral on the right-hand side is absolutely convergent if $4m - 2n < 0$.
- The second integral is of the form (1.44) with one less integration variable. Since the two integrals on the right-hand side are independent, it must also be convergent for $\mathcal{J}_j(C)$ to be finite.

This proves the announced convergence criterion.

7.c If we apply this criterion recursively, we get integrals of lower and lower dimensionality, until the last step where the analogue of (1.45) contains only the first factor (this corresponds to a one-loop integral). When we reach this point, we can conclude about the convergence from power counting only. Thus, the convergence of the original integral is ensured if its degree of divergence $4m - 2n$, and the degree of divergence of all the sub-integrals obtained by restricting the integration domain to hyperplanes of lower dimension, are negative. In this form, the convergence criterion is known as *Weinberg's convergence theorem* (note that here we have studied a less general situation than in the original theorem, which also considers the possibility of a polynomial of the integration variables in the numerator of the integrand). Let us add a final remark: the criterion derived in this problem is a criterion for *absolute convergence*. When it is not satisfied, it could still happen that the integral is nevertheless

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

(weakly) convergent because of cancellations among various parts of the integration domain. And of course, there could also be cancellations among the contributions of various Feynman graphs. This is typically what happens in gauge theories, where the gauge symmetry can induce cancellations among graphs that form a gauge invariant set (a single graph is usually not gauge invariant).
Source: Hahn, Y. and Zimmermann, W. (1968), *Commun Math Phys* 10: 330.

8. Electron Anomalous Magnetic Moment Consider the amputated renormalized photon–electron–positron vertex function $\Gamma_r^\mu(k, p, q)$ (with the convention that all momenta are incoming).

8.a Using the Dirac equation, prove the following relationship:

$$(p^\mu - q^\mu)\,\overline{u}(-\mathbf{q})u(\mathbf{p}) = 2m\,\overline{u}(-\mathbf{q})\left[\gamma^\mu + \frac{i}{m}M^{\mu\nu}(p+q)_\nu\right]u(\mathbf{p}),$$

known as *Gordon's identity* (with $M^{\mu\nu} \equiv \frac{i}{4}[\gamma^\mu, \gamma^\nu]$).

8.b Show that its contribution to electron scattering off an external field $\mathcal{A}^\mu$ can be parameterized as follows:

$$\mathcal{A}_\mu(k)\overline{u}(-\mathbf{q})\Gamma_r^\mu(k,p,q)u(\mathbf{p}) = e_r\mathcal{A}_\mu(k)\overline{u}(-\mathbf{q})\left[F_1(k^2)\gamma^\mu + i\,F_2(k^2)\frac{M^{\mu\nu}k_\nu}{m_r}\right]u(\mathbf{p}).$$

Hint: the most general form of Γ_r^μ *is*

$$\Gamma_r^\mu = C_1^\mu \mathbf{1} + C_2^{\mu\alpha}\gamma_\alpha + C_3^\mu\gamma_5 + C_4^{\mu\alpha}\gamma_\alpha\gamma_5 + C_5^{\mu\alpha\beta}M_{\alpha\beta}.$$

Then, use Lorentz invariance and the Ward–Takahashi identity to show that $C_1^\mu \propto (p^\mu - q^\mu)$, *and then the Gordon identity to bring this term to the announced form. The same reasoning can be used to bring the remaining four terms in* Γ_r^μ *to the announced form.*

8.c Approximate this formula for a constant magnetic field. In particular, for an electron at rest in a homogeneous magnetic field B in the x^3 direction, show that

$$\mathcal{A}_\mu(k)\overline{u}(-\mathbf{q})\Gamma_r^\mu(k,p,q)u(\mathbf{p}) \quad \text{becomes} \quad \frac{e_r B}{m_r}\left(1 + F_2(0)\right)\overline{u}(0)M^{12}u(0).$$

8.d Calculate the relevant parts of Γ_r^μ at one loop in order to show that $F_2(0) = \frac{\alpha}{2\pi}$. (This result led to one of the first experimental verifications of quantum electrodynamics.)

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

8.a Let us rewrite the Poincaré algebra generator $M^{\mu\nu}$ as follows:

$$M^{\mu\nu} = \frac{i}{4}\big(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\big) = \frac{i}{2}\big(\gamma^\mu\gamma^\nu - g^{\mu\nu}\big) = \frac{i}{2}\big(g^{\mu\nu} - \gamma^\nu\gamma^\mu\big).$$

Then, consider

$$\begin{aligned}\overline{u}(-\mathbf{q})M^{\mu\nu}(p_\nu + q_\nu)u(\mathbf{p}) &= \frac{i}{2}\overline{u}(-\mathbf{q})\big(\gamma^\mu\gamma^\nu - g^{\mu\nu}\big)p_\nu u(\mathbf{p}) \\ &\quad + \frac{i}{2}\overline{u}(-\mathbf{q})\big(g^{\mu\nu} - \gamma^\nu\gamma^\mu\big)q_\nu u(\mathbf{p}) \\ &= im\,\overline{u}(-\mathbf{q})\gamma^\mu u(\mathbf{p}) - \frac{i(p^\mu - q^\mu)}{2}\,\overline{u}(-\mathbf{q})u(\mathbf{p}).\end{aligned}$$

This relation, known as *Gordon's identity*, can be rewritten as

$$(p^\mu - q^\mu)\,\overline{u}(-\mathbf{q})u(\mathbf{p}) = 2m\,\overline{u}(-\mathbf{q})\left[\gamma^\mu + \frac{i}{m}M^{\mu\nu}(p+q)_\nu\right]u(\mathbf{p}). \tag{1.46}$$

8.b The object $\overline{u}(-\mathbf{q})\Gamma_r^\mu u(\mathbf{p})$ is a radiative correction to an electromagnetic current. Since the incoming and outgoing electrons are on-shell, it must satisfy the following Ward–Takahashi identity:

$$k_\mu\,\overline{u}(-\mathbf{q})\Gamma_r^\mu(k,p,q)u(\mathbf{p}) = 0.$$

Moreover, this identity must be satisfied even if the photon is off-shell, $k^2 \neq 0$. Obviously, the right-hand side of the announced formula obeys this identity, after one uses the Dirac equation and the antisymmetry of $M^{\mu\nu}$.

Let us now give a glimpse of how one would prove the converse, namely that this form is the only possible one. The starting point is to note that Γ_r^μ is a 4×4 matrix carrying a pair of Dirac indices, which we may decompose on the basis $\{\mathbf{1}, \gamma_\alpha, \gamma_5, \gamma_\alpha\gamma_5, M_{\alpha\beta}\}$:

$$\Gamma_r^\mu = C_1^\mu\mathbf{1} + C_2^{\mu\alpha}\gamma_\alpha + C_3^\mu\gamma_5 + C_4^{\mu\alpha}\gamma_\alpha\gamma_5 + C_5^{\mu\alpha\beta}M_{\alpha\beta}.$$

Since these terms are linearly independent, they must fulfill the Ward–Takahashi identity independently. From charge conjugation and parity symmetry, the term in γ_5 should be zero, $C_3^\mu \equiv 0$. The Lorentz indices of the remaining coefficients must be carried by the vectors p, q (or $k = -p - q$), the metric tensor, and possibly the Levi–Civita symbol in the case of $\gamma_\alpha\gamma_5$, with prefactors that depend only on Lorentz invariant quantities. In fact, all Lorentz invariant quantities can be expressed in terms of the electron mass and the photon virtuality k^2, since

$$p \cdot q = k^2 - m^2, \quad p \cdot k = q \cdot k = -\tfrac{k^2}{2}.$$

Consider for instance the coefficient C_1^μ of the identity. It may be written as

$$C_1^\mu = C_{1a}(k^2)\,p^\mu + C_{1b}(k^2)\,k^\mu.$$

The Ward–Takahashi identity implies

$$0 = C_{1a}\,k \cdot p + C_{1b}\,k^2 = k^2\left(C_{1b} - \tfrac{1}{2}C_{1a}\right).$$

Therefore, this coefficient must have the following form:

$$C_1^\mu = C_{1b}(k^2)\big(2p^\mu + k^\mu\big) = C_{1b}(k^2)\,(p^\mu - q^\mu).$$

Thanks to the Gordon identity (1.46), this term of Γ_r^μ indeed has the general form quoted in the statement of the problem. This turns out to be true for all the terms in this decomposition, as one

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

may check by first writing the most general Lorentz structure allowed by the Ward–Takahashi identity and then by using the Dirac equation to simplify the result after insertion between $\overline{u}(-\mathbf{q})\cdots u(\mathbf{p})$.

8.c The coefficients $F_1(k^2)$ and $F_2(k^2)$, called the *form factors* of the electron, describe the properties of the cloud of photons and virtual pairs that surround the electron, as one varies the virtuality k^2 of the photon that probes the electron (in a sense, k^2 plays the role of a the resolution scale at which the photon probes this cloud). In the limit $k^2 \to 0$, the photon probes this cloud on very large distance scales. Given its similarity with the bare vertex, the term proportional to γ^μ encodes the electrical charge seen by the photon. On very large distance scales, this must be the usual charge of the electron as we know it from atomic physics, which means that $F_1(0) = 1$. Let us now discuss the meaning of $F_2(0)$. Consider a constant external magnetic field $\mathcal{B}$ in the x^3 direction, corresponding to

$$\mathcal{A}_2(x) = \mathcal{B}x^1, \quad \mathcal{A}_2(k) = i\mathcal{B}\partial_{k^1}, \quad \mathcal{F}_{12}(x) = -\mathcal{F}_{12}(x) = \mathcal{B},$$

where $\mathcal{F}_{\mu\nu}$ is the field strength. When an electron at rest is embedded in this field, the dressed coupling of the photon to the current reads (in the limit $k \to 0$)

$$e_r\mathcal{B}\,\overline{u}(-\mathbf{p})\left[i\gamma^2\partial_{p_1} + F_2(0)\frac{M^{12}}{m_r}\right]u(\mathbf{p})\Big|_{\mathbf{p}=0}.$$

The derivative ∂_{p_1} acting on a spinor of small momentum acts like the boost K^1 in the direction x^1. More precisely, we have

$$\begin{aligned}\overline{u}(-\mathbf{p})i\gamma^2\partial_{p_1}u(\mathbf{p})\Big|_{\mathbf{p}=0} &= \overline{u}(-\mathbf{p})i\gamma^2\frac{iK^1}{m_r}u(\mathbf{p})\Big|_{\mathbf{p}=0} = -\overline{u}(-\mathbf{p})i\gamma^2\frac{\gamma^1\gamma^0}{2m_r}u(\mathbf{p})\Big|_{\mathbf{p}=0}\\ &= \frac{1}{m_r}\overline{u}(-\mathbf{p})M^{12}u(\mathbf{p})\Big|_{\mathbf{p}=0}.\end{aligned}$$

(The Dirac equation for zero momentum spinors reduces to $\gamma^0 u(\mathbf{p}) = u(\mathbf{p})$.) Therefore, the above coupling becomes

$$\frac{e_r\mathcal{B}}{m_r}\left(1 + F_2(0)\right)\overline{u}(-\mathbf{p})M^{12}u(\mathbf{p})\Big|_{\mathbf{p}=0}.$$

The 1 in $1 + F_2(0)$ encodes the bare coupling of the magnetic field to the electron, and $F_2(0)$ is therefore a correction to this coupling.

8.d At one loop, the QED Feynman rules lead to the following expression for the vertex function contracted between the spinors of the incoming and outgoing fermions:

$$\begin{aligned}\overline{u}(-\mathbf{q})\Gamma^\mu(k,p,q)u(\mathbf{p}) &= -i\,e^3\int\frac{d^D\ell}{(2\pi)^D}\,\frac{\overline{u}(-\mathbf{q})\gamma^\sigma(\not{\ell}-\not{q}+m)\gamma^\mu(\not{\ell}+\not{p}+m)\gamma_\sigma u(\mathbf{p})}{\ell^2((\ell-q)^2-m^2)((\ell+p)^2-m^2)}\\ &\equiv -i\,e^3\int\frac{d^D\ell}{(2\pi)^D}\,\frac{\mathcal{N}^\mu}{\mathcal{D}}.\end{aligned}$$

For the time being, we use dimensional regularization to make all intermediate expressions finite, but we shall see shortly that the form factor F_2 is ultraviolet finite. The denominators of

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

the three propagators can be combined into a single one by introducing Feynman parameters:

$$\frac{1}{ABC} = 2\int_0^1 dx_1 dx_2 dx_3 \; \frac{\delta(x_1 + x_2 + x_3 - 1)}{(x_1 A + x_2 B + x_3 C)^3}.$$

In the situation of interest here, the resulting denominator reads

$$\begin{aligned} & x_1((\ell + p)^2 - m^2) + x_2((\ell - q)^2 - m^2) + x_3\,\ell^2 \\ & \quad = L^2 + x_1(1 - x_1)p^2 + x_2(1 - x_2)q^2 - (x_1 + x_2)m^2 + x_1 x_2\, p \cdot q \\ & \quad = L^2 \underbrace{-(x_1 + x_2)^2 m^2 + x_1 x_2 k^2}_{\equiv -\Delta}, \end{aligned}$$

where we have defined $L \equiv \ell + x_1 p - x_2 q$ and where in the final line we have assumed the electrons to be on-shell ($p^2 = q^2 = m^2$, $k = -(p + q)$).

The next step is to express the numerator $\mathcal{N}^\mu$ in terms of the new integration variable L:

$$\begin{aligned} \mathcal{N}^\mu = {} & \overline{u}(-\mathbf{q})\gamma^\sigma \not{L} \gamma^\mu \not{L} \gamma_\sigma u(\mathbf{p}) \\ & + \text{terms linear in } L \\ & + \overline{u}(-\mathbf{q})\gamma^\sigma (m - x_1 \not{p} - (1 - x_2)\not{q})\gamma^\mu (m + (1 - x_1)\not{p} + x_2 \not{q})\gamma_\sigma u(\mathbf{p}). \end{aligned} \qquad (1.47)$$

The terms linear in L can be dropped since the denominator is even in L. The combination of Dirac matrices in the first line can be rewritten as

$$\gamma^\sigma \not{L} \gamma^\mu \not{L} \gamma_\sigma = \frac{(2 - D)^2}{D}\, L^2\, \gamma^\mu.$$

Since it is quadratic in L, this terms leads to a logarithmic ultraviolet divergence, but the proportionality to γ^μ indicates that it contributes only to the form factor F_1 (the ultraviolet divergence in F_1 leads to a renormalization of the electron electrical charge). Since our goal is to evaluate F_2, we can disregard this term from now on. The term on the last line of (1.47) does not depend on L and therefore gives an ultraviolet finite integral over L, which implies that we can perform the Dirac algebra for this term in $D = 4$. Using

$$\gamma^\mu \gamma^\nu \gamma_\mu = -2\,\gamma^\nu\,, \quad \gamma^\mu \gamma^\nu \gamma^\rho \gamma_\mu = 4\, g^{\nu\rho}\,, \quad \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma_\mu = -2\,\gamma^\sigma \gamma^\rho \gamma^\nu$$

and the Dirac equations obeyed by the two spinors, a straightforward but somewhat tedious calculation leads to

$$\begin{aligned} & \overline{u}(-\mathbf{q})\gamma^\sigma (m - x_1 \not{p} - (1 - x_2)\not{q})\gamma^\mu (m + (1 - x_1)\not{p} + x_2 \not{q})\gamma_\sigma u(\mathbf{p}) \\ & \quad = 4m\left(\overline{u}(-\mathbf{q})u(\mathbf{p})\right)\left(\tfrac{p^\mu - q^\mu}{2}(x_3 - x_3^2) + \tfrac{k^\mu}{2}(x_1 + x_1^2 - x_2 - x_2^2)\right) \\ & \qquad + 2\left(\overline{u}(-\mathbf{q})\gamma^\mu u(\mathbf{p})\right)\Big(m^2(x_3^2 + 2x_3 - 1) - k^2(x_3 + x_1 x_2)\Big). \end{aligned}$$

The term of the last line, in $\overline{u}(-\mathbf{q})\gamma^\mu u(\mathbf{p})$, contributes only to F_1 and can be ignored. In the first line on the right-hand side, the term in k^μ is odd under the exchange of x_1 and x_2, while the denominator is even. Its integral over the Feynman parameters is therefore zero. Thus, we

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

need only consider further the term in $p^\mu - q^\mu$, which can be rearranged thanks to the Gordon identity (1.46):

$$\begin{aligned}&\overline{u}(-\mathbf{q})\gamma^\sigma(m - x_1\not{p} - (1-x_2)\not{q})\gamma^\mu(m + (1-x_1)\not{p} + x_2\not{q})\gamma_\sigma u(\mathbf{p})\\&\quad = \frac{i}{m}\big(\overline{u}(-\mathbf{q})M^{\mu\nu}u(\mathbf{p})\big)k_\nu(-4m^2)(x_3 - x_3^2)\\&\qquad + \text{terms in } \overline{u}(-\mathbf{q})\gamma^\mu u(\mathbf{p})\\&\qquad + \text{terms that integrate to zero.}\end{aligned}$$

By performing the integral over L (using a Wick rotation to have a Euclidean integration momentum),

$$\int \frac{d^D L}{(2\pi)^D}\,\frac{1}{(L^2-\Delta)^3} = -i\,\frac{\Delta^{D/2-3}}{(4\pi)^{D/2}}\,\frac{\Gamma\big(3-\frac{D}{2}\big)}{\Gamma(3)} \underset{D=4}{=} -\frac{i}{32\pi^2\,\Delta},$$

and then by comparing the term in $M^{\mu\nu}$ with the expression for the vertex function in terms of the electron form factors, we obtain

$$F_2(k^2) = \frac{e^2}{4\pi^2}\int_0^1 dx_1 dx_2 dx_3\;\delta(x_1+x_2+x_3-1)\;\frac{x_3 - x_3^2}{(1-x_3)^2 - x_1 x_2(k^2/m^2)}.$$

Noting that

$$\int_0^1 dx_1 dx_2\;\delta(x_1+x_2+x_3-1) = 1 - x_3,$$

we finally get

$$F_2(0) = \frac{e^2}{8\pi^2} = \frac{\alpha}{2\pi}.$$

This is the celebrated result obtained by Schwinger in 1948. As of 2019, the QED calculation of this quantity has been pushed to order α^5, allowing a comparison with experimental measurements of the electron anomalous magnetic moment with an unprecedented accuracy.

9. Ward–Takahashi Identities and Lorentz Invariance The goal of this problem is to discuss the relationship between the Lorentz invariance of scattering cross-sections involving photons and the Ward–Takahashi identities obeyed by amplitudes with external photons. We will see that the Ward–Takahashi identities imply the Lorentz invariance, but also that the converse is true to some extent: namely a theory of massless spin-1 particles must obey at least some weak form of Ward–Takahashi identities in order to be Lorentz invariant. To that end, we consider an S-matrix element with an external photon, written as $\epsilon_\lambda^\mu(\mathbf{p})\mathcal{M}_\mu(p,\dots)$.

9.a What is the transformation law of $\mathcal{M}_\mu(p,\dots)$ under a Lorentz transformation Λ?

9.b Show that, under the same Lorentz transformation, the polarization vectors satisfy

$$\Lambda_\nu{}^\mu \epsilon_\pm^\nu(\Lambda \mathbf{p}) = e^{\mp i\theta} \epsilon_\pm^\mu(\mathbf{p}) + \text{terms in } p^\mu,$$

where we use the basis of positive and negative helicities as the two physical polarizations. *Hint: show that the polarization vectors* $\epsilon_\pm^\mu(\mathbf{p})$ *can be obtained from those of the reference momentum* $q^\mu \equiv (\omega, 0, 0, \omega)$ *by applying the Lorentz transformation* $L(p)$ *that maps* q^μ *into* p^μ*. Then, use the fact that* $R \equiv L^{-1}(p)\Lambda^{-1}L(\Lambda p)$ *belongs to the little group, and the result of Problem 2.*

9.c Show that

$$\left|\epsilon_\pm^\mu(\Lambda \mathbf{p})\, \mathcal{M}_\mu(\Lambda p, \dots)\right|^2 = \left|\epsilon_\pm^\mu(\mathbf{p})\, \mathcal{M}_\mu(p, \dots)\right|^2 + \text{terms in } \; p^\mu\, \mathcal{M}_\mu(p, \dots).$$

Thus, the S-matrix element is Lorentz invariant up to a phase, provided that $\mathcal{M}_\mu(p, \dots)$ satisfies the Ward–Takahashi identity. Discuss whether the converse is true.

9.a The matrix element $\mathcal{M}_\mu(p, \dots)$ transforms covariantly under a Lorentz transformation, i.e.,

$$\mathcal{M}_\mu(p, \dots) = \left[\Lambda^{-1}\right]_\mu{}^\nu \mathcal{M}_\nu(\Lambda p, \dots) = \Lambda^\nu{}_\mu \mathcal{M}_\nu(\Lambda p, \dots).$$

(Here, we are writing only the factor Λ that corresponds to the external line of momentum p^μ; there should be one extra such factor for each other external particle with non-zero spin.)

9.b In order to construct a basis of physical polarization vectors associated with the momentum p^μ, let us start from the two helicity polarization vectors of the reference momentum $q^\mu \equiv (\omega, 0, 0, \omega)$,

$$\epsilon_\pm^\mu(\mathbf{q}) \equiv \tfrac{1}{\sqrt{2}}(0, 1, \pm i, 0).$$

The polarization vectors of an arbitrary momentum p^μ can then be obtained by applying the spatial rotation $\mathcal{R}(\widehat{\mathbf{p}})$ that brings the axis 3 into the direction $\widehat{\mathbf{p}}$:

$$\epsilon_\pm^\mu(\mathbf{p}) = \left[\mathcal{R}(\widehat{\mathbf{p}})\right]^\mu{}_\nu\, \epsilon_\pm^\nu(\mathbf{q}). \tag{1.48}$$

Let us introduce the Lorentz transformation $L(p)$ that transforms the reference vector q^μ into the vector p^μ. This transformation can be decomposed into a boost $\mathcal{B}_3(p^0/\omega)$ in the direction 3 that rescales the vector q^μ so that it has the same temporal component as p^μ, followed by the rotation $\mathcal{R}(\widehat{\mathbf{p}})$ that rotates its spatial components into the correct orientation:

$$L(p) = \mathcal{R}(\widehat{\mathbf{p}})\, \mathcal{B}_3(p^0/\omega).$$

Note that the boost $\mathcal{B}_3$ does not affect the reference polarization vectors, since their only non-zero components are in the directions $1, 2$. Therefore, we also have

$$\epsilon_\pm^\mu(\mathbf{p}) = \left[L(p)\right]^\mu{}_\nu\, \epsilon_\pm^\nu(\mathbf{q}).$$

Recall now that, for any momentum p^μ and Lorentz transformation Λ, $R \equiv L^{-1}(p)\Lambda^{-1}L(\Lambda p)$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

is an element of the little group since it leaves the reference vector q^μ invariant:

$$q^\mu \underset{L(\Lambda \mathbf{p})}{\longrightarrow} (\Lambda p)^\mu \underset{\Lambda^{-1}}{\longrightarrow} p^\mu \underset{L^{-1}(\mathbf{p})}{\longrightarrow} q^\mu.$$

Let us apply this transformation to the reference polarization vectors, which leads to

$$\left[\Lambda^{-1}\right]^\mu{}_\nu \underbrace{\left[L(\Lambda p)\right]^\nu{}_\rho \, \epsilon^\rho_\pm(\mathbf{q})}_{\epsilon^\nu_\pm(\Lambda \mathbf{p})} = \left[L(p)\right]^\mu{}_\nu \, R^\nu{}_\rho \, \epsilon^\rho_\pm(\mathbf{q}).$$

In order to evaluate the right-hand side, we use the explicit representation of little-group elements obtained in Problem 2, which leads to

$$R^\nu{}_\rho \, \epsilon^\rho_\pm(\mathbf{q}) = e^{\mp i\theta} \epsilon^\nu_\pm(\mathbf{q}) + \beta_\pm q^\nu,$$

where we denote $\beta_\pm \equiv (\beta_1 \pm i\beta_2)/(\omega\sqrt{2})$. Multiplying this equation on the left by $[L(p)]^\mu{}_\nu$, we obtain

$$\left[\Lambda^{-1}\right]^\mu{}_\nu \, \epsilon^\nu_\pm(\Lambda \mathbf{p}) = e^{\mp i\theta} \, \epsilon^\mu_\pm(\mathbf{p}) + \beta_\pm \, p^\mu. \tag{1.49}$$

This formula shows that the physical polarization vectors do not transform as Lorentz vectors: they transform as 4-vectors only up to a phase and up to an additive term proportional to the momentum.

9.c By multiplying Eqs. (1.48) and (1.49), we obtain

$$\begin{aligned} \left(e^{\mp i\theta} \, \epsilon^\mu_\pm(\mathbf{p}) + \beta_\pm \, p^\mu\right) \mathcal{M}_\mu(p, \dots) &= \underbrace{\Lambda^\nu{}_\mu \, [\Lambda^{-1}]^\mu{}_\rho}_{\delta^\nu{}_\rho} \, \epsilon^\rho_\pm(\Lambda \mathbf{p}) \mathcal{M}_\nu(\Lambda p, \dots) \\ &= \epsilon^\mu_\pm(\Lambda \mathbf{p}) \, \mathcal{M}_\mu(\Lambda p, \dots). \end{aligned}$$

The right-hand side is simply the S-matrix element evaluated in a boosted frame. The first term on the left-hand side is the same matrix element in the original frame, multiplied by a phase. Thus, we have

$$\left|\epsilon^\mu_\pm(\Lambda \mathbf{p}) \, \mathcal{M}_\mu(\Lambda p, \dots)\right|^2 = \left|\epsilon^\mu_\pm(\mathbf{p}) \, \mathcal{M}_\mu(p, \dots)\right|^2 + \text{terms in } \; p^\mu \, \mathcal{M}_\mu(p, \dots).$$

From this equation, we see immediately that the Ward–Takahashi identities (i.e., the fact that the last term on the right-hand side is zero) imply the Lorentz invariance of squared matrix elements with photons.

Let us now consider the converse. The Lorentz invariance of cross-sections requires that

$$\left|\epsilon^\mu_\pm(\Lambda \mathbf{p}) \, \mathcal{M}_\mu(\Lambda p, \dots)\right|^2 = \left|\epsilon^\mu_\pm(\mathbf{p}) \, \mathcal{M}_\mu(p, \dots)\right|^2,$$

which by the above identity implies that $p^\mu \, \mathcal{M}_\mu(p, \dots) = 0$. This looks like the usual Ward–Takahashi identity, except for an important restriction: all the other lines are on-shell and contracted with physical polarizations if they are photons (while in the Ward–Takahashi identities derived from current conservation, the photons do not need to be on-shell, nor do they need to have physical polarizations). In other words, any theory of massless spin-1 particles must satisfy a weak form of the Ward–Takahashi identities (in fact, identical to the non-Abelian version of the Ward–Takahashi identity), in order for its physical predictions to be consistent with Lorentz symmetry.

Additional Note: In the limit of soft massless spin-1 particles, the eikonal approximation leads to the following universal (i.e., independent of the spin of the emitters) form for the amplitude for producing a photon of momentum k in addition to hard particles of momenta p_i:

$$\mathcal{M}^\mu(k; p_1, \ldots, p_n) = \mathcal{M}(p_1, \ldots, p_n) \sum_{i=1}^{n} \frac{e_i\, p_i^\mu}{p_i \cdot k},$$

where the e_i are the couplings of the spin-1 particles to the emitters (conventionally defined to be all outgoing). In this limit, the condition $k_\mu \mathcal{M}^\mu = 0$ is equivalent to $\sum_i e_i = 0$. In other words, the Lorentz invariance of scattering amplitudes implies that any soft massless spin-1 particle couples to other fields via a conserved charge. The limitation of this argument to *soft* momenta (i.e., long-distance interactions) is easy to understand: a hard spin-1 would probe the substructure of composite objects, instead of its total charge (this does not mean that the charge is not conserved on these shorter-distance scales, but that we need to change our description of the emitters to make it manifest).

10. Equivalence Principle and Lorentz Invariance Recall that the emission of a soft photon of momentum k^μ off a hard scattering amplitude $\mathcal{M}(p_1, \ldots, p_n)$ is given by

$$\mathcal{M}^\mu(k; p_1, \ldots, p_n) \underset{k \to 0}{=} \mathcal{M}(p_1, \ldots, p_n) \sum_{i=1}^{n} \frac{e_i\, p_i^\mu}{p_i \cdot k}, \tag{1.50}$$

where the p_i^μ are the momenta of the hard external particles (with the convention that they are all outgoing) and e_i the electrical charge they carry.

10.a Generalize the formula (1.50) to the emission of a soft graviton in a hard scattering process.

10.b Extend the arguments of Problem 9 to show that an S-matrix element with an external soft graviton is Lorentz invariant up to a phase, provided that soft gravitons couple with the same strength to all the other fields (i.e., provided that long-distance gravitational interactions satisfy the equivalence principle).

10.a Since they have spin 2, gravitons carry two Lorentz indices. Moreover, since they are massless, they have only two physical polarizations. It is possible to choose a gauge in which their "polarization tensors" are obtained from two copies of the photon polarization vectors:

$$\epsilon_\pm^{\mu\nu}(\mathbf{k}) = \epsilon_\pm^\mu(\mathbf{k})\, \epsilon_\pm^\nu(\mathbf{k}). \tag{1.51}$$

The only information needed to generalize (1.50) is to recall that gravitons couple to the energy–momentum tensor. Therefore, instead of a single factor p^μ, the coupling of a soft graviton should contain $p^\mu p^\nu$. The denominators in (1.50) are unchanged since they come

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

from the propagators of the other external lines (fermions, scalars, etc.). Therefore, (1.50) becomes

$$\mathcal{M}^{\mu\nu}(k; p_1, \ldots, p_n) \underset{k\to 0}{=} \mathcal{M}(p_1, \ldots, p_n) \sum_{i=1}^{n} \frac{g_i\, p_i^\mu p_i^\nu}{p_i \cdot k}, \tag{1.52}$$

where g_i is the strength of the coupling of a soft graviton to the particle on the i-th external line.

10.b Thanks to Eq. (1.51) and the results of Problem 9, the Lorentz transform of the polarization tensor of a graviton is given by

$$\left[\Lambda^{-1}\right]^\mu{}_\rho \left[\Lambda^{-1}\right]^\nu{}_\sigma\, \epsilon_\pm^\rho(\Lambda\mathbf{k})\epsilon_\pm^\sigma(\Lambda\mathbf{k}) = e^{\mp 2i\theta}\, \epsilon_\pm^\mu(\mathbf{k})\epsilon_\pm^\nu(\mathbf{k})$$
$$+ \text{ terms in } k^\mu\epsilon_\pm^\nu,\ \epsilon_\pm^\mu k^\nu \text{ or } k^\mu k^\nu.$$

Then, the reasoning of Problem 9 indicates that S-matrix elements with external gravitons are Lorentz invariant (up to a phase), provided that the corresponding amplitude satisfies the following identities:

$$k_\mu\, \mathcal{M}^{\mu\nu}(k, \ldots) = k_\nu\, \mathcal{M}^{\mu\nu}(k, \ldots) = 0.$$

Specializing to the case where the graviton is soft and all other external lines are hard, we can use (1.52), and this becomes

$$\sum_{i=1}^{n} g_i p_i^\nu = 0. \tag{1.53}$$

This identity must be true for any allowed configuration of the momenta p_i of the hard external lines. Obviously, this holds if we have $g_i = \text{const}$ (the identity then follows from energy–momentum conservation). This corresponds to the situation where soft gravitons couple to all other fields with the same strength, i.e., to the *equivalence principle*. Thus, we have shown that the equivalence principle implies the Lorentz invariance of cross-sections involving gravitons.

As in Problem 9, the converse is also true. Indeed, $g = \text{const}$ is the only way to satisfy (1.53) for all p_i. Therefore, *the Lorentz invariance of cross-sections with soft external gravitons implies the equivalence principle for long-distance gravitational interactions.* (Here, "long-distance" should be understood as long compared to the other distance scales of the process, but certainly short compared to the size of the Universe, since translation invariance does not hold on such large scales.) This observation was first made by Weinberg (see Weinberg, S. (1964), *Phys Rev* 135: 1049).

Higher Spins: This argument can be extended in order to exclude the possibility of interacting higher-spin massless particles. For instance, for a spin-3 massless particle, Lorentz invariance would imply

$$\sum_{i=1}^{n} g_i\, p_i^\mu p_i^\nu = 0.$$

One may check that the only way this can be true for all p_i's and all μ, ν is to have $g_i = 0$. (But interacting *massive* spin-3 particles fall out of the scope of this argument, and are in fact allowed.)

11. Lee–Nauenberg Theorem In quantum mechanics and quantum field theory, there exist states that are practically undistinguishable (e.g., states that differ by features that are too dim to be experimentally detected). Transition amplitudes with such initial or final states are in general plagued by singularities. The goal of this problem is to discuss these pathologies in the simple setting of quantum mechanics, and to show that they can be avoided by averaging transition probabilities over all degenerate states. To that end, consider a Hamiltonian $\mathcal{H} = \mathcal{H}_0 + V$ decomposed into a free Hamiltonian $\mathcal{H}_0$ and an interaction potential V. We recall that the scattering matrix S can be written as

$$S_{\beta\alpha} = \delta_{\beta\alpha} - 2\pi i\, \delta(E_\alpha - E_\beta)\, T_{\beta\alpha},$$

$$\text{with } T_{\beta\alpha} \equiv V_{\beta\alpha} + \int d\gamma\, \frac{V_{\beta\gamma}T_{\gamma\alpha}}{E_\alpha - E_\gamma + i0^+}, \qquad V_{\beta\alpha} \equiv \big\langle \Psi_\beta^{(0)} \big| V \big| \Psi_\alpha^{(0)} \big\rangle,$$

where the E_α's and $\big|\Psi_\alpha^{(0)}\big\rangle$'s are the eigenvalues and eigenstates of $\mathcal{H}_0$.

11.a Show that $\big(\Omega_{\rm in}\big)_{\beta\alpha} \equiv \delta_{\beta\alpha} + \frac{T_{\beta\alpha}}{E_\alpha - E_\beta + i0^+}$ is a unitary matrix. *Hint: define* $\big|\Psi_{\alpha,{\rm in}}\big\rangle \equiv \int d\gamma\, \big(\Omega^t_{\rm in}\big)_{\alpha\gamma} \big|\Psi_\gamma^{(0)}\big\rangle$ *and consider the matrix element* $\big\langle \Psi_{\beta,{\rm in}} \big| V \big| \Psi_{\alpha,{\rm in}} \big\rangle$.

11.b Solve formally the equation for $T_{\beta\alpha}$, to obtain

$$T_{\beta\alpha} = V_{\beta\alpha} + \sum_{n=1}^{+\infty} \int d\gamma_1 \cdots d\gamma_n\, \frac{V_{\beta\gamma_n} V_{\gamma_n\gamma_{n-1}} \cdots V_{\gamma_1\alpha}}{(E_\alpha - E_{\gamma_1} + i0^+) \cdots (E_\alpha - E_{\gamma_n} + i0^+)}.$$

11.c Denote by $\mathcal{D}(\alpha)$ the set of states obtained from α by adding/removing ultrasoft particles, by splitting collinearly hard massless particles or by recombining collinear particles, i.e., the set of states that are undistinguishable from the state α. Explain why $T_{\beta\alpha}$ has divergences if $\gamma_1, \gamma_2, \ldots, \gamma_i$ all belong to $\mathcal{D}(\alpha)$ or if $\gamma_i, \gamma_{i+1}, \ldots, \gamma_n$ all belong to $\mathcal{D}(\beta)$.

11.d Introduce P_α, P_β, P_3, the projectors on $\mathcal{D}(\alpha)$, $\mathcal{D}(\beta)$ and on the rest of the Hilbert space, respectively. Show that squared matrix elements of T become finite if they are summed over all the states in $\mathcal{D}(\alpha)$ and $\mathcal{D}(\beta)$, i.e., $\sum_{a\in\mathcal{D}(\alpha), b\in\mathcal{D}(\beta)} \big|T_{ba}\big|^2 < \infty$.

Reminders on Scattering Theory in Quantum Mechanics: The setting of this study is the formulation of scattering in quantum mechanics. Let us consider a Hamiltonian $\mathcal{H}$, which we decompose into a free Hamiltonian and an interaction potential:

$$\mathcal{H} = \mathcal{H}_0 + V.$$

Denote by $\big|\Psi_\alpha^{(0)}\big\rangle$ the eigenstates of the free Hamiltonian,

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

$$\mathcal{H}_0 \left|\Psi_\alpha^{(0)}\right\rangle = E_\alpha \left|\Psi_\alpha^{(0)}\right\rangle,$$

and by $\left|\Psi_\alpha\right\rangle$ those of the full Hamiltonian,

$$\mathcal{H} \left|\Psi_\alpha\right\rangle = E_\alpha \left|\Psi_\alpha\right\rangle.$$

The eigenvalues of the free and interacting Hamiltonians are the same, provided we disregard the formation of bound states that may happen with certain interactions, and provided the free Hamiltonian is expressed in terms of the physical masses of the particles. The last equation may be rewritten as

$$(E_\alpha - \mathcal{H}_0) \left|\Psi_\alpha\right\rangle = V \left|\Psi_\alpha\right\rangle. \tag{1.54}$$

Note that the operator $E_\alpha - \mathcal{H}_0$ is not invertible, since $\left|\Psi_\alpha^{(0)}\right\rangle$ lies in its kernel (as well as any other state degenerate with it). Consider the following states:

$$\begin{aligned}\left|\Psi_{\alpha,\text{in}}\right\rangle &\equiv \left|\Psi_\alpha^{(0)}\right\rangle + \left(E_\alpha - \mathcal{H}_0 + i0^+\right)^{-1} V \left|\Psi_{\alpha,\text{in}}\right\rangle,\\ \left|\Psi_{\alpha,\text{out}}\right\rangle &\equiv \left|\Psi_\alpha^{(0)}\right\rangle + \left(E_\alpha - \mathcal{H}_0 - i0^+\right)^{-1} V \left|\Psi_{\alpha,\text{out}}\right\rangle.\end{aligned}$$

These states are obviously solutions of (1.54). These states are time-independent (i.e., they are states in the Heisenberg representation), but may look different to observers in different frames. In particular, in a frame that differs by a translation by a time T, these states are multiplied by $e^{-i\mathcal{H}T}$. Since they are energy eigenstates, the action of this operator is just a multiplication by an irrelevant phase $\exp(-iE_\alpha T)$. But these phases have a non-trivial effect on superpositions of states. Consider, for instance,

$$\left|\Psi_{h,\text{in}}\right\rangle \equiv \int d\alpha\, h(\alpha) \left|\Psi_{\alpha,\text{in}}\right\rangle$$

(the notation $d\alpha$ is a shorthand for the integration/summation measure over all the quantum numbers of the particles in the state α, such as 3-momenta, spins, etc.), where $h(\alpha)$ is a smooth function that spans a small range of energies E_α. Under such a translation in time, this state becomes

$$\begin{aligned}\left|\Psi_{h,\text{in}}\right\rangle \to \left|\Psi_{h,\text{in}}\right\rangle_T &= \int d\alpha\, h(\alpha)\, e^{-iE_\alpha T} \left|\Psi_{\alpha,\text{in}}\right\rangle\\ &= e^{-i\mathcal{H}_0 T} \left|\Psi_h^{(0)}\right\rangle + \int d\alpha\, h(\alpha)\, e^{-iE_\alpha T} \left(E_\alpha - \mathcal{H}_0 + i0^+\right)^{-1} V \left|\Psi_{\alpha,\text{in}}\right\rangle.\end{aligned} \tag{1.55}$$

In the second term, we must close the contour for the integral over the energy E_α in the upper half-plane in order to ensure its convergence if $T < 0$. Then, the theorem of residues gives the integral in terms of the singularities of the integrand. Except for the denominator, the other factors generically have their singularities at finite imaginary part, $\text{Im}\,(E_\alpha) > 0$ (because interactions typically lead to a finite lifetime for the particles), and their contribution vanishes when $T \to -\infty$. The denominator leads to poles located at $E_\alpha = E_* - i0^+$, i.e., below the real axis, and therefore it does not contribute to the result. Thus, we have

$$e^{-i\mathcal{H}T} \left|\Psi_{h,\text{in}}\right\rangle \underset{T\to-\infty}{=} e^{-i\mathcal{H}_0 T} \left|\Psi_h^{(0)}\right\rangle, \quad e^{-i\mathcal{H}T} \left|\Psi_{h,\text{out}}\right\rangle \underset{T\to+\infty}{=} e^{-i\mathcal{H}_0 T} \left|\Psi_h^{(0)}\right\rangle.$$

Note that, since the operators $\exp(-i\mathcal{H}T)$ and $\exp(-i\mathcal{H}_0 T)$ are unitary, the states $\left|\Psi_{h,\text{in}}\right\rangle$ and $\left|\Psi_{h,\text{out}}\right\rangle$ are normalized like the states $\left|\Psi_h^{(0)}\right\rangle$. Since this is true for any function $h(\alpha)$, the

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

states $|\Psi_{\alpha,\mathrm{in}}\rangle$ and $|\Psi_{\alpha,\mathrm{out}}\rangle$ each form an orthonormal set of states. In other words, the *in* and *out* states become identical to the eigenstates of the free Hamiltonian in frames translated by a time $T \to -\infty$ or $T \to +\infty$, respectively. Thus, the physical interpretation of the (time-independent) state $|\Psi_{\alpha,\mathrm{in}}\rangle$ is that it encodes all the possible outcomes of an interacting system initialized at $T = -\infty$ in the free state $|\Psi^{(0)}_\alpha\rangle$. Likewise, $|\Psi_{\alpha,\mathrm{out}}\rangle$ encodes all the possible histories that lead to the free state $|\Psi^{(0)}_\alpha\rangle$ at $T = +\infty$. Therefore, the transition amplitude from a state α to a state β should be defined as

$$S_{\beta\alpha} \equiv \langle\Psi_{\beta,\mathrm{out}}|\Psi_{\alpha,\mathrm{in}}\rangle.$$

Note that the definition of the state $|\Psi_{\alpha,\mathrm{in}}\rangle$ can be rewritten as

$$|\Psi_{\alpha,\mathrm{in}}\rangle = \int d\beta\, \underbrace{\left(\delta_{\beta\alpha} + \frac{T_{\beta\alpha}}{E_\alpha - E_\beta + i0^+}\right)}_{\equiv\, \Omega_{\beta\alpha}} |\Psi^{(0)}_\beta\rangle, \quad \text{with } T_{\beta\alpha} \equiv \langle\Psi^{(0)}_\beta|V|\Psi_{\alpha,\mathrm{in}}\rangle$$

($\delta_{\beta\alpha}$ is the combination of delta functions and Kronecker symbols such that $\int d\beta\, f(\beta)\delta_{\beta\alpha} = f(\alpha)$). Inserting the definition of $|\Psi_{\alpha,\mathrm{in}}\rangle$ in $T_{\beta\alpha}$, we obtain an integral equation for T:

$$T_{\beta\alpha} = V_{\beta\alpha} + \int d\gamma\, \frac{V_{\beta\gamma} T_{\gamma\alpha}}{E_\alpha - E_\gamma + i0^+} = \int d\gamma\, V_{\beta\gamma}\, \Omega_{\gamma\alpha}, \tag{1.56}$$

with $V_{\beta\alpha} \equiv \langle\Psi^{(0)}_\beta|V|\Psi^{(0)}_\alpha\rangle$. This equation, known as the *Lipmann–Schwinger equation*, shows how the matrix elements of T are obtained from the matrix elements of the interaction potential in the free basis.

11.a Given the above discussion, we have

$$|\Psi_{\alpha,\mathrm{in}}\rangle \underset{T\to-\infty}{=} e^{i\mathcal{H}T}\, e^{-i\mathcal{H}_0 T}\, |\Psi^{(0)}_\alpha\rangle,$$

which implies that $\Omega_{\beta\alpha} = \lim_{T\to-\infty} \langle\Psi^{(0)}_\beta|e^{i\mathcal{H}T}\, e^{-i\mathcal{H}_0 T}|\Psi^{(0)}_\alpha\rangle$. In this form, we see immediately that the $\Omega_{\beta\alpha}$'s are the matrix elements of a unitary operator.

Alternatively, we can reach the same conclusion using only the information provided in the statement of the problem. For this, it is convenient to write

$$\Omega \equiv 1 + \Delta, \quad T_{\beta\alpha} = (E_\alpha - E_\beta + i0^+)\,\Delta_{\beta\alpha},$$
$$|\Psi_{\alpha,\mathrm{in}}\rangle = |\Psi^{(0)}_\alpha\rangle + \int d\gamma\, \Delta^t_{\alpha\gamma}\, |\Psi^{(0)}_\gamma\rangle \quad (\Delta^t_{\alpha\gamma} \equiv \Delta_{\gamma\alpha}).$$

Given the integral equation that relates $V_{\beta\alpha}$ and $T_{\beta\alpha}$ and the definition of $|\Psi_{\alpha,\mathrm{in}}\rangle$, we have also

$$T_{\beta\alpha} = \langle\Psi^{(0)}_\beta|V|\Psi_{\alpha,\mathrm{in}}\rangle.$$

Then, the matrix element $\langle\Psi_{\beta,\mathrm{in}}|V|\Psi_{\alpha,\mathrm{in}}\rangle$ can be written in two ways, depending on whether

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

we expand the state on the left or the state on the right:

$$\begin{aligned}\big\langle\Psi_{\beta,\mathrm{in}}\big|V\big|\Psi_{\alpha,\mathrm{in}}\big\rangle &= \underbrace{\big\langle\Psi_\beta^{(0)}\big|V\big|\Psi_{\alpha,\mathrm{in}}\big\rangle}_{T_{\beta\alpha}} + \int d\gamma\, \Delta^\dagger_{\beta\gamma}\, \underbrace{\big\langle\Psi_\gamma^{(0)}\big|V\big|\Psi_{\alpha,\mathrm{in}}\big\rangle}_{T_{\gamma\alpha}} \\ &= \underbrace{\big\langle\Psi_{\beta,\mathrm{in}}\big|V\big|\Psi_\alpha^{(0)}\big\rangle}_{T^*_{\alpha\beta}} + \int d\gamma\, \underbrace{\big\langle\Psi_{\beta,\mathrm{in}}\big|V\big|\Psi_\gamma^{(0)}\big\rangle}_{T^*_{\gamma\beta}=T^\dagger_{\beta\gamma}}\, \Delta_{\gamma\alpha}.\end{aligned}$$

Equating these two expressions for the matrix element leads to

$$\begin{aligned}T^*_{\alpha\beta} - T_{\beta\alpha} &= \int d\gamma\, \Big(\Delta^\dagger_{\beta\gamma} T_{\gamma\alpha} - T^\dagger_{\beta\gamma}\Delta_{\gamma\alpha}\Big) \\ &= (E_\alpha - E_\beta + i0^+) \int d\gamma\, \Delta^\dagger_{\beta\gamma}\Delta_{\gamma\alpha},\end{aligned}$$

i.e., $\Delta + \Delta^\dagger + \Delta^\dagger\Delta = 0$. From this relationship, we conclude that $\Omega = 1 + \Delta$ is unitary.

S-matrix: The matrix element $S_{\beta\alpha}$ can also be expressed in terms of $T_{\beta\alpha}$. In order to obtain this relationship, let us start again from (1.55), but this time we take the limit $T \to +\infty$. Since $T > 0$, we must close the integration contour in the lower half-plane. In the limit $T \to +\infty$, the only term that survives when we apply the theorem of residues comes from the zero in the denominator, and we get

$$\begin{aligned}\big|\Psi_{h,\mathrm{in}}\big\rangle_T &\underset{T\to+\infty}{=} e^{-i\mathcal{H}_0 T}\big|\Psi_h^{(0)}\big\rangle - 2\pi i \int d\alpha d\beta\; h(\alpha)\, e^{-iE_\beta T}\,\delta(E_\alpha - E_\beta)\, T_{\beta\alpha}\big|\Psi_\beta^{(0)}\big\rangle \\ &\underset{T\to+\infty}{=} \int d\alpha d\beta\; h(\alpha)\, e^{-iE_\beta T}\Big(\delta_{\beta\alpha} - 2\pi i\,\delta(E_\alpha - E_\beta)\, T_{\beta\alpha}\Big)\big|\Psi_\beta^{(0)}\big\rangle.\end{aligned}$$

Alternatively, we can also write

$$\begin{aligned}\big|\Psi_{h,\mathrm{in}}\big\rangle_T &= \int d\alpha\; h(\alpha)\, e^{-iE_\alpha T}\big|\Psi_{\alpha,\mathrm{in}}\big\rangle \\ &= \int d\alpha\; h(\alpha)\, e^{-iE_\alpha T}\int d\beta\; \big|\Psi_{\beta,\mathrm{out}}\big\rangle \underbrace{\big\langle\Psi_{\beta,\mathrm{out}}\big|\Psi_{\alpha,\mathrm{in}}\big\rangle}_{S_{\beta\alpha}} \\ &\underset{T\to+\infty}{=} \int d\alpha d\beta\; h(\alpha)\, e^{-iE_\beta T}\, S_{\beta\alpha}\big|\Psi_\beta^{(0)}\big\rangle.\end{aligned}$$

(In the last line, we also used the fact that $S_{\beta\alpha}$ is proportional to a delta function of energy conservation to change E_α into E_β in the exponential.) By comparing the two formulas, we conclude that

$$S_{\beta\alpha} = \delta_{\beta\alpha} - 2\pi i\,\delta(E_\alpha - E_\beta)\, T_{\beta\alpha}.$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

11.b The Lipmann–Schwinger equation (1.56) can be solved iteratively, leading to the following series in powers of the interaction potential:

$$T_{\beta\alpha} = V_{\beta\alpha} + \sum_{n=1}^{+\infty} \int d\gamma_1 \cdots d\gamma_n \, \frac{V_{\beta\gamma_n} V_{\gamma_n\gamma_{n-1}} \cdots V_{\gamma_1\alpha}}{(E_\alpha - E_{\gamma_1} + i0^+) \cdots (E_\alpha - E_{\gamma_n} + i0^+)}.$$

The physical interpretation of this expression is that the system starts in the state α and evolves to the state β, undergoing a sequence of interactions that take it through the intermediate states $\gamma_1, \gamma_2, \ldots, \gamma_n$.

11.c In this formula, the matrix elements of V are regular, and the only potential singularities come from the poles in the other factors under the integral. However, these poles do not always lead to a divergence in the result. This can be seen, for instance, in the following two toy examples:

$$\int_{-1}^{+1} \frac{dx}{x + i0^+} = -i\pi, \qquad \int_{0}^{+1} \frac{dx}{x + i0^+} = \infty.$$

In these examples, we see that in order to lead to a divergence, the pole should be at one of the endpoints of the integration domain; otherwise we may deform the integration contour to completely avoid the pole. Another configuration leading to an unavoidable divergence is when the integrand has a pair of poles "pinching" the integration contour so that it cannot be deformed.

Two states α and γ may accidentally have the same energy when one varies the momentum of a particle. However, when varying the state γ, these accidental degeneracies occur in the middle of the allowed domain of the variables that parameterize γ. The corresponding pole in $(E_\alpha - E_\gamma + i0^+)^{-1}$ is therefore avoidable by deforming the integration contour, and this does not lead to a divergence.

The state γ also has the same energy as α if it differs from α by one or more ultrasoft particles. In this situation, the resulting pole in $(E_\alpha - E_\gamma + i0^+)^{-1}$ occurs when the energy of these extra particles goes to zero, which is on the boundary of the integration domain for γ, and therefore produces an actual divergence in the result. Another situation where γ has the same energy as α is when a hard massless particle in one of the two states is replaced in the other state by two or more collinear particles with the same total momentum. This case also leads to a divergence, because the pole is reached on the boundary of the angular integration domain.

In order to keep track of these two problematic situations, let us denote by $\mathcal{D}(\alpha)$ the set of states that are degenerate with α, excluding the accidental degeneracies. Consider now the sequence of states

$$\alpha \to \gamma_1 \to \cdots \to \gamma_{i-1} \to \gamma_i \to \gamma_{i+1} \to \cdots \to \gamma_n \to \beta.$$

Each state γ_i in this sequence is related to the preceding and following states by an elementary interaction V. Obviously, in order to produce a pole, the state γ_i must have the same energy as α. But for this pole to lead to a divergence, a stronger condition should be satisfied: the pole must occur on the boundary of the integration domain of γ_i, which is made of the states in $\mathcal{D}(\gamma_{i-1})$ or in $\mathcal{D}(\gamma_{i+1})$.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

Assume for instance that $\gamma_i \in \mathcal{D}(\gamma_{i-1})$, and has the same energy as α. But this implies that γ_{i-1} also has the same energy as α, which happens on a harmless zero measure subset of the phase-space of γ_{i-1} unless γ_{i-1} itself belongs to the degeneracy set of γ_{i-2}, $\mathcal{D}(\gamma_{i-2})$. By repeating this inductive argument, we see that this divergence occurs only if all of $\gamma_1, \gamma_2, \dots, \gamma_i$ belong to $\mathcal{D}(\alpha)$. If instead we make the assumption that $\gamma_i \in \mathcal{D}(\gamma_{i+1})$, then it is the entire chain $\gamma_i, \gamma_{i+1}, \dots, \gamma_n$ that must be in $\mathcal{D}(\beta)$ in order to have a divergence. This result is consistent with the well-known fact that infrared divergences happen when soft photons are attached to the external legs of a hard process, but not when one attaches the soft photon to a hard internal line of a graph. The same is true of the divergences due to collinear splittings.

11.d We assume now that the initial and final states are not degenerate, i.e., that $\mathcal{D}(\alpha)$ and $\mathcal{D}(\beta)$ do not overlap. Then, we introduce projectors P_α, P_β and P_3 that project respectively on $\mathcal{D}(\alpha)$, on $\mathcal{D}(\beta)$ and on the remaining portion of the Hilbert space. Since the union of these sets is the complete Hilbert space, the sum of these projectors is the identity

$$P_\alpha + P_\beta + P_3 = 1.$$

Moreover, we assume that the volume of $\mathcal{D}(\alpha)$ and $\mathcal{D}(\beta)$ is very small (this is controlled by what is experimentally meant by “soft” or “collinear”) compared to the rest, implying that we may neglect these states except when they produce a divergence. In terms of these projectors, the matrix element $T_{\beta\alpha}$ can be written as

$$T_{\beta\alpha} = \sum_{n=0}^{+\infty} \left[V \left(\frac{P_\alpha + P_\beta + P_3}{E_\alpha - \mathcal{H}_0 + i0^+} V \right)^n \right]_{\beta\alpha}.$$

This formula is still exact. The next step is to expand this expression, and to drop P_α and P_β unless they occur in configurations that produce a divergence. Thus we keep the P_α’s if they form an uninterrupted chain on the right of the product, and the P_β’s if they form a chain on the left of the product:

$$T_{\beta\alpha} \underset{\text{sing.}}{=} \sum_{r,s,n=0}^{+\infty} \left[\left(V \frac{P_\beta}{E_\alpha - \mathcal{H}_0 + i0^+} \right)^s V \left(\frac{P_3}{E_\alpha - \mathcal{H}_0 + i0^+} V \right)^n \left(\frac{P_\alpha}{E_\alpha - \mathcal{H}_0 + i0^+} V \right)^r \right]_{\beta\alpha},$$

which can be written compactly as

$$T_{\beta\alpha} = \left(U^\dagger_{\beta,\text{out}}\, T_3\, U_{\alpha,\text{in}} \right)_{\beta\alpha}, \quad \text{with} \quad T_3 \equiv V \sum_{n=0}^{+\infty} \left(\frac{P_3}{E_\alpha - \mathcal{H}_0 + i0^+} V \right)^n,$$

and

$$U_{\alpha,\text{in}} \equiv \sum_{r=0}^{+\infty} \left(\frac{P_\alpha}{E_\alpha - \mathcal{H}_0 + i0^+} V \right)^r, \quad U_{\beta,\text{out}} \equiv \sum_{s=0}^{+\infty} \left(\frac{P_\beta}{E_\alpha - \mathcal{H}_0 - i0^+} V \right)^s.$$

The factors $U_{\alpha,\text{in}}$ and $U_{\beta,\text{out}}$ contain divergences, while the factor T_3 is divergence-free.

Note now that the operator $U_{\alpha,\text{in}}$ is the restriction to the subspace $\mathcal{D}(\alpha)$ of the operator Ω that we introduced just above (1.56), and that we have shown to be unitary. Therefore, $U_{\alpha,\text{in}}$ is

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

also unitary if restricted to $\mathcal{D}(\alpha)$, which means that

$$U_{\alpha,\mathrm{in}}\, P_\alpha\, U^\dagger_{\alpha,\mathrm{in}} = P_\alpha, \quad U_{\beta,\mathrm{out}}\, P_\beta\, U^\dagger_{\beta,\mathrm{out}} = P_\beta.$$

(We have also written an analogous relationship for $U_{\beta,\mathrm{out}}$.) Then, summing the squared matrix elements over all the states degenerate with α and β, we obtain

$$\begin{aligned}\sum_{a\in\mathcal{D}(\alpha),b\in\mathcal{D}(\beta)} \left|T_{ba}\right|^2 &= \sum_{a\in\mathcal{D}(\alpha),b\in\mathcal{D}(\beta)} \left(U^\dagger_{\beta,\mathrm{out}}\, T_3\, U_{\alpha,\mathrm{in}}\right)_{ba} \left(U^\dagger_{\alpha,\mathrm{in}}\, T_3^\dagger\, U_{\beta,\mathrm{out}}\right)_{ab} \\ &= \mathrm{tr}\left(U_{\beta,\mathrm{out}}\, P_\beta\, U^\dagger_{\beta,\mathrm{out}}\, T_3\, U_{\alpha,\mathrm{in}}\, P_\alpha\, U^\dagger_{\alpha,\mathrm{in}}\, T_3^\dagger\right) \\ &= \mathrm{tr}\left(P_\beta\, T_3\, P_\alpha\, T_3^\dagger\right) = \sum_{a\in\mathcal{D}(\alpha),b\in\mathcal{D}(\beta)} \left|(T_3)_{ba}\right|^2,\end{aligned}$$

which is a finite quantity since it contains only the finite operator T_3. In the second line, we have rewritten the sum over the states in $\mathcal{D}(\alpha)$ and $\mathcal{D}(\beta)$ in the form of a trace by introducing the appropriate projectors, and we have used the cyclic invariance of the trace to reorder the factors.

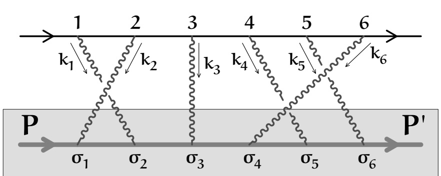
This result, established by Lee and Nauenberg, shows that finite cross-sections are obtained provided one sums over all the initial and final degenerate states. This is rather natural in quantum mechanics since all these states are undistinguishable in practice, given the unavoidable limited resolutions (in energy and in angle) of any detector.

12. Classical External Field Approximation The goal of this problem is to study the scattering between an electron of charge $-e$ and a large atomic nucleus of charge Ze, and to show that, in the limit of a recoilless nucleus at rest, the nucleus can be approximated by its classical Coulomb potential.

12.a Why are multiple scatterings important at large Z?

12.b We are interested in the limit where the momentum exchanges k_i are soft. Why is this the relevant limit for a scattering at large impact parameter? How can we simplify the treatment of the nucleus in this limit?

12.c Express the lower part $\mathcal{A}^{\mu_{\sigma_1}\cdots\mu_{\sigma_n}}_{ss'}(P, P'; \{k_{\sigma_i}\})$ of the graph shown below (surrounded by a box) in this limit, for an arbitrary number of exchanged photons.



Perform the sum explicitly over all the permutations σ of the attachments of the photons to the nucleus, in order to obtain

$$\sum_{\sigma\in\mathfrak{S}_n} \mathcal{A}^{\mu_{\sigma_1}\cdots\mu_{\sigma_n}}_{ss'}(P, P'; \{k_{\sigma_i}\})$$

$$\underset{k_i\ll P}{\approx} 2\,E_{\mathbf{P}}\,\delta_{ss'}\,(2\pi)^3\delta\big(\mathbf{P} + \sum \mathbf{k}_i - \mathbf{P}'\big) \prod_i 2\pi\,(-iZe)\,P^{\mu_i}\,\delta(P\cdot k_i).$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

12.d Show that when the nucleus is at rest, the scattering amplitude of an electron off the nucleus can be obtained by replacing the nucleus by its classical Coulomb potential.

12.a Each photon exchanged between the electron and the nucleus can potentially bring a factor Ze^2. Despite the smallness of the fine-structure constant ($\alpha \sim 137^{-1}$), a large Z can lead to significant multi-photon effects. Here, we must mention an important limitation to this power counting: the coupling of the photon to the nucleus gives a factor Ze only if the photon sees coherently the entire electrical charge of the nucleus. In other words, the wavelength of the photon should be large enough so that it cannot resolve the more elementary charged constituents contained inside the nucleus (protons, or quarks at an even smaller distance scale). Given that the size of a nucleus is of the order of 10 fm, the momentum of the exchanged photon should not exceed 20 MeV in order to benefit from this enhancement.

12.b If we denote by x and y the spacetime coordinates at the endpoints of a photon propagator, the momentum k carried by this photon is Fourier conjugate to the difference $x-y$. In the case of the photons exchanged between the electron and the nucleus propagators, this difference is related to the impact parameter of the scattering. Thus, a small momentum carried by the exchanged photons is equivalent to a large impact parameter.

In fact, the criterion that the photon momentum is small enough for the photon to couple coherently with the entire nucleus is equivalent to having an impact parameter larger than the size of the nucleus. In other words, the electron should pass outside of the nucleus. In this limit, the photon sees only a point-like charge Ze instead of a complicated arrangement of quarks. Moreover, in this limit, the photon momenta ($\lesssim 20$ MeV) are much smaller than the mass of the nucleus, implying that the nucleus suffers a negligible recoil in the scattering.

12.c Using the QED Feynman rules, and the approximation of a point-like nucleus, the lower part of the graph shown in the figure can be written as

$$\begin{aligned}\mathcal{A}_{ss'}^{\mu_{\sigma_1}\cdots\mu_{\sigma_n}}(P,P';\{k_{\sigma_i}\}) \equiv{}& (2\pi)^4\delta\big(P+\sum_{i=1}^{n}k_i-P'\big)\,\overline{u}_{s'}(\mathbf{P}')\\ &\times\Big[(-iZe\gamma^{\mu_{\sigma_n}})\cdots\frac{i(\not{P}+\not{k}_{\sigma_1}+\not{k}_{\sigma_2}+M)}{(P+k_{\sigma_1}+k_{\sigma_2})^2-M^2+i0^+}\\ &\times(-iZe\gamma^{\mu_{\sigma_2}})\frac{i(\not{P}+\not{k}_{\sigma_1}+M)}{(P+k_{\sigma_1})^2-M^2+i0^+}(-iZe\gamma^{\mu_{\sigma_1}})\Big]\,u_s(\mathbf{P}).\end{aligned}$$

Note a limitation of our treatment here, since the above formula assumes that the target has spin 1/2 (although our final result will in fact be valid for arbitrary spin). Using the fact that $P^2 = M^2$ and the fact that the exchanged photons are soft, the intermediate propagators can be approximated as

$$\frac{i(\not{P}+\not{k}_{\sigma_1}+\cdots+\not{k}_{\sigma_i}+M)}{(P+k_{\sigma_1}+\cdots+k_{\sigma_i})^2-M^2+i0^+}\approx\frac{i\sum_{s_i=\pm}u_{s_i}(\mathbf{P})\overline{u}_{s_i}(\mathbf{P})}{2P\cdot(k_{\sigma_1}+\cdots+k_{\sigma_i})+i0^+}.$$

Note also that $\overline{u}_{s'}(\mathbf{P}') \approx \overline{u}_{s'}(\mathbf{P})$, $\overline{u}_r(\mathbf{P})\gamma^\mu u_s(\mathbf{P}) = 2\,\delta_{rs}P^\mu$ and

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

$$P'^0 - P^0 - \sum_i k_i^0 \approx \sqrt{M^2 + \mathbf{P}^2 + 2\,\mathbf{P}\cdot\sum_i \mathbf{k}_i} - E_{\mathbf{P}} - \sum_i k_i^0 \approx -\frac{P\cdot\sum_i k_i}{E_{\mathbf{P}}}.$$

Therefore, we have

$$\begin{aligned}\mathcal{A}_{ss'}^{\mu_{\sigma_1}\cdots\mu_{\sigma_n}}(P,P';\{k_{\sigma_i}\}) \underset{k_i\ll P}{=} &\; 2\,E_{\mathbf{P}}\,\delta_{ss'}\,(-iZe)^n\,P^{\mu_1}P^{\mu_2}\cdots P^{\mu_n}\\ &\times (2\pi)^4\delta\big(\mathbf{P}+\sum \mathbf{k}_i-\mathbf{P}'\big)\,\delta\big(P\cdot\sum k_i\big)\\ &\times \frac{i}{P\cdot k_{\sigma_1}+i0^+}\,\frac{i}{P\cdot(k_{\sigma_1}+k_{\sigma_2})+i0^+}\cdots\\ &\times \frac{i}{P\cdot(k_{\sigma_1}+\cdots+k_{\sigma_{n-1}})+i0^+}.\end{aligned}$$

The next step is to perform the sum over all the permutations $\sigma\in\mathfrak{S}_n$. One may get a hint about the answer by considering the case $n=2$, where there are only two permutations:

$$\begin{aligned}&2\pi\delta\big(P\cdot(k_1+k_2)\big)\left(\frac{i}{P\cdot k_1+i0^+}+\frac{i}{P\cdot k_2+i0^+}\right)\\ &\quad= 2\pi\delta\big(P\cdot(k_1+k_2)\big)\left(\frac{i}{P\cdot k_1+i0^+}+\frac{i}{-P\cdot k_1+i0^+}\right)\\ &\quad= 2\pi\delta\big(P\cdot(k_1+k_2)\big)\,2\pi\delta\big(P\cdot k_1\big) = 2\pi\delta\big(P\cdot k_1\big)\,2\pi\delta\big(P\cdot k_2\big).\end{aligned}$$

This can be generalized to any n, by starting from the trivial identity

$$\sum_{\sigma\in\mathfrak{S}_n}\theta(t_{\sigma_1}-t_{\sigma_2})\,\theta(t_{\sigma_2}-t_{\sigma_3})\cdots\theta(t_{\sigma_{n-1}}-t_{\sigma_n}) = 1.$$

Then, we multiply it by $\prod_{i=1}^n \exp\big(i(P\cdot k_i)t_i\big)$ and integrate over all the t_i's. This leads immediately to

$$2\pi\,\delta\big(P\cdot\sum k_i\big)\sum_{\sigma\in\mathfrak{S}_n}\frac{i}{P\cdot k_{\sigma_1}+i0^+}\cdots\frac{i}{P\cdot(k_{\sigma_1}+\cdots+k_{\sigma_{n-1}})+i0^+} = \prod_{i=1}^n 2\pi\,\delta(P\cdot k_i).$$

Therefore, we obtain

$$\begin{aligned}&\sum_{\sigma\in\mathfrak{S}_n}\mathcal{A}_{ss'}^{\mu_{\sigma_1}\cdots\mu_{\sigma_n}}(P,P';\{k_{\sigma_i}\})\\ &\qquad\underset{k_i\ll P}{\approx} 2\,E_{\mathbf{P}}\,\delta_{ss'}\,(2\pi)^3\delta\big(\mathbf{P}+\sum_i \mathbf{k}_i-\mathbf{P}'\big)\prod_i 2\pi\,(-iZe)\,P^{\mu_i}\,\delta(P\cdot k_i),\end{aligned} \tag{1.57}$$

where the only remaining entanglement among the photon momenta is the delta function of momentum conservation.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

12.d In order to go further with the interpretation of this formula, one should recall that with our conventions a state $\big|\mathbf{P}, s\big\rangle$ represents a beam of particles of momentum $\mathbf{P}$ and spin s, with a uniform density of $2E_\mathbf{P}$ particles per unit volume, rather than a single particle. (The delta function of momentum conservation in (1.57) is consistent with this: indeed, momentum conservation is equivalent to invariance under spatial translations, which is only possible for states that are themselves translation invariant.) It is easy to construct a normalizable state that represents a single particle, by a linear superposition such as

$$\big|\Psi\big\rangle \equiv \sum_s \int \frac{d^3\mathbf{P}}{(2\pi)^3}\,\frac{e^{i\mathbf{P}\cdot\mathbf{X}_0}}{\sqrt{2E_\mathbf{p}}}\,\Psi_s(\mathbf{P})\,\big|\mathbf{P}, s\big\rangle.$$

Indeed, one may check that

$$\big\langle\Psi\big|\Psi\big\rangle = 1 \quad \text{provided that} \quad \sum_s \int \frac{d^3\mathbf{P}}{(2\pi)^3}\,\big|\Psi_s(\mathbf{P})\big|^2 = 1.$$

In the definition of the state $\big|\Psi\big\rangle$, the coordinate $\mathbf{X}_0$ may be viewed as the "center" of the object represented by this state. (For instance, if the momentum-space wavefunction $\Psi_s(\mathbf{P})$ is a Gaussian, the wavefunction of the state in the coordinate representation is a Gaussian centered at X_0.) The analogue of (1.57) for the state $\big|\Psi\big\rangle$ is

$$\begin{aligned}
&\mathcal{A}^{\mu_1\cdots\mu_n}(\Psi;\{k_i\}) \\
&\quad\equiv \sum_{s,s'} \int \frac{d^3\mathbf{P}d^3\mathbf{P}'}{(2\pi)^6}\,\frac{e^{i(\mathbf{P}-\mathbf{P}')\cdot\mathbf{X}_0}}{\sqrt{4E_\mathbf{P}E_{\mathbf{P}'}}}\,\Psi_\mathbf{s}(\mathbf{P})\Psi^*_{s'}(\mathbf{P}') \sum_{\sigma\in\mathfrak{S}_n} \mathcal{A}_{ss'}^{\mu_{\sigma_1}\cdots\mu_{\sigma_n}}(P,P';\{k_{\sigma_i}\}) \\
&\quad\underset{k_i\ll P}{\approx} \sum_s \int \frac{d^3\mathbf{P}}{(2\pi)^3}\,\Psi_\mathbf{s}(\mathbf{P})\Psi^*_{s'}(\mathbf{P}) \prod_i 2\pi\,(-iZe)\,P^{\mu_i}\,e^{-ik_i\cdot X_0}\,\delta(P\cdot k_i),
\end{aligned}$$

where we have used $\mathbf{P} \approx \mathbf{P}'$ whenever possible. Let us now attach the n photon propagators to the Lorentz indices of this expression:

$$\begin{aligned}
\mathcal{M}^{\nu_1\cdots\nu_n}(\Psi;\{k_i\}) &\equiv \mathcal{A}^{\mu_1\cdots\mu_n}(\Psi;\{k_i\}) \prod_{i=1}^{n} \frac{-i\,g_{\mu_i}{}^{\nu_i}}{k_i^2 + i0^+} \\
&\underset{k_i\ll P}{=} \sum_s \int \frac{d^3\mathbf{P}}{(2\pi)^3}\,\Psi_\mathbf{s}(\mathbf{P})\Psi^*_{s'}(\mathbf{P}) \prod_i 2\pi\,\delta(P\cdot k_i)\,\frac{-Ze\,P^{\nu_i}\,e^{-ik_i\cdot X_0}}{k_i^2 + i0^+}.
\end{aligned}$$

(Note that the result does not depend on the gauge used to express the photon propagators, because they are contracted into P^{μ_i} and we have $P\cdot k_i = 0$.) Assume now that the nucleus described by the state $\big|\Psi\big\rangle$ is at rest. In other words, the wavefunction $\Psi_s(\mathbf{P})$ is a narrow peak centered at $\mathbf{P} = 0$, with a support such that the typical $\mathbf{P}$ is much smaller than the mass M. In

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

this static limit, we have

$$2\pi\,\delta(P\cdot k_i)\,P^{\nu_i} \approx 2\pi\,\delta(Mk_i^0)\,M\,\delta^{\nu_i}{}_0 = 2\pi\,\delta(k_i^0)\,\delta^{\nu_i}{}_0,$$

and

$$\mathcal{M}^{\nu_1\cdots\nu_n}(\Psi;\{k_i\}) \underset{\substack{k_i\ll P\\ \text{static}}}{=} \underbrace{\sum_s\int\frac{d^3\mathbf{P}}{(2\pi)^3}\,\Psi_s(\mathbf{P})\Psi^*_{s'}(\mathbf{P})}_{=1}\;\prod_i 2\pi\,\delta(k_i^0)\,\frac{-Ze\,\delta^{\nu_i}{}_0\,e^{-ik_i\cdot X_0}}{k_i^2+i0^+}$$

$$=\prod_i 2\pi\,\delta(k_i^0)\,\frac{Ze\,\delta^{\nu_i}{}_0\,e^{-ik_i\cdot X_0}}{\mathbf{k}_i^2-i0^+}.$$

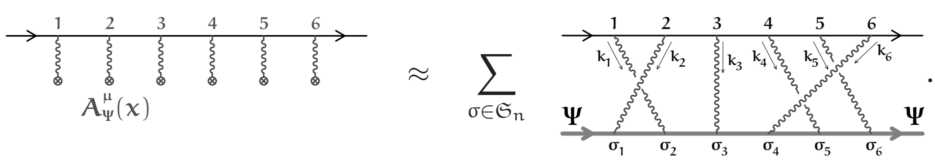
The interpretation of this expression is more transparent in coordinate space, after a Fourier transformation:

$$\mathcal{M}^{\nu_1\cdots\nu_n}(\Psi;\{x_i\}) \equiv \int\prod_{i=1}^n\frac{d^4k_i}{(2\pi)^4}\,e^{-ik_i\cdot x_i}\,\mathcal{M}^{\nu_1\cdots\nu_n}(\Psi;\{k_i\}) \approx \prod_i \underbrace{\delta^{\nu_i}{}_0\,\frac{Ze}{4\pi\,|\mathbf{x}_i-\mathbf{X}_0|}}_{\equiv A_\Psi^{\nu_i}(x_i)}.$$

In this formula, each factor is nothing but the Coulomb potential produced at $\mathbf{x}_i$ by a point-like electrical charge Ze located at $\mathbf{X}_0$. Therefore, in this limit, the scattering of an electron off the target Ψ is equivalent to a scattering off an external classical field. One would have obtained the same answer simply by adding the Coulomb potential $A_\Psi^\mu(x)$ into the covariant derivative:

$$D^\mu = \partial^\mu - ie\,A^\mu \quad\to\quad \partial^\mu - ie\,(A^\mu + A_\Psi^\mu).$$

This approximation is known as the *external field approximation* or *Weizsäcker–Williams approximation*. Its derivation shows unambiguously that the term of order n in this external field is equivalent to summing over Feynman graphs with n exchanged photons, *summed over all the ways of permuting the attachments of the photons*. Thus, the (single) term of order n in the expansion in powers of the external field contains the contributions from $n!$ Feynman diagrams:



13. Subleading Soft Radiation in Scalar QED The goal of this problem and the next one is to study the first correction to the formula that gives the amplitude for the emission of a soft photon off a hard process. In the present problem, we consider the simpler case where the charged particles are scalar; the result is generalized to spin-1/2 charges in Problem 14.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

13.a Show that the formula (1.50) in the statement of Problem 10, which describes the emission of a soft photon of momentum k from an amplitude with hard spin-1/2 charged particles of momenta p_i, is unchanged when the charged particles are scalar, i.e.,

$$\mathcal{M}^\mu(k; p_1, \ldots, p_n) = \mathcal{M}(p_1, \ldots, p_n) \sum_{i=1}^{n} \frac{e_i\, p_i^\mu}{p_i \cdot k} + \mathcal{O}(k^0).$$

We recall that, at this order in k, the only contributions are obtained when the soft photon is attached to an external line of the hard process.

13.b Our goal is now to calculate the first subleading correction in k to this formula, in order to obtain the following improved result:

$$\mathcal{M}^\mu(k; p_1, \ldots, p_n) = \sum_{i=1}^{n} \frac{e_i}{p_i \cdot k} \Big(p_i^\mu - i\, k_\nu\, J_i^{\mu\nu}\Big)\, \mathcal{M}(p_1, \ldots, p_n) + \mathcal{O}(k^1),$$

where $J_i^{\mu\nu} \equiv i\big(p_i^\mu \frac{\partial}{\partial p_{i\nu}} - p_i^\nu \frac{\partial}{\partial p_{i\mu}}\big)$ is the angular momentum operator. *Hints: consider more accurately the insertions of the photon on the external lines. Then, note that there is also a term where the photon is attached to an inner line, but that the leading piece of this contribution is fully constrained by the Ward–Takahashi identities.*

13.a Consider an amplitude $\mathcal{M}(p_1, \ldots, p_n)$ in *scalar QED*, with all external particles hard. The emission of a soft photon of momentum k off this amplitude is dominated by Feynman graphs where the soft photon is attached to one of the external lines (the analysis of the denominators is the same in scalar and in spinor QED). The only change comes from the coupling of the photon to a scalar line: the Feynman rule for attaching a photon of momentum k and Lorentz index μ to a scalar of momentum p and charge e is a factor $-ie(2p + k)^\mu$. Therefore, when attaching a soft photon to an *outgoing scalar* particle of momentum p, the amplitude is modified by a factor

$$-i\bar{e}\,(2p + k)^\mu \frac{i}{(p + k)^2 - m^2} \underset{k \to 0}{\approx} \frac{\bar{e}\, p^\mu}{p \cdot k}.$$

(All the momenta are defined to be outgoing.) Note that here $\bar{e}$ is the coupling constant as it appears in the definition of the covariant derivative, $D_\mu \equiv \partial_\mu - i\bar{e}A_\mu$. In the case of an outgoing scalar, this parameter is also the electrical charge that flows outwards. But consider now an *outgoing antiscalar* of momentum p. In this case, the amplitude is modified by a factor

$$-i\bar{e}\,(-2p - k)^\mu \frac{i}{(p + k)^2 - m^2} \underset{k \to 0}{\approx} \frac{-\bar{e}\, p^\mu}{p \cdot k},$$

which differs by a sign from the case of an outgoing scalar. It is possible to make the subsequent formulas look more uniform if we introduce for each external line the corresponding outwards flowing electrical charge, denoted e. The relationship between this charge and the coupling constant $\bar{e}$ is simply

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

- *Outgoing scalar:* $e = \overline{e}$
- *Outgoing antiscalar:* $e = -\overline{e}$
- *Incoming scalar:* $e = -\overline{e}$
- *Incoming antiscalar:* $e = \overline{e}$.

In terms of this outgoing electrical charge, the modification factor for the emission of a soft photon is $ep^\mu/p \cdot k$ for any kind of external line (in this respect, scalar QED is therefore identical to spinor QED).

Summing over all the charged external lines, we thus obtain

$$\mathcal{M}^\mu(k; p_1, \dots, p_n) \underset{k \ll p_i}{\approx} \mathcal{M}(p_1, \dots, p_n) \sum_{i=1}^{n} \frac{e_i\, p_i^\mu}{p_i \cdot k}, \tag{1.58}$$

where the sum runs over all the charged external lines of the hard process. The first neglected terms in this formula are of order zero in k^μ. Recall that, with the conventions used in this problem, conservation of momentum and conservation of electrical charge in the hard process read

$$\sum_{i=1}^{n} p_i^\mu = 0, \quad \sum_{i=1}^{n} e_i = 0.$$

13.b Our task is now to determine the first subleading correction to (1.58), i.e., the next term in the formal expansion in powers of k/p_i. In order to calculate this correction, we need two things:

i. Improve the approximation made when the soft photon is attached to a charged external line of the hard process, so that it remains accurate at least up to the next order in k.

ii. Consider graphs where the soft photon is attached to an internal line of the hard process.

Let us start with **i**. Without doing any approximation, the attachment of a soft on-shell (i.e., $k^2 = 0$) photon to an outgoing scalar line leads to

$$\sum_{i=1}^{n} e_i \frac{(2p_i + k)^\mu}{(p_i + k)^2 - m_i^2}\, \mathcal{M}(\dots, p_i + k, \dots) = \sum_{i=1}^{n} e_i \frac{(2p_i + k)^\mu}{2p_i \cdot k}\, \mathcal{M}(\dots, p_i + k, \dots).$$

Note that this will simplify further when we contract the Lorentz index μ of the soft photon with a physical polarization vector, since $k^\mu \epsilon_{\lambda\mu}(\boldsymbol{k}) = 0$. This contraction will get rid of the term in k^μ, leaving the same eikonal factor $p_i^\mu/(p_i \cdot k)$ as in (1.58). Anticipating this contraction, the only difference from (1.58) is that the hard amplitude must be evaluated at $p_i + k$ instead of p_i when the i-th line emits the soft photon. This is true for all the possible types of external scalar lines (incoming/outgoing, scalar/antiscalar).

Now consider **ii**, i.e., the contributions where the soft photon comes from the interior of the hard process. As we shall show, these extra contributions are constrained by gauge invariance,

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

and these constraints are sufficient to determine the dominant terms in these contributions. Let us write the total amplitude as follows:

$$\mathcal{M}^\mu(k;p_1,\dots,p_n) = \underbrace{\sum_{i=1}^{n} e_i \frac{(2p_i+k)^\mu}{2p_i\cdot k}\,\mathcal{M}(\dots,p_i+k,\dots)}_{\text{photon attached to external lines}} + \underbrace{\mathcal{N}^\mu(k;p_1,\dots,p_n)}_{\text{photon attached to an internal line}} .$$

An important point is that the two terms are in general not gauge invariant separately, but their sum is. In particular, the full amplitude must obey the following Ward–Takahashi identity:

$$\begin{aligned} 0 &= k_\mu\,\mathcal{M}^\mu(k;p_1,\dots,p_n) \\ &= \sum_{i=1}^{n} e_i\,\mathcal{M}(\dots,p_i+k,\dots) + k_\mu\,\mathcal{N}^\mu(k;p_1,\dots,p_n). \end{aligned} \tag{1.59}$$

(We have used the fact that the photon is on-shell, $k^2 = 0$, in order to simplify the first term.) Note that the term $\mathcal{N}^\mu$ is not singular when $k \to 0$, since the leading k^{-1} behavior of the amplitude is entirely contained in the first term. Thus, the expansion of $\mathcal{N}^\mu$ in powers of k starts with a constant, followed by a term linear in k, etc.:

$$\mathcal{N}^\mu(k;p_1,\dots,p_n) = \mathcal{N}_0^\mu(p_1,\dots,p_n) + k_\nu\,\mathcal{N}_1^{\mu\nu}(p_1,\dots,p_n) + \cdots .$$

We can similarly expand the quantity $\mathcal{M}(\dots,p_i+k,\dots)$:

$$\mathcal{M}(\dots,p_i+k,\dots) = \mathcal{M}(p_1,\dots,p_n) + k_\mu \frac{\partial}{\partial p_{i\mu}}\mathcal{M}(p_1,\dots,p_n) + \cdots .$$

The next step is to investigate the consequence of the Ward–Takahashi identity (1.59) order-by-order in k. At order k^{-1}, only the term in $\mathcal{M}$ contributes, and the Ward–Takahashi identity is trivially satisfied thanks to charge conservation in the hard process. At the next order (k^0), the Ward–Takahashi identity reads

$$0 = \sum_{i=1}^{n} e_i\; k_\mu \frac{\partial}{\partial p_{i\mu}}\mathcal{M}(p_1,\dots,p_n) + k_\mu\,\mathcal{N}_0^\mu(p_1,\dots,p_n).$$

Therefore, we have

$$\mathcal{N}_0^\mu(p_1,\dots,p_n) = -\sum_{i=1}^{n} e_i\,\frac{\partial}{\partial p_{i\mu}}\mathcal{M}(p_1,\dots,p_n) + G^\mu(p_1,\dots,p_n),$$

where G^μ is a term that fulfills the Ward–Takahashi identity by itself, $k_\mu G^\mu = 0$. But note that such a G^μ would have to be independent of the photon momentum k, and yet satisfy this identity for any k. The only possibility is therefore $G^\mu(p_1,\dots,p_n) = 0$. Thus, the leading term in $\mathcal{N}^\mu$ is in fact fully determined by gauge invariance, and we have

$$\begin{aligned} &\epsilon_{\lambda\mu}(\mathbf{k})\,\mathcal{M}^\mu(k;p_1,\dots,p_n) \\ &\quad= \sum_{i=1}^{n} e_i \Bigg(\underbrace{\frac{p_i\cdot\epsilon_\lambda(\mathbf{k})}{p_i\cdot k}}_{\mathcal{O}(k^{-1})} + \underbrace{\left(\frac{p_i\cdot\epsilon_\lambda(\mathbf{k})}{p_i\cdot k}\,k_\nu - \epsilon_{\lambda\nu}(\mathbf{k})\right)\frac{\partial}{\partial p_{i\nu}}}_{\mathcal{O}(k^{0})} \Bigg)\mathcal{M}(p_1,\dots,p_n) + \mathcal{O}(k^1). \end{aligned}$$

Quite remarkably, it is not just the leading term, but also the first subleading correction in the photon momentum, that we can write in this factorized form (provided we allow derivatives

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

with respect to the p_i in the formula). The first non-factorizable term occurs only two orders down, with a suppression $(k/p_i)^2$ relative to the leading term. Note that the above expression can be written in a somewhat more symmetric way:

$$\epsilon_{\lambda\mu}(\mathbf{k})\,\mathcal{M}^\mu(k;p_1,\ldots,p_n)$$
$$= \sum_{i=1}^{n} \frac{e_i\,p_i^\mu}{p_i\cdot k}\Big(\epsilon_{\lambda\mu}(\mathbf{k}) + \big(\epsilon_{\lambda\mu}(\mathbf{k})k_\nu - k_\mu\epsilon_{\lambda\nu}(\mathbf{k})\big)\frac{\partial}{\partial p_{i\nu}}\Big)\,\mathcal{M}(p_1,\ldots,p_n) + \mathcal{O}(k^1).$$

Note also that

$$\Big(\underbrace{\epsilon_{\lambda\mu}(\mathbf{k})k_\nu - k_\mu\epsilon_{\lambda\nu}(\mathbf{k})}_{\text{antisymmetric in }(\mu,\nu)}\Big)\,p_i^\mu\,\frac{\partial}{\partial p_{i\nu}} = \frac{\epsilon_{\lambda\mu}(\mathbf{k})k_\nu - k_\mu\epsilon_{\lambda\nu}(\mathbf{k})}{2}\Big(p_i^\mu\,\frac{\partial}{\partial p_{i\nu}} - p_i^\nu\,\frac{\partial}{\partial p_{i\mu}}\Big)$$
$$= -i\,\epsilon_{\lambda\mu}(\mathbf{k})k_\nu\,\underbrace{\Big(i\,p_i^\mu\,\frac{\partial}{\partial p_{i\nu}} - i\,p_i^\nu\,\frac{\partial}{\partial p_{i\mu}}\Big)}_{\text{angular momentum op. }J_i^{\mu\nu}}.$$

Therefore, a very compact way of writing the amplitude for soft emission is

$$\epsilon_{\lambda\mu}(\mathbf{k})\,\mathcal{M}^\mu(k;p_1,\ldots,p_n)$$
$$= \epsilon_{\lambda\mu}(\mathbf{k})\sum_{i=1}^{n}\frac{e_i}{p_i\cdot k}\Big(p_i^\mu - i\,k_\nu\,J_i^{\mu\nu}\Big)\,\mathcal{M}(p_1,\ldots,p_n) + \mathcal{O}(k^1).$$

This result was first derived by Low (Low, F. (1958), *Phys Rev* 110: 974). It has been subsequently extended to Dirac fermions by Burnett and Kroll (see Problem 14), and it is now known as the *Low–Burnett–Kroll theorem.*

A consequence of this factorization is that, if a hard process is forbidden because of some conflicting quantum numbers, emitting a soft photon will not resolve the conflict (this is true in the first two orders in an expansion in the soft photon energy). For instance, if some decay is forbidden by charge parity (e.g., via Furry's theorem), one may naively think that emitting an extra photon – no matter how soft – lifts the obstruction since the charge parity of a photon is -1. Moreover, since extra soft photons have a probability of order one of being emitted, thanks to the terms in k^{-1}, this extra emission would not lead to a suppressed decay rate, implying that charge parity is in practice not conserved, at odds with experimental evidence. But the above result tells us that this conclusion is not true: it is only via the terms of order k^1 that the extra photon stands a chance of lifting the obstruction, and therefore it cannot be soft (thus, emitting it truly brings a suppression factor $\sim \alpha$). For this reason, another way this theorem is sometimes phrased is by saying that "soft photons do not carry quantum numbers."

14. Low–Burnett–Kroll Theorem In this problem, we extend the result of Problem 13 to the case of QED with spin-1/2 charges. As we shall see, the generalization requires that one sum the squared amplitude over the spins of the particles in the hard process, and the resulting formula reads

$$\sum_{\text{spins}}\Big|\epsilon_{\lambda\mu}(\mathbf{k})\mathcal{M}^\mu\Big|^2 = \sum_{i,j=1}^{n}\frac{e_ie_j(p_i\cdot\epsilon_\lambda)}{(p_i\cdot k)(p_j\cdot k)}\Big(p_j\cdot\epsilon_\lambda - i\,\epsilon_{\lambda\mu}k_\nu J_j^{\mu\nu}\Big)\sum_{\text{spins}}|\mathcal{M}|^2 + \mathcal{O}(k^0).$$

This result implies that corrections to a hard process by the emission of an extra soft photon are in general of order $\alpha(\Lambda/k)^2$ relative to the hard process, where Λ is the hard scale. (However,

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

the divergence that would arise when integrating over the energy of the soft photon is canceled when one also includes virtual corrections with a soft photon in a loop.) More importantly, this formula also shows that, when a hard process is forbidden by charge parity, emitting an extra soft photon is not sufficient to lift the obstruction despite the fact that the extra photon flips the charge parity, at least up to a relative order $\alpha(\Lambda/k)^0$. For this reason, this result is often phrased by saying that "soft photons do not carry quantum numbers."

14.a Adapt the QED Feynman rules to the convention where all the external momenta in the amplitude $\mathcal{M}$ are defined to be outgoing.

14.b Consider an outgoing fermion of momentum $\mathbf{p}_i$, spin s_i and charge e_i in the hard amplitude, and single out the corresponding external spinor by writing $\mathcal{M} \equiv \overline{u}_{s_i}(\mathbf{p}_i)\mathcal{A}_{s_i}$. Show that attaching a soft photon to this external line can be done by the substitution

$$\overline{u}_{s_i}(\mathbf{p}_i)\,\mathcal{A}_{s_i} \quad \to \quad e_i\,\overline{u}_{s_i}(\mathbf{p}_i)\,\frac{p_i^\mu + p_i^\mu k_\nu \partial_{p_i}^\nu + \frac{1}{2}\gamma^\mu \not{k}}{p_i \cdot k}\,\mathcal{A}_{s_i} + \mathcal{O}(k).$$

Determine the substitution rules for attaching a soft photon to all the other types of external lines.

14.c Use the Ward–Takahashi identity to determine the term where the soft photon is attached to an internal line of the hard process. Write the total amplitude as

$$\mathcal{M}^\mu(k;p_1,\ldots,p_n) = \sum_i e_i\,\frac{p_i^\mu - ik_\nu\,\left[J_i^{\mu\nu}\right]_{\mathcal{A},\mathcal{B}}}{p_i \cdot k}\,\mathcal{M}(p_1,\ldots,p_n) + \mathcal{E}^\mu + \mathcal{O}(k^1),$$

$$\mathcal{E}^\mu \equiv \sum_{i\in\{\overline{u},\overline{v}\}} \left\{\begin{matrix}\overline{u}_{s_i}(\mathbf{p}_i)\\ \overline{v}_{s_i}(-\mathbf{p}_i)\end{matrix}\right\}\frac{e_i\gamma^\mu\not{k}}{2p_i\cdot k}\,\mathcal{A}_{s_i} + \sum_{i\in\{u,v\}} \overline{\mathcal{B}}_{s_i}\,\frac{e_i\not{k}\gamma^\mu}{2p_i\cdot k}\left\{\begin{matrix}u_{s_i}(-\mathbf{p}_i)\\ v_{s_i}(\mathbf{p}_i)\end{matrix}\right\},$$

where the subscript $\mathcal{A},\mathcal{B}$ on the angular momentum operators $J^{\mu\nu}$ indicates that they should not act on the external spinors contained in the hard amplitude $\mathcal{M}$.

14.d Square this amplitude and sum over the spins of the hard external fermions in order to prove the announced formula.

14.a In order to provide an easy connection with Problem 13, we use a convention in which all the momenta are outgoing. In spinor quantum electrodynamics, this implies somewhat unusual assignments for the arguments of the spinors that represent the external fermions, and the Dirac equations they satisfy are also modified:

	Feynman rule	Dirac equation
Outgoing fermion	$\overline{u}_s(\mathbf{p})$	$\overline{u}(\mathbf{p})(\not{p}-m)=0$
Outgoing anti-fermion	$v_s(\mathbf{p})$	$(\not{p}+m)v(\mathbf{p})=0$
Incoming fermion	$u_s(-\mathbf{p})$	$(\not{p}+m)u(-\mathbf{p})=0$
Incoming anti-fermion	$\overline{v}_s(-\mathbf{p})$	$\overline{v}(-\mathbf{p})(\not{p}-m)=0$

14.b Consider an amplitude $\mathcal{M}(p_1, \ldots, p_n)$ with n hard external fermions or anti-fermions. When an extra soft photon is emitted in this process, the leading contributions come from Feynman diagrams where the extra photon is attached to one of these external lines. However, since we now want to calculate the next-to-leading contribution (in an expansion in powers of the photon momentum k), the calculation of the attachments of the photon to the external lines must be refined. Since we have to consider in turn every external line, it is convenient to write the hard amplitude in a factorized form that highlights the spinor corresponding to a given external line. Depending on the type of external line under consideration, we write

$$\mathcal{M} \equiv \overline{u}_{s_i}(\mathbf{p}_i)\,\mathcal{A}_{s_i} \equiv \overline{v}_{s_i}(-\mathbf{p}_i)\,\mathcal{A}_{s_i} \equiv \overline{\mathcal{B}}_{s_i}\,u_{s_i}(-\mathbf{p}_i) \equiv \overline{\mathcal{B}}_{s_i}\,v_{s_i}(\mathbf{p}_i).$$

(In these expressions, there is no sum on the index i.) Let us start with the case of an *outgoing fermion* of momentum $\mathbf{p}_i$. Attaching the soft photon of momentum k to this line gives a contribution that corresponds to the following substitution:

$$\begin{aligned}
&\overline{u}_{s_i}(\mathbf{p}_i)\,\mathcal{A}_{s_i}(\ldots, p_i, \ldots) \\
&\quad\to \quad \overline{u}_{s_i}(\mathbf{p}_i)\,\overline{e}\,\gamma^\mu \frac{\not{p}_i + m + \not{k}}{2\,p_i \cdot k}\,\mathcal{A}_{s_i}(\ldots, p_i + k, \ldots) \\
&\qquad = \frac{\overline{e}}{p_i \cdot k}\,\overline{u}_{s_i}(\mathbf{p}_i)\left\{p_i^\mu + p_i^\mu k_\nu \partial_{p_i}^\nu + \tfrac{1}{2}\gamma^\mu \not{k}\right\}\mathcal{A}_{s_i}(\ldots, p_i, \ldots) + \mathcal{O}(k).
\end{aligned}$$

In the second line, the denominator has been simplified thanks to the assumption that the external fermion as well as the photon are on-shell. In the last line, we have used the Dirac equation and we have expanded the function $\mathcal{A}_{s_i}(\ldots, p_i + k, \ldots)$ in order to extract the linear order in k. As in Problem 13, $\overline{e}$ denotes the coupling constant that enters in the definition of the covariant derivative, which is equal to the charge of a fermion and is opposite to that of an anti-fermion.

Consider now a soft photon attached to an *outgoing anti-fermion* of momentum $\mathbf{p}_i$. This leads to

$$\begin{aligned}
&\overline{\mathcal{B}}_{s_i}(\ldots, p_i, \ldots)\,v_{s_i}(\mathbf{p}_i) \\
&\quad\to \quad \overline{\mathcal{B}}_{s_i}(\ldots, p_i + k, \ldots)\,\frac{-\not{p}_i + m - \not{k}}{2\,p_i \cdot k}\,\overline{e}\,\gamma^\mu\,v_{s_i}(\mathbf{p}_i) \\
&\qquad = \overline{\mathcal{B}}_{s_i}(\ldots, p_i, \ldots)\left\{p_i^\mu + p_i^\mu k_\nu \overleftarrow{\partial}{}_{p_i}^{\nu} + \tfrac{1}{2}\not{k}\gamma^\mu\right\}v_{s_i}(\mathbf{p}_i)\,\frac{-\overline{e}}{p_i \cdot k} + \mathcal{O}(k).
\end{aligned}$$

Two remarks are in order about this formula. First, note that the derivative with respect to p_i acts on the left, as indicated by an arrow. Second, in the eikonal factor $-\overline{e}/p_i \cdot k$, the numerator $-\overline{e}$ is nothing but the electrical charge *flowing outwards*. In order to further uniformize the notation, it is convenient to express everything in terms of the outgoing electrical charge e, whose relationship with $\overline{e}$ depends on the type of external line under consideration:

- *Outgoing fermion:* $e = \overline{e}$
- *Outgoing anti-fermion:* $e = -\overline{e}$
- *Incoming fermion:* $e = -\overline{e}$
- *Incoming anti-fermion:* $e = \overline{e}$.

With these conventions, it is easy to work out the remaining two cases (incoming fermions and anti-fermions). All the rules for attaching a soft photon, up to terms of order k, to the four types of external fermions lines can be summarized by

$$\begin{Bmatrix} \overline{u}_{s_i}(\mathbf{p}_i) \\ \overline{v}_{s_i}(-\mathbf{p}_i) \end{Bmatrix} \mathcal{A}_{s_i} \quad \to \quad e_i \begin{Bmatrix} \overline{u}_{s_i}(\mathbf{p}_i) \\ \overline{v}_{s_i}(-\mathbf{p}_i) \end{Bmatrix} \frac{p_i^\mu + p_i^\mu k_\nu \overrightarrow{\partial}{}_{p_i}^\nu + \frac{1}{2}\gamma^\mu \not{k}}{p_i \cdot k} \mathcal{A}_{s_i},$$

$$\overline{\mathcal{B}}_{s_i} \begin{Bmatrix} u_{s_i}(-\mathbf{p}_i) \\ v_{s_i}(\mathbf{p}_i) \end{Bmatrix} \quad \to \quad e_i \overline{\mathcal{B}}_{s_i} \frac{p_i^\mu + p_i^\mu k_\nu \overleftarrow{\partial}{}_{p_i}^\nu + \frac{1}{2}\not{k}\gamma^\mu}{p_i \cdot k} \begin{Bmatrix} u_{s_i}(-\mathbf{p}_i) \\ v_{s_i}(\mathbf{p}_i) \end{Bmatrix}.$$

Recall also that with outgoing momenta and electrical charges, the conservation laws in the hard process take the following simple forms:

$$\sum_i p_i^\mu = 0, \quad \sum_i e_i = 0.$$

14.c The amplitude for the emission of an extra soft photon can be split into a term where the photon is attached to one of the external lines of the hard process and a term where the photon is attached to an internal line:

$$\mathcal{M}^\mu(k; p_1, \ldots, p_n) \equiv \mathcal{M}^\mu_{\text{ext}}(k; p_1, \ldots, p_n) + \mathcal{M}^\mu_{\text{inner}}(k; p_1, \ldots, p_n).$$

The first term is obtained by summing the results of the preceding section over all the external lines:

$$\mathcal{M}^\mu_{\text{ext}}(k; p_1, \ldots, p_n) = \sum_{i \in \{\overline{u}, \overline{v}\}} e_i \begin{Bmatrix} \overline{u}_{s_i}(\mathbf{p}_i) \\ \overline{v}_{s_i}(-\mathbf{p}_i) \end{Bmatrix} \frac{p_i^\mu + p_i^\mu k_\nu \overrightarrow{\partial}{}_{p_i}^\nu + \frac{1}{2}\gamma^\mu \not{k}}{p_i \cdot k} \mathcal{A}_{s_i}$$
$$+ \sum_{i \in \{u, v\}} e_i \overline{\mathcal{B}}_{s_i} \frac{p_i^\mu + p_i^\mu k_\nu \overleftarrow{\partial}{}_{p_i}^\nu + \frac{1}{2}\not{k}\gamma^\mu}{p_i \cdot k} \begin{Bmatrix} u_{s_i}(-\mathbf{p}_i) \\ v_{s_i}(\mathbf{p}_i) \end{Bmatrix},$$

where the first line contains the contributions from outgoing fermions and incoming anti-fermions, and the second line that of incoming fermions and outgoing anti-fermions (in each case, one should choose the appropriate spinor from the two listed). The above expression can also be rearranged as follows:

$$\mathcal{M}^\mu_{\text{ext}}(k; p_1, \ldots, p_n) = \left(\sum_i e_i \frac{p_i^\mu + p_i^\mu k_\nu \partial_{p_i}^\nu \big|_{\mathcal{A}, \mathcal{B}}}{p_i \cdot k} \right) \mathcal{M}(p_1, \ldots, p_n)$$
$$+ \sum_{i \in \{\overline{u}, \overline{v}\}} \begin{Bmatrix} \overline{u}_{s_i}(\mathbf{p}_i) \\ \overline{v}_{s_i}(-\mathbf{p}_i) \end{Bmatrix} \frac{\frac{e_i}{2}\gamma^\mu \not{k}}{p_i \cdot k} \mathcal{A}_{s_i} + \sum_{i \in \{u, v\}} \overline{\mathcal{B}}_{s_i} \frac{\frac{e_i}{2}\not{k}\gamma^\mu}{p_i \cdot k} \begin{Bmatrix} u_{s_i}(-\mathbf{p}_i) \\ v_{s_i}(\mathbf{p}_i) \end{Bmatrix},$$

where the subscript $\mathcal{A}, \mathcal{B}$ added to the derivative with respect to p_i in the first line indicates that it does not act on the spinors associated with the external lines in the hard amplitude $\mathcal{M}$.

We now need to determine the term $\mathcal{M}^\mu_{\text{inner}}$ where the photon is attached to an inner line of the hard process. This term starts at order k^0, and therefore, to the accuracy of the expansion

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

considered here, it is independent of k. As in the scalar QED case, this term can be constrained by requesting that the full amplitude obey the Ward–Takahashi identity when we contract it with the photon momentum k_μ. Note that the terms in $\not{k}$ vanish identically in this contraction since $\not{k}\not{k} = k^2 = 0$. Using the fact that the total outgoing charge in the hard process is zero, the Ward–Takahashi identity gives

$$\mathcal{M}^\mu_{\text{inner}}(p_1, \dots, p_n) = -\left(\sum_i e_i \left.\frac{\partial}{\partial p_{i\mu}}\right|_{\mathcal{A},\mathcal{B}} \right) \mathcal{M}(p_1, \dots, p_n).$$

(A priori, we could add to this expression a term G^μ such that $k_\mu G^\mu = 0$. But since G^μ should be independent of k, the only way to satisfy this for all k_μ is to have $G^\mu \equiv 0$.)

When we combine this term with $\mathcal{M}^\mu_{\text{ext}}$, the terms with derivatives with respect to the external momenta p_i can be arranged into angular momentum operators $J_i^{\mu\nu} \equiv -i(p_i^\mu \partial_{p_i\nu} - p^\nu \partial_{p_i\mu})$. This gives

$$\mathcal{M}^\mu(k; p_1, \dots, p_n) = \sum_i e_i \frac{p_i^\mu - ik_\nu \left[J_i^{\mu\nu}\right]_{\mathcal{A},\mathcal{B}}}{p_i \cdot k} \mathcal{M}(p_1, \dots, p_n) + \mathcal{E}^\mu + \mathcal{O}(k^1),$$

$$\mathcal{E}^\mu \equiv \sum_{i \in \{\overline{u}, \overline{v}\}} \begin{Bmatrix} \overline{u}_{s_i}(\mathbf{p}_i) \\ \overline{v}_{s_i}(-\mathbf{p}_i) \end{Bmatrix} \frac{e_i \gamma^\mu \not{k}}{2p_i \cdot k} \mathcal{A}_{s_i} + \sum_{i \in \{u, v\}} \overline{\mathcal{B}}_{s_i} \frac{e_i \not{k} \gamma^\mu}{2p_i \cdot k} \begin{Bmatrix} u_{s_i}(-\mathbf{p}_i) \\ v_{s_i}(\mathbf{p}_i) \end{Bmatrix}.$$

The first term is identical to the scalar case, except for the fact that the angular momentum operators do not act on the external spinors (as indicated by the subscript $\mathcal{A}, \mathcal{B}$). Another difference from the scalar QED case is that the second term, $\mathcal{E}^\mu$, does not vanish when contracted with a physical polarization vector.

14.d It turns out that, after squaring this amplitude and summing over the spins of the external hard fermions, the terms in $\mathcal{E}^\mu$ provide the missing terms to enlarge the scope of the action of the angular momentum operators to the full hard amplitude. In order to be consistent with the accuracy to which we have determined $\mathcal{M}^\mu$, we should keep only terms of order k^{-2} and k^{-1} in its square, since the higher orders would be incomplete. At this order, the squared amplitude reads

$$\sum_{\text{spins}} \left| \epsilon_{\lambda\mu}(\mathbf{k}) \mathcal{M}^\mu \right|^2 = \sum_{i,j} \underbrace{\frac{e_i e_j (p_i \cdot \epsilon_\lambda)}{(p_i \cdot k)(p_j \cdot k)} \Big((p_j \cdot \epsilon_\lambda) - i\, \epsilon_{\lambda\mu} k_\nu \left[J_j^{\mu\nu}\right]_{\mathcal{A},\mathcal{B}} \Big)}_{\mathcal{O}(k^{-2}) \oplus \mathcal{O}(k^{-1})} \sum_{\text{spins}} |\mathcal{M}|^2$$
$$+ \sum_i \underbrace{\frac{e_i (p_i \cdot \epsilon_\lambda)}{p_i \cdot k} \sum_{\text{spins}} \Big(\mathcal{M}^* \epsilon_{\lambda\mu} \mathcal{E}^\mu + \mathcal{M} \epsilon_{\lambda\mu} \mathcal{E}^{\mu *} \Big)}_{\mathcal{O}(k^{-1})}. \tag{1.60}$$

Note that in the first line, the derivatives contained in the angular momentum operators act on

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

the two factors of $|\mathcal{M}|^2 = \mathcal{M}^*\mathcal{M}$. In order to calculate the term of the second line, note that

$$\sum_{\text{spins}} \Big(\mathcal{M}^*\epsilon_{\lambda\mu}\mathcal{E}^\mu + \mathcal{M}\epsilon_{\lambda\mu}\mathcal{E}^{\mu*}\Big) = \sum_{j\in\{\overline{u},\overline{v}\}} e_j\,\overline{\mathcal{A}}_{s_j}\,\frac{\begin{Bmatrix}\not{p}_j+m\\ -\not{p}_j-m\end{Bmatrix}\not{\epsilon}\not{k}+\not{k}\not{\epsilon}\begin{Bmatrix}\not{p}_j+m\\ -\not{p}_j-m\end{Bmatrix}}{2p_j\cdot k}\mathcal{A}_{s_j}$$
$$+\sum_{j\in\{u,v\}} e_j\,\overline{\mathcal{B}}_{s_j}\,\frac{\begin{Bmatrix}-\not{p}_j+m\\ \not{p}_j-m\end{Bmatrix}\not{\epsilon}\not{k}+\not{k}\not{\epsilon}\begin{Bmatrix}-\not{p}_j+m\\ \not{p}_j-m\end{Bmatrix}}{2p_j\cdot k}\mathcal{B}_{s_j}.$$

Using the fact that $k\cdot\epsilon = 0$, we have

$$\begin{aligned}\frac{(\not{p}\pm m)\not{\epsilon}\not{k}+\not{k}\not{\epsilon}(\not{p}\pm m)}{2p\cdot k} &= \frac{(p\cdot\epsilon)\not{k}-(p\cdot k)\not{\epsilon}}{p\cdot k}\\ &= \frac{1}{p\cdot k}\big(-i\epsilon_\mu k_\nu\big[J_p^{\mu\nu}\big]\big)(\not{p}\pm m).\end{aligned}$$

The remarkable feature of this result is that it takes the form of an angular momentum operator acting on spin sums such as $\sum_s u_s(\mathbf{p})\overline{u}_s(\mathbf{p}) = \not{p}+m$. Thanks to this formula, the terms in the second line of (1.60) are precisely what is needed in order to lift the restrictions on the action of the angular momentum operators in the terms of the first line, allowing us to write the complete answer up to (and including) the order k^{-1} in a much more compact fashion:

$$\sum_{\text{spins}}\Big|\epsilon_{\lambda\mu}(\mathbf{k})\mathcal{M}^\mu\Big|^2 = \sum_{i,j}\frac{e_ie_j(p_i\cdot\epsilon_\lambda)}{(p_i\cdot k)(p_j\cdot k)}\Big((p_j\cdot\epsilon_\lambda)-i\,\epsilon_{\lambda\mu}k_\nu J_j^{\mu\nu}\Big)\sum_{\text{spins}}|\mathcal{M}|^2.$$

This formula, obtained by Burnett and Kroll (Burnett, T. and Kroll, N. M. (1968), *Phys Rev Lett* 20: 86), generalizes Low's theorem (established in Problem 13) to the case of hard charged external particles of spin $1/2$. And the main consequence is the same: if some unpolarized hard process is forbidden by charge parity, emitting extra soft photons is not sufficient to make it possible. (The emission of an extra photon may make this process possible via terms of order $\mathcal{O}(k^0)$, but these terms are not enhanced by inverse powers of the photon energy.)

15. Coherent States in Quantum Field Theory In quantum mechanics, *coherent states* are defined as eigenstates of the annihilation operators. In a certain sense, they are the closest analogues to classical states where the position and momentum are both well defined. The goal of this problem is to generalize the concept of coherent state to a non-interacting scalar quantum field theory, starting from the following definition:

$$\big|\chi_{\text{in}}\big\rangle \equiv \mathcal{N}_\chi\ \exp\Big\{\int\frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}}\,\chi(\mathbf{k})\,a^\dagger_{\mathbf{k},\text{in}}\Big\}\,\big|0_{\text{in}}\big\rangle,$$

where $\chi(\mathbf{k})$ is a function of 3-momentum and $\mathcal{N}_\chi$ a normalization constant.

15.a Check that this state is an eigenstate of $a_{\mathbf{k},\text{in}}$.

15.b Determine the prefactor $\mathcal{N}_\chi$ so that $\langle\chi_{\text{in}}|\chi_{\text{in}}\rangle = 1$. *Hint: introduce the operator*

$$T_\chi \equiv \exp \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \left(\chi(\mathbf{k})a^\dagger_{\mathbf{k},\text{in}} - \chi^*(\mathbf{k})a_{\mathbf{k},\text{in}}\right)$$

and write it in normal-ordered form.

15.c Show that $|\chi_{\text{in}}\rangle$ is the ground state of a free scalar field theory with a shifted field:

$$\phi(x) \to \phi(x) - \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \left(\chi^*(\mathbf{k})e^{ik\cdot x} + \chi(\mathbf{k})e^{-ik\cdot x}\right).$$

15.d Consider two coherent states $|\chi, \vartheta_{\text{in}}\rangle$, with $\chi(\mathbf{k}), \vartheta(\mathbf{k}) \propto \delta(\mathbf{k})$, and calculate $\langle\vartheta_{\text{in}}|\chi_{\text{in}}\rangle$.

15.a Let us denote

$$E_\chi \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \chi(\mathbf{k})\, a^\dagger_{\mathbf{k},\text{in}}.$$

The commutation relation between creation and annihilation operators leads to

$$\left[a_{\mathbf{k},\text{in}}, E_\chi\right] = \chi(\mathbf{k}), \qquad \left[a_{\mathbf{k},\text{in}}, f(E_\chi)\right] = \chi(\mathbf{k})\, f'(E_\chi).$$

This implies

$$\begin{aligned} a_{\mathbf{k},\text{in}}|\chi_{\text{in}}\rangle &= \mathcal{N}_\chi\, a_{\mathbf{k},\text{in}} e^{E_\chi}|0_{\text{in}}\rangle = \mathcal{N}_\chi \left[a_{\mathbf{k},\text{in}}, e^{E_\chi}\right]|0_{\text{in}}\rangle \\ &= \mathcal{N}_\chi\, \chi(\mathbf{k})\, e^{E_\chi}|0_{\text{in}}\rangle = \chi(\mathbf{k})\, |\chi_{\text{in}}\rangle. \end{aligned}$$

Thus, coherent states are eigenstates of the annihilation operators.

15.b Consider now the following operator:

$$T_\chi \equiv \exp \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \left(\chi(\mathbf{k})a^\dagger_{\mathbf{k},\text{in}} - \chi^*(\mathbf{k})a_{\mathbf{k},\text{in}}\right).$$

First, it is trivial to check that $T^\dagger_\chi = T_{-\chi} = T^{-1}_\chi$. (Therefore, T_χ is unitary.) Then, by using the Baker–Campbell–Hausdorff formula, we can put this operator in normal-ordered form:

$$\begin{aligned} T_\chi = \exp\Big(-\frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} |\chi(\mathbf{k})|^2\Big) \exp\Big(\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \chi(\mathbf{k})a^\dagger_{\mathbf{k},\text{in}}\Big) \\ \times \exp\Big(-\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \chi^*(\mathbf{k})a_{\mathbf{k},\text{in}}\Big). \end{aligned}$$

From this expression, we see that $|\chi_{\text{in}}\rangle \propto T_\chi|0_{\text{in}}\rangle$. Since T_χ is unitary, this implies that the normalization $\langle\chi_{\text{in}}|\chi_{\text{in}}\rangle = 1$ is obtained by choosing precisely $|\chi_{\text{in}}\rangle = T_\chi|0_{\text{in}}\rangle$, from which we read the value of the normalization prefactor $\mathcal{N}_\chi$:

$$\mathcal{N}_\chi = \exp\Big(-\frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} |\chi(\mathbf{k})|^2\Big).$$

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

15.c If $\mathcal{H}_0$ denotes the Hamiltonian of a free theory, let us now define $\mathcal{H}_\chi \equiv T_\chi \mathcal{H}_0 T_\chi^\dagger$. Note first that $\mathcal{H}_\chi$ is Hermitian. Then, we have

$$\mathcal{H}_\chi\big|\chi_{\rm in}\big\rangle = T_\chi \mathcal{H}_0 \underbrace{T_\chi^\dagger T_\chi}_{=1} \big|0_{\rm in}\big\rangle = 0.$$

Thus, the coherent state is the ground state of the modified Hamiltonian $\mathcal{H}_\chi$.

Using the commutation relations

$$\big[a_{\mathbf{k},\rm in}, T_\chi\big] = \chi(\mathbf{k})\,T_\chi, \quad \big[a^\dagger_{\mathbf{k},\rm in}, T_\chi\big] = \chi^*(\mathbf{k})\,T_\chi,$$

we obtain

$$T_\chi a_{\mathbf{k},\rm in} T_\chi^\dagger = a_{\mathbf{k},\rm in} - \chi(\mathbf{k}), \quad T_\chi a^\dagger_{\mathbf{k},\rm in} T_\chi^\dagger = a^\dagger_{\mathbf{k},\rm in} - \chi^*(\mathbf{k}),$$

and the modified Hamiltonian is more explicitly given by

$$\mathcal{H}_\chi = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}}\, E_\mathbf{k}\,\big(a^\dagger_{\mathbf{k},\rm in} - \chi^*(\mathbf{k})\big)\big(a_{\mathbf{k},\rm in} - \chi(\mathbf{k})\big).$$

(We have discarded the zero-point energy term, since it is not modified by the action of $T_\chi \cdots T_\chi^\dagger$.) Recalling the formulas that relate a free field operator and the creation and annihilation operators, we see that this is the Hamiltonian of a free theory with a shifted field:

$$\phi(x) \to \phi(x) - \Phi_\chi(x), \quad \Phi_\chi(x) \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}} \Big(\chi^*(\mathbf{k})e^{ik\cdot x} + \chi(\mathbf{k})e^{-ik\cdot x}\Big).$$

(Note that $\Phi_\chi(x)$ is an ordinary real-valued function, not an operator.)

15.d Consider two such coherent states, $\big|\chi_{\rm in}\big\rangle, \big|\vartheta_{\rm in}\big\rangle$, in the special case where their defining functions only have support at $\mathbf{k} = 0$: $\chi(\mathbf{k}) \equiv (2\pi)^3\chi_0\delta(\mathbf{k})$, $\vartheta(\mathbf{k}) \equiv (2\pi)^3\vartheta_0\delta(\mathbf{k})$ (we assume $\chi_0, \vartheta_0 \in \mathbb{R}$). The overlap of these two states is given by

$$\begin{aligned}\big\langle\vartheta_{\rm in}\big|\chi_{\rm in}\big\rangle &= \exp\Big(-\frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}}\Big[\big|\chi(\mathbf{k})\big|^2 + \big|\vartheta(\mathbf{k})\big|^2\Big]\Big)\\ &\quad\times \big\langle 0_{\rm in}\big|\exp\Big(\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}}\,\vartheta^*(\mathbf{k})\,a_{\mathbf{k},\rm in}\Big)\exp\Big(\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}}\,\chi(\mathbf{k})\,a^\dagger_{\mathbf{k},\rm in}\Big)\big|0_{\rm in}\big\rangle\\ &= \exp\Big(-\frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_\mathbf{k}}\big(\big|\chi(\mathbf{k}) - \vartheta(\mathbf{k})\big|^2 + \chi^*(\mathbf{k})\vartheta(\mathbf{k}) - \vartheta^*(\mathbf{k})\chi(\mathbf{k})\big)\Big)\\ &= \exp\Big(-\frac{V\,|\chi_0 - \vartheta_0|^2}{4m}\Big).\end{aligned}$$

Therefore, spatially homogeneous coherent states (i.e., ground states of quadratic theories shifted by a uniform field) have an exponentially suppressed overlap, and the argument of the exponential is proportional to the volume. Thus, pairs of coherent states of this type are mutually orthogonal if the volume is infinite.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

16. Running Couplings in a Two-Field Scalar Field Theory The purpose of this problem is to study at one loop the scale dependence of the coupling in a scalar field theory with quartic coupling, first with a single field and then with two coupled fields. In the latter case, we show that the theory evolves towards a $U(1)$ symmetry at large distance.

16.a Draw the graphs that contribute to the β function at lowest order (i.e., one-loop) in a scalar field theory with $\frac{\lambda}{4!}\phi^4$ interaction. Why don't we need to consider self-energy graphs at this order?

16.b Calculate the relevant part of these graphs in dimensional regularization, at the renormalization point where the Mandelstam variables are set to $s = t = u = -\mu^2$. Write their sum as $\frac{3}{2}\lambda^2 L(\mu)$, where $L(\mu)$ is a quantity that should be made explicit. From this, check that the β function is $\beta = 3\lambda^2/(4\pi)^2$.

16.c Next, we consider a theory with two scalar fields ϕ_1 and ϕ_2, with the following Lagrangian:

$$\mathcal{L} \equiv \tfrac{1}{2}(\partial_\mu\phi_1)(\partial^\mu\phi_1) + \tfrac{1}{2}(\partial_\mu\phi_2)(\partial^\mu\phi_2) - \tfrac{\lambda}{4!}(\phi_1^4 + \phi_2^4) - \tfrac{2\rho}{4!}\phi_1^2\phi_2^2.$$

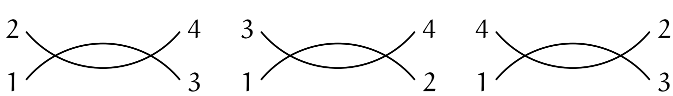
(Note that this theory has a $U(1)$ invariance in the (ϕ_1, ϕ_2) plane when $\lambda = \rho$.) What are the free Feynman propagators for the fields ϕ_1 and ϕ_2 (in momentum space)? What are the vertices in this theory and the corresponding Feynman rules?

16.d Draw the graphs that give, at lowest order, the β function that controls the scale dependence of the coupling constant λ. Show that the relevant part of these graphs is given by $\left(\frac{3}{2}\lambda^2 + \frac{1}{6}\rho^2\right)L(\mu)$, where $L(\mu)$ is the quantity obtained in question 16.b. Check that this β function is $\beta_\lambda = (3\lambda^2 + \rho^2/3)/(4\pi)^2$.

16.e Repeat question 16.d for the β function that controls the scale dependence of the coupling constant ρ. Check that it is $\beta_\rho = (2\lambda\rho + 4\rho^2/3)/(4\pi)^2$.

16.f Derive the renormalization group equation for the ratio ρ/λ. What are the fixed points of this equation? Is $\rho/\lambda = 1$ an attractive fixed point in the ultraviolet? in the infrared?

16.a The calculation of the β function requires the coupling counterterm δ_λ and the wave-function renormalization counterterm δ_z. δ_λ is obtained from the one-loop corrections to the vertex function (i.e., the four-point function in the case of the ϕ^4 theory):



(Note that there are three ways of attaching the four external lines to the loop.) δ_z comes from the momentum dependence of the self-energy. But in the ϕ^4 theory, the self-energy is a constant at one loop, and therefore one has $\delta_z = 0$ at this order.

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

16.b We need to calculate only one of the above graphs, since the renormalization point we have chosen is symmetric under the exchange of the channels s, t, u. Moreover, since only the ultraviolet divergent part is necessary, we do not need to keep the mass in the propagators. With this in mind, let us consider the first graph. Its contribution to the amputated correlation function Γ_4 reads (we use dimensional regularization, with $d = 4 - 2\epsilon$):

$$\begin{aligned}\Gamma_4^{(1)}(p_{1,2,3,4}) &= i\mu^{2\epsilon}\frac{(-i\lambda)^2}{2}\int\frac{d^dk}{(2\pi)^d}\,\frac{i}{k^2}\,\frac{i}{(k+p_1+p_2)^2}\\ &= i\frac{\lambda^2\mu^{2\epsilon}}{2}\int_0^1 dx\int\frac{d^dk}{(2\pi)^d}\,\frac{1}{(k^2-\Delta)^2} = -\frac{\lambda^2\mu^{2\epsilon}}{2}\int_0^1 dx\int\frac{d^dk_{\mathrm{E}}}{(2\pi)^d}\,\frac{1}{(k_{\mathrm{E}}^2+\Delta)^2}\\ &= -\frac{\lambda^2\mu^{2\epsilon}}{2}\int_0^1 dx\,\frac{\Delta^{d/2-2}}{(4\pi)^{d/2}}\,\frac{\Gamma(2-\frac{d}{2})}{\Gamma(2)},\end{aligned}$$

with $\Delta \equiv -x(1-x)(p_1+p_2)^2 = x(1-x)\mu^2$. The x dependence of the integrand is of the form $(x(1-x))^\epsilon$ and therefore this integral produces a result of the form $1 + \mathcal{O}(\epsilon)$. Since we are only interested in the ultraviolet divergent part, only the 1 is needed. Thus, we can write

$$\Gamma_4^{(1)}(p_{1,2,3,4}) = -\frac{\lambda^2}{2(4\pi)^2\epsilon}\,\frac{\mu^{2\epsilon}}{\mu^{2\epsilon}}.$$

Since the renormalization point is symmetric in the s, t, u channels, the above result is simply multiplied by 3 when we include all the graphs:

$$\Gamma_4^{(1+2+3)}(p_{1,2,3,4}) = \frac{3\lambda^2}{2}L(\mu), \quad \text{with } L(\mu) \equiv -\frac{1}{(4\pi)^2\epsilon}\,\frac{\mu^{2\epsilon}}{\mu^{2\epsilon}}.$$

In this notation, the prefactor $\frac{3}{2}$ is the number of graphs times their symmetry factor, the factor λ^2 is the contribution from the two vertices, and $L(\mu)$ is the value of the loop integral. The counterterm that removes this divergence is

$$\delta_\lambda = -\frac{3\lambda^2}{2}L(\mu),$$

which leads to the following β function:

$$\beta = -\lim_{\epsilon\to 0}\mu\partial_\mu\,\delta_\lambda = \frac{3\lambda^2}{(4\pi)^2}.$$

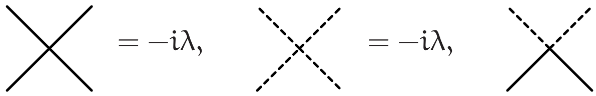
16.c Consider now an extension of this theory that has two scalar fields ϕ_1 and ϕ_2, with the following Lagrangian:

$$\mathcal{L} \equiv \tfrac{1}{2}(\partial_\mu\phi_1)(\partial^\mu\phi_1) + \tfrac{1}{2}(\partial_\mu\phi_2)(\partial^\mu\phi_2) - \tfrac{\lambda}{4!}(\phi_1^4+\phi_2^4) - \tfrac{2\rho}{4!}\phi_1^2\phi_2^2.$$

The two propagators of this theory have identical expressions:

$$\text{———} = \frac{i}{p^2+i0^+}, \qquad \text{- - - - - -} = \frac{i}{p^2+i0^+}.$$

(Even though the propagators are the same, it is important to use different symbols for the two fields in the diagrams, since there will be various vertices, depending on which fields are attached to them.) This theory has only quartic vertices, ϕ_1^4, ϕ_2^4 and $\phi_1^2\phi_2^2$:



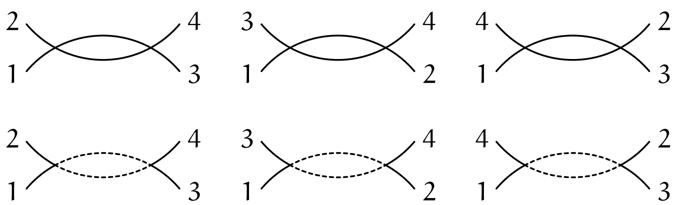
$$= -i\lambda, \qquad = -i\lambda, \qquad = -i\frac{\rho}{3}.$$

For the first two vertices, this is just the Feynman rule for the vertex of the standard ϕ^4 theory, when the interaction term in the Lagrangian is normalized as $\frac{\lambda}{4!}\phi^4$. For the third vertex, the Feynman rule is obtained as

$$-i \underbrace{\frac{2\rho}{4!}}_{\text{coeff. in } \mathcal{L}} \times 2! \times 2! = -i\,\frac{\rho}{3},$$

where the first factor 2! corresponds to the two ways of attaching the fields $\phi_1\phi_1$ to the two solid lines of the vertex, and the second factor 2! counts the number of ways of attaching the fields $\phi_2\phi_2$ to the two dotted lines of the vertex.

16.d In order to determine the β function that controls the scale dependence of the coupling λ, it is sufficient to calculate at one loop the four-point function with only external fields of type 1. The corresponding graphs are:



The values of the loops do not depend on the type of field running in them since we neglect masses in this calculation, and are therefore given by the same function $L(\mu)$ in all six graphs. We just need to count the graphs and their symmetry factors ($\frac{1}{2}$ in all cases), and assign them the appropriate vertices to find the correct four-point function:

$$\Gamma_4(p_{1,2,3,4}) = \left(\frac{3}{2}\lambda^2 + \frac{3}{2}\left(\frac{\rho}{3}\right)^2\right) L(\mu) = \left(\frac{3\lambda^2}{2} + \frac{\rho^2}{6}\right) L(\mu).$$

Since $L(\mu)$ is already known, we do not need any new calculation in order to obtain the β function in this case. It is sufficient to appropriately adjust the prefactor, to obtain

$$\beta_\lambda = \frac{3\lambda^2 + \frac{\rho^2}{3}}{(4\pi)^2}.$$

16.e To obtain the β function that drives the evolution of ρ, we need to consider the 4-point function with two external fields ϕ_1 and two external fields ϕ_2. To be definite, let us assume

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

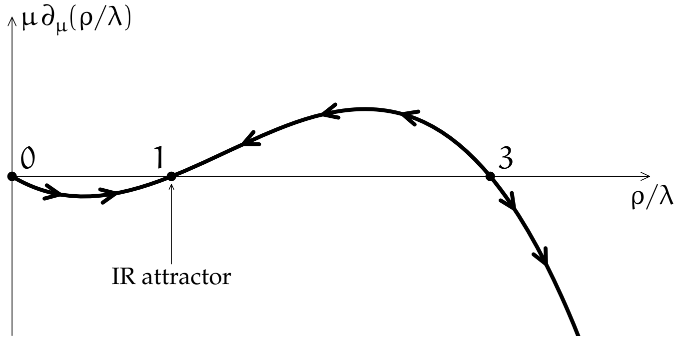
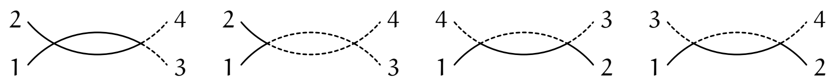


Figure 1.5 The β function that controls the running of the ratio ρ/λ. The arrows indicate the direction of the flow towards long distances.

that the legs $1, 2$ carry the field ϕ_1 and the legs $3, 4$ are attached to a field ϕ_2. Thus, the one-loop graphs that contribute to this function are the following:



We can directly write the corresponding expression for the four-point function in terms of the integral $L(\mu)$:

$$\Gamma_4(p_{1,2,3,4}) = \left(\frac{2}{2}\frac{\lambda\rho}{3} + 2\left(\frac{\rho}{3}\right)^2\right) L(\mu).$$

(The first two graphs have a symmetry factor $\frac{1}{2}$, while the remaining two have a symmetry factor 1.) From this, we directly obtain the following β function:

$$\beta_\rho = 3 \times \frac{\frac{2\lambda\rho}{3} + \frac{4\rho^2}{9}}{(4\pi)^2} = \frac{2\lambda\rho + \frac{4\rho^2}{3}}{(4\pi)^2}.$$

(On the right-hand side, the prefactor 3 is due to the fact that the four-point function we have calculated is a correction to $\rho/3$, and must therefore be multiplied by 3 in order to obtain a correction to ρ itself.)

16.f From the above results, we have

$$\mu\partial_\mu\lambda = \beta_\lambda, \quad \mu\partial_\mu\rho = \beta_\rho,$$

$$\mu\partial_\mu\left(\frac{\rho}{\lambda}\right) = \frac{\beta_\rho}{\lambda} - \frac{\rho\beta_\lambda}{\lambda^2} = -\frac{\rho}{3}\left(\frac{3 - \frac{4\rho}{\lambda} + \frac{\rho^2}{\lambda^2}}{(4\pi)^2}\right) = -\frac{\rho}{3(4\pi)^2}\left(\frac{\rho}{\lambda} - 1\right)\left(\frac{\rho}{\lambda} - 3\right).$$

From the factorized form of the right-hand side in the preceding equation, the fixed points are

$$\frac{\rho}{\lambda} = 0, 1, 3.$$

To decide whether they are attractive or repulsive, we just need to determine the sign of the right-hand side of this equation in the various intervals between the fixed points. Figure 1.5 shows that $\rho/\lambda = 1$ is an attractive fixed point in the infrared (and repulsive in the ultraviolet).

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

17. Solution of the Running Equation at Two Loops Consider a theory, such as QCD, in which the β function has the following perturbative expansion up to two loops: $\beta(g) = \beta_0 g^3 + \beta_1 g^5 + \mathcal{O}(g^7)$, with $\beta_0 < 0$. Show that the scale dependence of the coupling is given by

$$g(\mu^2) = -\frac{1}{\beta_0 \ln\left(\frac{\mu^2}{\Lambda^2}\right)}\left(1 + \frac{\beta_1}{\beta_0^2}\frac{\ln\ln\left(\frac{\mu^2}{\Lambda^2}\right)}{\ln\left(\frac{\mu^2}{\Lambda^2}\right)}\right) + \mathcal{O}\left(\ln^{-2}\left(\frac{\mu^2}{\Lambda^2}\right)\right).$$

Hint: recall that Λ is defined as the infrared scale where the coupling constant becomes infinite.

Given the β function, the coupling constant evolves according to $\mu\partial_\mu g = \beta$. This can be turned into a differential equation for the scale dependence of g^2:

$$\mu\frac{\partial g^2}{\partial\mu} = 2\beta_0 g^4 (1 + \underbrace{\frac{\beta_1}{\beta_0}}_{\equiv\, b\,\sim\, g^2} g^2),$$

which can be rewritten as

$$\frac{dg^2}{g^4(1+bg^2)} = \beta_0 \frac{d\mu^2}{\mu^2}.$$

To be definite, we integrate this differential equation between a coupling $g^2(\mu^2)$ at the scale μ^2 and a coupling $g^2(\Lambda^2) = \infty$ at the scale Λ^2. In other words, we define Λ to be the scale at which the running coupling becomes infinite. Note that this scale always exists if $\beta_0, \beta_1 < 0$, as is the case in QCD when the number of quark flavors is not too large (see Problem 47). In the following integral, it is legitimate to integrate up to $g^2 = \infty$ since the denominator $1 + bg^2$ does not vanish if $b > 0$. We get

$$\int_{g^2(\mu^2)}^{+\infty} \frac{dg^2}{g^4(1+bg^2)} = \beta_0 \int_{\mu^2}^{\Lambda^2} \frac{d\mu^2}{\mu^2} = -\beta_0 \ln\left(\frac{\mu^2}{\Lambda^2}\right).$$

In order to evaluate the integral on the left-hand side, we use

$$\int_{g^2(\mu^2)}^{+\infty} \frac{dg^2}{g^4(1+\underbrace{bg^2}_{x})} = b\int_{bg^2(\mu^2)}^{+\infty} \frac{dx}{x^2(1+x)} = \frac{1}{g^2(\mu^2)} + b\ln\left(\frac{b\, g^2(\mu^2)}{1 + b\, g^2(\mu^2)}\right).$$

Therefore, we have

$$\frac{1}{g^2(\mu^2)} + b\ln\left(\frac{b\, g^2(\mu^2)}{1 + b\, g^2(\mu^2)}\right) = -\beta_0 \ln\left(\frac{\mu^2}{\Lambda^2}\right).$$

Now, we must invert this equation in order to express $g^2(\mu^2)$ in terms of the logarithm that appears on the right-hand side. Unfortunately, this cannot be done in closed form, and we must

https://doi.org/10.1017/9781108976688.004 Published online by Cambridge University Press

perform some kind of expansion. Let us denote by $g_0^2(\mu^2)$ the solution obtained when keeping only the one-loop β function,

$$g_0^2(\mu^2) = -\frac{1}{\beta_0 \, \ln\left(\frac{\mu^2}{\Lambda^2}\right)},$$

and then write the full g^2 as

$$g^2(\mu^2) = g_0^2(\mu^2) + g_1^2(\mu^2).$$

When solving for g_1^2, we may replace g^2 by g_0^2 in the term in $b\ln(\cdot)$ because this term is formally of higher order. Therefore, we have

$$\frac{1}{g_0^2(\mu^2)}\left(1 - \frac{g_1^2(\mu^2)}{g_0^2(\mu^2)} + \cdots\right) \approx -\beta_0\,\ell - b\ln\left(\frac{b\,g_0^2(\mu^2)}{1 + b\,g_0^2(\mu^2)}\right),$$

where we denote $\ell \equiv \ln\left(\frac{\mu^2}{\Lambda^2}\right)$ for compactness. This leads to

$$\begin{aligned} g_1^2(\mu^2) &\approx b\,g_0^4(\mu^2)\,\ln\left(\frac{b\,g_0^2(\mu^2)}{1 + b\,g_0^2(\mu^2)}\right) \\ &= -\frac{\beta_1}{\beta_0^3\ell^2}\left(\ln(\ell) + \underbrace{\ln\left(\frac{\beta_0^2}{\beta_1} + \ell^{-1}\right)}_{\text{const } + \mathcal{O}(\ell^{-1})}\right) \\ &= -\frac{\beta_1 \ln(\ell)}{\beta_0^3\ell^2} + \mathcal{O}(\ell^{-2}). \end{aligned}$$

Therefore, we have

$$g(\mu^2) = -\frac{1}{\beta_0\,\ln\left(\frac{\mu^2}{\Lambda^2}\right)}\left(1 + \frac{\beta_1}{\beta_0^2}\,\frac{\ln\ln\left(\frac{\mu^2}{\Lambda^2}\right)}{\ln\left(\frac{\mu^2}{\Lambda^2}\right)}\right) + \mathcal{O}\left(\ln^{-2}\left(\frac{\mu^2}{\Lambda^2}\right)\right).$$