Lecture Notes for Quantum Field Theory

by

Charles B. Thorn

1. Introduction

One of the most basic facts about relativistic physics is the equivalence of mass and energy, as in Einstein's famous relation $E = mc^2$. This fact has profound consequences for relativistic quantum mechanics, because of Heisenberg's uncertainty principle

$$\Delta x \Delta p \geq \hbar. \tag{1.1}$$

In ordinary nonrelativistic quantum mechanics this follows from the fact that the position and momentum are represented by operators that do not commute: $[x, p] = i\hbar$. According to (1.1) an accurate position measurement can be made only at the expense of a compensatingly large inaccuracy in the momentum of the particle. This implies a correspondingly large expense of energy in the accurate measurement of position. In nonrelativistic quantum mechanics the rest energy of a massive particle is effectively infinite so one can subject a particle to arbitrarily large energy changes without compromising its integrity as a particle, and accurate position measurements are compatible with particle number conservation. The concept of an N particle wave function as a function of the N particle coordinates is sustainable. This is drastically changed in relativistic quantum mechanics. As soon as one attempts to measure the position of a particle to an accuracy much less than the particle's Compton wavelength, $\hbar/mc$, the associated momentum uncertainty is $> mc$, corresponding to an energy change larger than mc^2. Thus accurate position measurements require the supply of energy sufficient to produce additional particles. If the particle possesses a conserved charge, additional such particles can be produced but only in association with particles of opposite charge. Thus we should expect any consistent version of relativistic quantum mechanics to require abandoning the concept of a wave function for a system with a definite number of particles. Further, in retrospect, it is not surprising that when there are conserved charges, relativistic quantum physics requires antiparticles.

The preceding is not meant to be a rigorous argument, but rather a simple explanation of why we must expect at least the complexities of quantum field theory when we attempt to extend quantum mechanics into the relativistic domain. The many-body aspects of relativistic quantum mechanics are universal, but the manner in which they make their appearance varies. In quantum electrodynamics it is immediate once one tries to interpret classical radiation in terms of photons. For massive particles, like the electron, it arises from the existence of negative energy solutions of the Dirac equation.

The first semester of quantum field theory will focus primarily on quantum electrodynamics, the theory of electrons and positrons interacting with the quantized electromagnetic field. We shall begin however with a much simpler quantum field theory– that of a self interacting scalar field. Then we shall show how Dirac's theory of electrons and positrons can be converted to a quantum field theory through the device of "second quantization". We shall study this Dirac quantum field theory first in the presence of external (classical) electromagnetic fields. Only then will we turn to the problem of quantizing the electromagnetic field and studying it in interaction with electrons and positrons.

2. Lorentz Invariance

In elementary discussions of special relativity we learn that frames of reference with a constant relative velocity V have their coordinates related by the Lorentz "boost"

$$x' = \gamma(x + Vt), y' = y, z' = z, t' = \gamma(t + Vx)$$

where $\gamma = 1/\sqrt{1 - V^2}$. But more generally the boost could be in any direction specified by a three vector $\vec{V}$. in addition to boost invariance, we also require rotational invariance, a symmetry we are very familiar with in nonrelativistic mechanics. We define a general Lorentz transformation of space-time coordinates $x^\mu = (t, x, y, z) \equiv (x^0, x^1, x^2, x^3)$ as a linear transformation

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

where Λ preserves Minkowski scalar products $v \cdot w = v^1 w^1 + v^2 w^2 + v^3 w^3 - v^0 w^0 \equiv v^\mu v^\nu \eta_{\mu\nu}$. ($\eta_{11} = \eta_{22} = \eta_{33} = -\eta_{00} = 1$, $\eta_{\mu\nu} = 0$ for $\mu \neq \nu$. NOte that we use the convention that repeated indices are summed. This requirement implies the following constraints on Λ:

$$\eta_{\rho\sigma} \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu = \eta_{\mu\nu}.$$

You should convince yourself that the special boost in the x direction does indeed preserve Minkowski scalar products. The Poincaré group consists of Lorentz transformations together with translations

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu + a^\mu.$$

According to special relativity the laws of physics should look the same in all frames related by Lorentz transformations. The systematic way to implement this requirement is to identify all physical quantites as the components of 4 tensors $A^{\mu\nu\cdots}$ which transform as

$$A'^{\mu\nu\cdots} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \cdots A^{\rho\sigma\cdots}$$

For example energy and momentum are the components of a four vector $p^\mu = (E, \vec{p})$ which transform just like the coordinates x^μ.

$$p^\mu \to p'^\mu = \Lambda^\mu{}_\nu p^\nu$$

Tensors written with index superscripts are called contravariant tensors. It is also useful to introduce covariant tensors $A_{\mu\nu\ldots}$, written with index subscripts. A covariant index transforms like $\partial/\partial x^\mu$:

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial}{\partial x^\rho} = \frac{\partial}{\partial x^\rho}(\Lambda^{-1})^\rho{}_\mu = \frac{\partial}{\partial x^\rho}\eta^{\rho\kappa}\Lambda^\sigma{}_\kappa \eta_{\sigma\mu} \equiv \Lambda_\mu{}^\rho \frac{\partial}{\partial x^\rho}$$

A general tensor can have any nunmber of upper and lower indices. the metric tensor is used to raise and lower indices.

The Λ's can be divided into 4 disjoint sets according to the signs of $\det\Lambda$ and $\Lambda^0{}_0$. This is because it is easy to show from the above property that $(\det\Lambda)^2 = 1$ and $(\Lambda^0{}_0)^2 \geq 1$. Thus a continuous variation of Λ always stays within one of these sets. In the following we restrict ourselves to the *proper* Lorentz Group, *i.e.*with $\det\Lambda = +1$ and $\Lambda^0{}_0 \geq +1$. The complete Lorentz group is then obtained by adjoining parity and time reversal.

Lorentz transformations with ${\Lambda^0}_k = {\Lambda^k}_0 = 0$ are simply rotations and form a subgroup. We know from basic quantum mechanics all the unitary irreducible representations of the Rotation group, namely those labelled by angular momentum $j = 0, 1/2, 1, 3/2, \ldots$. This conclusion is reached by considering infinitesimal rotations which are generated by the angular momentum $\vec{J}$, $U(R) = e^{-i\theta\hat{u}\cdot\vec{J}} \approx 1 - i\theta\hat{u}\cdot\vec{J}$. The commutators of J^k are well-known:

$$[J^k, J^l] = i\hbar\epsilon_{klm}J^m$$

and the $|jm\rangle$ basis of eigenstates of $\vec{J}^2$, J_3 is constructed in the familiar way. The unitary representations of the Lorentz group must be extensions of these, $U(\Lambda) = U(R)$ when ${\Lambda^0}_k = {\Lambda^k}_0 = 0$. To find the generators of boosts, consider the boost in the x direction of the momentum of a particle of mass m:

$$p'^1 = \gamma(p^1 + V\sqrt{\vec{p}^2 + m^2}) \approx p^1 + V\sqrt{\vec{p}^2 + m^2}$$

for infinitesimal V. This should be identified with $p^1 - iV[p^1, K^1]$, from which we infer

$$K^1 = -(x^1\sqrt{\vec{p}^2 + m^2} + \sqrt{\vec{p}^2 + m^2}x^1)/2 + F(\vec{p})$$

. The symmetrized product in the first term is to keep K hermitian. To find F we consider the transform of the coordinate

$$x'^1(t') \approx x^1(t) + Vt \approx x^1(t') - V(x^1\dot{x}^1(t) + \dot{x}^1(t)x^1)/2 + Vt$$

But

$$-iV[x^1, K^1] = -\frac{1}{2}\left(x^1\frac{p^1}{\sqrt{\vec{p}^2 + m^2}} + \frac{p^1}{\sqrt{\vec{p}^2 + m^2}}x^1\right) + \frac{\partial F}{\partial p^1} = -\frac{1}{2}\left(x^1\dot{x}^1 + \dot{x}^1x^1\right) + \frac{\partial F}{\partial p^1}$$

Thus we are led to

$$K^1 = -(x^1\sqrt{\vec{p}^2 + m^2} + \sqrt{\vec{p}^2 + m^2}x^1)/2 + p^1t$$

We easily see that K^1 is constant in time. Boosts in all three directions are generated by

$$K^k = -(x^k\sqrt{\vec{p}^2 + m^2} + \sqrt{\vec{p}^2 + m^2}x^k)/2 + p^kt$$

It is now a straightforward exercise to complete the commutator algebra of Lorentz generators:

$$[J^k, J^l] = i\hbar\epsilon_{klm}J^m, \qquad [K^k, J^l] = i\hbar\epsilon_{klm}K^m, \qquad [K^k, K^l] = -i\hbar\epsilon_{klm}J^m$$

An alternative more covariant presentation of the Lorentz algebra is to define $M_{ij} = \epsilon_{ijk}J^k$, $M_{0i} = K^i$, and find

$$[M_{\mu\nu}, M_{\rho\sigma}] = i\left(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\nu\rho}M_{\mu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} + \eta_{\nu\sigma}M_{\mu\rho}\right)$$

where now and henceforth we will choose units so that $\hbar = c = 1$.

In classical electomagnetic theory we learn that Maxwell's equations are covariant under Lorentz transformations of the electromagnetic field:

$$F'^{\mu\nu}(x') = {\Lambda^\mu}_\rho{\Lambda^\nu}_\sigma F^{\rho\sigma}(x) = {\Lambda^\mu}_\rho{\Lambda^\nu}_\sigma F^{\rho\sigma}(\Lambda^{-1}x')$$

Since $F^{\mu\nu}(x)$ is a field with a different value at each point, notice that two changes are going on simultaneously: the components of F are mixed *and* the old and new fields are compared at different space-time

coordinates x', x related by $x' = \Lambda x$. We are going to consider at the beginning a scalar field $\phi(x)$ with only one component. The scalar field then has the much simmpler Lorentz transformation

$$\phi'(x') = \phi(\lambda^{-1}x')$$

The simplest Lorentz covariant scalar field equation is the Klein-Gordon equation

$$-\eta^{\mu\nu}\frac{\partial^2}{\partial x^\mu \partial x^\nu}\phi + \mu^2\phi = 0$$

To show Lorentz covariance we want to prove

$$-\eta^{\mu\nu}\frac{\partial^2}{\partial x'^\mu \partial x'^\nu}\phi(\Lambda^{-1}x') + \mu^2\phi(\Lambda^{-1}x') = 0.$$

this follows through the chain rule

$$\eta^{\mu\nu}\frac{\partial^2}{\partial x'^\mu \partial x'^\nu} = \eta^{\mu\nu}(\Lambda^{-1})^\rho{}_\mu(\Lambda^{-1})^\sigma{}_\nu\frac{\partial^2}{\partial x^\rho \partial x^\sigma} = \eta^{\rho\sigma}\frac{\partial^2}{\partial x^\rho \partial x^\sigma}$$

3. Scalar Quantum Field Thoery

3.1. The Free Scalar Field

The field concept goes all the way back to Faraday in the first part of the nineteenth century. It was of course indispensable to a proper understanding of electromagnetism which he pioneered. This first field theory is actually quite complex involving three components for each of the electric and magnetic fields. For our first look at quantum field theory we consider the much simpler case of the single component scalar field.

A field is a dynamical variable $\phi(\mathbf{x}, t)$ assigned to each point of space. One is therefore positing from the beginning an infinite number of degrees of freedom. The simplest relativistic equation of motion for such a field is a linear wave equation (henceforth we shall use units in which the speed of light $c = 1$):

$$(\frac{\partial^2}{\partial t^2} - \nabla^2 + \mu^2)\phi \equiv (-\partial^2 + \mu^2)\phi = 0.$$

The parameter μ clearly has dimensions of 1/Length, and represents an inverse wavelength (or wave number) rather than a mass. After quantization, we shall see that the particles associated with the quantum field have mass $\hbar\mu$. We are of course very familiar with the nature of the solutions to these equations: general superpositions of plane waves, called wave packets:

$$\phi(\mathbf{x}, t) = \int d^3k f(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x} - i\omega(\mathbf{k})t} \approx \int d^3k f(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x} - i\omega(\mathbf{k}_0)t - i(\mathbf{k}-\mathbf{k}_0)\cdot\nabla\mathbf{k}\omega|_{\mathbf{k}0}} \approx e^{-i\omega(\mathbf{k}_0)t + i\mathbf{k}_0\cdot\mathbf{v}_g t}\phi(\mathbf{x} - \mathbf{v}_g t, 0)$$

where $\omega(\mathbf{k}) = \sqrt{\mathbf{k}^2 + \mu^2}$, moving with group velocity $v_g = \mathbf{k}/\sqrt{\mathbf{k}^2 + \mu^2}$.

Notice that the wave equation is invariant in form under Lorentz transformations $\phi'(x') = \phi(x)$ where $x'^\mu = \Lambda^\mu_\nu x^\nu$ and Λ is a Lorentz transformation satisfying $\eta_{\rho\sigma}\Lambda^\rho_\mu\Lambda^\sigma_\nu = \eta_{\mu\nu}$. This condition on Λ is just the requirement that scalar products of four vectors $v \cdot w \equiv \eta_{\mu\nu}v^\mu w^\nu = \vec{v}\cdot\vec{w} - v^0w^0$ are invariant. One easily can see that $\partial/\partial x_\mu$ transforms as a four-vector so that $\partial^2 = \partial^\mu\partial^\nu\eta_{\mu\nu}$ is invariant.

To discuss the quantum mechanics of this field we must identify the canonical variables of the system. Since the equation of motion is of second order we can easily guess that the momentum conjugate to $\phi(\mathbf{x})$ is just $\pi(\mathbf{x}) = \dot{\phi}$. A more systematic approach is to find an action principle which implies the field equation. Recall that for dynamical systems with discretely labelled degrees of freedom $q_i(t)$ one forms the action from the Lagrangian $L(q(t), \dot{q}(t), t)$ via $S = \int_{t_1}^{t_2} dt L$. Hamilton's principle states that the equations of motion follow from the requirement that S is stationary under veriations $\delta q_i(t)$ with $\delta q_i(t_1) = \delta q_i(t_2) = 0$. By definition the momentum conjugate to q_i is $p_i \equiv \partial L/\partial \dot{q}_i$. For our field the spatial coordinates $\mathbf{x}$ play the role of the index i, and the Lagrangian will be an integral over spatial coordinates. It is easy to see that Hamilton's principle applied to the action

$$S = \int_{t_1}^{t_2} dt \int d^3x \frac{1}{2}(\dot{\phi}^2 - (\nabla\phi)^2 - \mu^2\phi^2) \equiv \int d^4x \mathcal{L}$$

$$\mathcal{L} = -\frac{1}{2}[\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + \mu^2\phi^2]$$

implies the Klein-Gordon equation. The second form shows that the Lagrange density $\mathcal{L}$ is a Lorentz covariant scalar field, so its integral is a Lorentz invariant.

From this action, it is evident that the above guess for $\pi \equiv \partial\mathcal{L}/\partial\dot{\phi} = \dot{\phi}$ is correct. The canonical Hamiltonian is

$$H \equiv \int d^3x[\dot{\phi}\pi - \mathcal{L}] = \int d^3x \frac{1}{2}(\pi^2 + (\nabla\phi)^2 + \mu^2\phi^2).$$

One can easily check that Hamilton's equations obtained from this Hamiltonian give that same old wave equation. To quantize canonically, we promote ϕ, π to operators with the canonical commutation relations

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\hbar\delta(\mathbf{x} - \mathbf{y}).$$

In the language of Hamiltonian mechanics with canonical variables q_i, p_i, ϕ is analogus to q amd π to p, and the spatial coordinate to the index i that distinguishes independent degrees of freedom.

To find the energy eigenstates, we note the close resemblance of our Hamiltonian to that of a system of **coupled harmonic oscillators**. The only thing a little strange is that the oscillator coordinates are labelled by a continuous index. This can be remedied by replacing space by a lattice $a\mathbf{n}$ of spacing a which we send to zero after finding the eigenstates. Thus we replace $\phi(a\mathbf{n})$ with $\phi_{\mathbf{n}}$ and $\pi(a\mathbf{n})$ by $\pi_{\mathbf{n}}/a^3$. We divide by a^3 so that $[\phi_{\mathbf{n}}, \pi_{\mathbf{m}}] = i\hbar\delta_{\mathbf{n},\mathbf{m}}$ Doing this the Hamiltonian becomes

$$H_{lattice} = \frac{a^3}{2}\sum_{\mathbf{n}}\left(\frac{1}{a^6}\pi_{\mathbf{n}}^2 + \mu^2\phi_{\mathbf{n}}^2 + \frac{1}{a^2}\sum_{\hat{i}}(\phi_{\mathbf{n}+\hat{\mathbf{i}}} - \phi_{\mathbf{n}})^2\right).$$

On the lattice our system *is* a coupled system of oscillators. To solve it, we just have to find the normal modes. This is done by a change of variables which renders diagonal the coupling matrix

$$V_{\mathbf{n},\mathbf{m}} = (\mu^2 + \frac{2d}{a^2})\delta_{\mathbf{n},\mathbf{m}} - \frac{1}{a^2}\sum_{\hat{i}}(\delta_{\mathbf{n}+\hat{\mathbf{i}},\mathbf{m}} + \delta_{\mathbf{n},\mathbf{m}+\hat{\mathbf{i}}}),$$

i.e. we want to find the eigenvectors and eigenvalues of V. It is easy to see that the eigenvectors are

$u_{\mathbf{m}} = e^{i\mathbf{K}\cdot\mathbf{m}}$, $-\pi < K_i < \pi$, belonging to eigenvalue

$$\lambda \equiv \omega^2 = \mu^2 + \frac{2}{a^2}(d - \sum_{\hat{i}} \cos \mathbf{K}\cdot\hat{\mathbf{i}}).$$

Thus the transformation to normal modes is given by

$$\phi_{\mathbf{n}} = \int_{-\pi}^{\pi} \frac{d^3K}{(2\pi)^{3/2}} e^{i\mathbf{K}\cdot\mathbf{n}} Q(\mathbf{K})$$
$$\pi_{\mathbf{n}} = \int_{-\pi}^{\pi} \frac{d^3K}{(2\pi)^{3/2}} e^{i\mathbf{K}\cdot\mathbf{n}} P(\mathbf{K})$$

or inversely by

$$Q(\mathbf{K}) = \sum_{\mathbf{n}} \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{K}\cdot\mathbf{n}} \phi_{\mathbf{n}}$$
$$P(\mathbf{K}) = \sum_{\mathbf{n}} \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{K}\cdot\mathbf{n}} \pi_{\mathbf{n}}.$$

The inverse equations directly imply the commutation relations

$$[Q(\mathbf{K}), P(\mathbf{L})] = i\hbar\delta(\mathbf{K}+\mathbf{L}).$$

Note also that if ϕ and π are hermitian, as we assume, then we have $Q(\mathbf{K})^\dagger = Q(-\mathbf{K})$ and $P(\mathbf{K})^\dagger = P(-\mathbf{K})$.

Expressed in terms of normal modes the lattice Hamiltonian becomes

$$H_{lattice} = \int_{-\pi}^{\pi} d^3K \left(\frac{1}{2a^3} P(\mathbf{K})P(-\mathbf{K}) + \frac{a^3}{2}\omega^2(\mathbf{K})Q(\mathbf{K})Q(-\mathbf{K}) \right).$$

In the standard fashion raising and lowering operators can be constructed in terms of which

$$Q(\mathbf{K}) = \sqrt{\frac{\hbar}{2\omega(\mathbf{K})a^3}}(A(\mathbf{K}) + A^\dagger(-\mathbf{K}))$$
$$P(\mathbf{K}) = -i\sqrt{\frac{\hbar\omega(\mathbf{K})a^3}{2}}(A(\mathbf{K}) - A^\dagger(-\mathbf{K}))$$

with $[A(\mathbf{K}), A^\dagger(\mathbf{L})] = \delta(\mathbf{K}-\mathbf{L})$, and the Hamiltonian becomes the familiar

$$H_{lattice} = \int_{-\pi}^{\pi} d^3K \hbar\omega(\mathbf{K}) \frac{1}{2}(A(\mathbf{K})A^\dagger(\mathbf{K}) + A^\dagger(\mathbf{K})A(\mathbf{K})).$$

$A^\dagger$ and A are of course eigenoperators of $H_{Lattice}$ with eigenvalues $\pm\omega(\mathbf{K})$.

Now we can describe the exact energy eigenstates of the system. The ground state (the vacuum) $|0\rangle$ is annihilated by all the A's: $A(\mathbf{K})|0\rangle = 0$, and its energy $E_0 = (\hbar/2)\delta(\mathbf{0})\int d^3K\omega(\mathbf{K})$. We shall measure all

energies relative to E_0, *i.e.* from now on we take $E_0 = 0$. This means our energy operator is redefined to be

$$H - E_0 = \int_{-\pi}^{\pi} d^3K \hbar\omega(\mathbf{K}) A^\dagger(\mathbf{K}) A(\mathbf{K})$$

The excited states are obtained by applying any number of $A^\dagger$'s to the vacuum. The simplest one is $A^\dagger(\mathbf{K})\,|0\rangle$ with energy

$$E(\mathbf{K}) = \hbar\sqrt{\mu^2 + \frac{2}{a^2}(d - \sum_{\hat{i}} \cos \mathbf{K}\cdot\hat{\mathbf{i}})}.$$

At this point we can consider the continuum limit $a \to 0$. All excited states with finite $\mathbf{K}$ will have infinite energy and be dynamically irrelevant. However since the components of $\mathbf{K}$ are continuous, we may consider states with $\mathbf{K} = a\mathbf{k}$. The energy of these states becomes $\hbar\sqrt{\mu^2 + k^2}$ independent of a in the continuum limit. We identify this energy as that of a relativistic particle of mass $\hbar\mu$ and momentum $\hbar\mathbf{k}$.

We can pass to the continuum limit in all of our results by changing variables from $\mathbf{K}$ to $\mathbf{k}$ whose components range from $-\pi/a$ to π/a which become $-\infty$ to ∞ as $a \to 0$. At the same time we replace $A(\mathbf{K}) \equiv a(\mathbf{k})/a^{3/2}$ so that $[a(\mathbf{k}), a^\dagger(\mathbf{q})] = \delta(\mathbf{k} - \mathbf{q})$. After these changes we find as $a \to 0$:

$$\begin{aligned}
\phi_{\mathbf{n}} \to \phi(\mathbf{x}) &= \int \frac{d^3k}{(2\pi)^{3/2}} \sqrt{\frac{\hbar}{2\omega(\mathbf{k})}} (a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}) \\
\frac{\pi_{\mathbf{n}}}{a^3} \to \pi(\mathbf{x}) &= -\,i \int \frac{d^3k}{(2\pi)^{3/2}} \sqrt{\frac{\hbar\omega(\mathbf{k})}{2}} (a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} - a^\dagger(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}) \\
H - E_0 \to H_\phi - E_0 &= \int d^3k \hbar\omega(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k})
\end{aligned}$$

where now $\omega(\mathbf{k}) = \sqrt{\mu^2 + \mathbf{k}^2}$. It should of course be clear that we really never needed the lattice in obtaining these results: we simply had to substitute the above expansions for ϕ, π into H_ϕ. The lattice only served to make absolutely concrete our assertion that the field system was a set of coupled oscillators.

We should also note that the momentum operator of the continuum field theory, identified as the generator of translations, is

$$\mathbf{P} = -\int d^3x \pi(\mathbf{x}) \nabla\phi(\mathbf{x}) = \int d^3k \hbar\mathbf{k} a^\dagger(\mathbf{k}) a(\mathbf{k})$$

confirming the interpretation of $\hbar\mathbf{k}$ as the momentum of the one particle state $a^\dagger(\mathbf{k})\,|0\rangle$.

If we apply two $a^\dagger$'s to the vacuum we get a two particle state

$$a^\dagger(\mathbf{k}_1) a^\dagger(\mathbf{k}_2)\,|0\rangle$$

with total momentum $\hbar(\mathbf{k}_1 + \mathbf{k}_2)$ and total energy $\hbar(\omega(\mathbf{k}_1) + \omega(\mathbf{k}_2))$. Clearly this energetics is that of noninteracting particles. It is highly significant that all multiparticle states are completely symmetric under interchange of the labels of any pair of particles: the scalar field theory predicts Bose statistics for the associated particles!

The example of the scalar field illustrates the main physical aspects of quantum field theory

1. It predicts multi-particle states together with their statistics (Bose or Fermi).
2. It incorporates the requirements of Special Relativity (Poincaré invariance).

3. One can consider two classical limits ($\hbar \to 0$):
   (a) With μ fixed the limit gives a classical field theory. In this case the Compton wavelength $1/\mu$ is finite in the classical limit.
   (b) With $m = \hbar\mu$ fixed it gives a classical theory of point particles (the Compton wavelength vanishes).
   (c) Thus quantum field theory unites the particle and field concepts.
4. Quantum field theories with linear field equations predict noninteracting (free) particles. Interactions will arise if the field equations have nonlinear terms. Such terms are associated with terms in the action or Hamiltonian with three or more powers of fields. The presence of such terms implies that particle number is not conserved.
5. The scalar field describes spinless particles; more general fields *e.g.* spinor, vector, tensor describe particles of higher spin.

Finally, let us consider how an approximate quantum particle interpretation can be retrieved from this quantum field theory. Remember the observables are quantum fields, not particle coördinates! To discover properties of a single particle wave packet,

$$|f\rangle \equiv \int d^3k f(\mathbf{k}) a^\dagger(\mathbf{k}) |0\rangle ,$$

we must consider what we can get from measurements of the quantum field

$$\phi(\mathbf{x},t) = \int \frac{d^3k}{(2\pi)^{3/2}} \sqrt{\frac{\hbar}{2\omega(\mathbf{k})}} (a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} + a^\dagger(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x} + i\omega t}).$$

First it is easily shown that $\langle f| \phi(\mathbf{x},t) |f\rangle = 0$. That is, if we make many measurements of ϕ, we find zero on average. This means nothing more than ϕ is negative as often as it is positive. A more sophisticated measurement is to measure $\phi(\mathbf{x},t)\phi(\mathbf{y},t)$, *i.e.* the product of the results of simultaneous measurement of the fields at $\mathbf{x}$ and at $\mathbf{y}$. If no particle is present, the average of many repeats of this measurement is

$$\langle 0| \phi(\mathbf{x},t)\phi(\mathbf{y},t) |0\rangle = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} \frac{\hbar}{\omega(\mathbf{k})} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}.$$

Note that due to the $1/\omega(\mathbf{k})$, this is not 0 even when the fields are not measured at the same point. This implies a correlation between the measurements at separate points, but does not contradict causality: $[\phi(\mathbf{x},t), \phi(\mathbf{y},t)] = 0$! The vacuum is an energy eigenstate which requires an infinite time to set up. Thus there is plenty of time to set up correlations at distant points.

Now, suppose a particle is present in the wave packet f, and the same quantity is measured.

$$\langle f| \phi(\mathbf{x},t)\phi(\mathbf{y},t) |f\rangle = \langle 0| \phi(\mathbf{x},t)\phi(\mathbf{y},t) |0\rangle + \psi^*(\mathbf{x},t)\psi(\mathbf{y},t) + \psi^*(\mathbf{y},t)\psi(\mathbf{x},t).$$

The presence of the particle causes the change in average results given by the last 2 terms on the r.h.s. which are modulated by the function

$$\begin{aligned}\psi(\mathbf{x},t) \equiv & \int \frac{d^3k}{(2\pi)^{3/2}} \sqrt{\frac{\hbar}{2\omega(\mathbf{k})}} e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} f(\mathbf{k}) \\ \approx & \psi(\mathbf{x} - \mathbf{v}_g t, 0) \\ \approx & \sqrt{\frac{\hbar}{2\omega(\mathbf{k_0})}} \psi_S(\mathbf{x} - \mathbf{v}_g t, 0).\end{aligned}$$

The last two forms hold if the packet function f is narrowly peaked about $\mathbf{k_0}$. ψ_S is the one particle

Schrödinger wave function corresponding to momentum wave function f. $\mathbf{v}_g \equiv \partial\omega/\partial\mathbf{k}|_{\mathbf{k}=\mathbf{k_0}}$ is the usual group velocity. The disturbance is nonzero only when the point is in the support of the wave function.

3.2. Interacting Scalar Field Theory

So far we have only considered linear field equations, whose associated Lagrangians are quadratic functions of the the field. As soon as the field equations become non-linear, the particles associated with the quantum fields will ineract with each other. In particular they can scatter.

The simplest way to introduce non-linear terms in the field equations, while preserving Lorentz invariance, is to add Lorentz scalar terms cubic and higher in the fields to the Lagrangian density. for scalar fields this is easy to do. For example any term

$$-\frac{g_n}{n!}\phi^n(x), \qquad\qquad \text{for } n > 2$$

is a scalar which produces a term $g_n\phi^{n-1}/(n-1)!$ in the field equation. Since such terms do not involve time derivatives, they do not alter the definition of the conjugate momentum nor the commutation relations. But clearly plane waves no longer solve the new field equations so there will be interactions.

Nonlinear quantum field theory is too complicated to solve exactly, so we shall first assume the nonlinear terms are small so they can be handled using perturbation theory. In the next chapter we review the formalism of time dependent perturbation theory.

4. Time Dependent Perturbation Theory

Although there are special quantum field theories, for example some in 2 space-time dimensions, that can be solved exactly, the exact solution of realistic interacting quantum field theories in 4 space-time dimensions is beyond reach. There are important cases for which linear QFT's in the presence of certain external fields can be solved. For example the Dirac equation in a Coulomb potential admits an exact solution for which one can find all the energy eigenvalues and eigenstates. The Coulomb potential is of special importance since the exact solution in that case is the starting point for the relativistic theory of atomic energy levels. Still the complete dynamics is never exactly given by these special cases and perturbation theory is the important tool for evaluating corrections to the exactly soluble (idealized) case, which can be free field theory with no external fields or one of the above cases. When we quantize the electromagnetic field, perturbation theory is essentially our *only* tool for computing radiative corrections due to the quantum nature of the electromagnetic field.

4.1. Heisenberg and Schrödinger Pictures

In the Schrödinger picture the quantum dynamics is given by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\left|\Phi,t\right\rangle = H_S(t)\left|\Phi,t\right\rangle$$

where we stress that we allow time varying external forces to be present (hence the time dependence of H). The field equations arise in the Heisenberg picture wherein the time dependence resides in the operators rather than in the system states which are constant in time. To pass to the Heisenberg picture we write

$$\left|\Phi,t\right\rangle = U(t)\left|\Phi,0\right\rangle$$

where

$$i\hbar\frac{\partial}{\partial t}U = H_S(t)U \qquad\qquad U(0) = I,$$

and give the time independent Schrödinger picture operators Ω time dependence according to

$$\Omega(t) \equiv U^\dagger \Omega U.$$

The Heisenberg picture Hamiltonian is similarly related to the Schrödinger picture one by

$$H(t) \equiv U^\dagger H_S(t) U.$$

Then the Heisenberg picture operators corresponding to constant Schrödinger picture operators satisfy the Heisenberg equations

$$i\hbar\dot{\Omega}(t) = [\Omega(t), H(t)].$$

It is most natural to formulate the time dependence in quantum field theory using Heisenberg picture, since the field operators will then satisfy equations of motion that are the direct quantum analogue of the classical field equations. We shall therefore always understand $H(t)$ without subscripts to be the Hamiltonian in Heisenberg picture. When we woek with any other picture we will atach a subscript to H, e.g. $H_S(t)$ is the Hamiltonian in Schrödinger picture.

4.2. Asymptotic States and Matrix Elements

In discussing time dependent processes, it is convenient to introduce asymptotic states which are eigenstates of $H(\pm\infty)$. We denote by $|in\rangle$ the ground state of $H(-\infty)$ and by $|out\rangle$ the ground state of $H(+\infty)$. The normal situation will be one in which all external fields vanish at sufficiently early and late times. Thus $|in\rangle$ and $|out\rangle$ will typically be ground states of $H_0(-\infty)$ and $H_0(+\infty)$ respectively. Although these operators are not the same (because their time evolution is governed by H not H_0), the spectra of the two Hamiltonians *are* identical: $H_0(t) = U^{-1}(t,-\infty)H_{0S}U(t,-\infty)$. By convention we are identifying the Schrödinger and Heisenberg pictures at $t = -\infty$. Thus, if $|in\rangle$ is the ground state of $H_0(-\infty) = H_{0S}$, the state $\langle in|\, U(\infty,-\infty)$ is an eigenstate of $H_0(+\infty)$ with the same eigenvalue and hence the ground state. Thus we can and shall fix phases by defining

$$\langle out| \equiv \langle in|\, U(\infty,-\infty).$$

We stress that this is the true "out" state only when $H_S(\infty) = H_S(-\infty) \equiv H_0$.

If the time dependence of H_S is *adiabatic*, *i.e.* very slow on the time scale set by the level spacings, the Adiabatic Theorem assures us that an eigenstate of $H_S(-\infty)$ evolves to an eigenstate of $H_S(t)$ for all t for which adiabatic conditions apply, even after a long enough time to change H_S by a finite amount. For example, the state $|in\rangle$ will be an eigenstate of $H(t)$ for all t for which adiabatic time variation applies. In particular, the ground state eigenvalue $E_G(t)$ must not get close to the next higher eigenvalue as t varies. If this situation holds for all time, it follows that the state $|in\rangle$ is a phase times the state $|out\rangle$, or $\langle out|$ is this same phase times $\langle in|$. This phase is easily evaluated in terms of the time dependent ground state energy $E_G(t)$ of $H_S(t)$ by applying the Schrödinger equation to $\langle in|\, U(t,-\infty)\, |in\rangle$ and using the adiabatic theorem $H_S(t)U(t,-\infty)\, |in\rangle = E_G(t)U(t,-\infty)\, |in\rangle$:

$$\langle out|in\rangle = \exp\left\{-i\int_{-\infty}^{\infty} dt E_G(t)\right\} \qquad \text{Adiabatic Conditions.}$$

Note carefully that adiabatic conditions would *not* apply if the ground state energy got close to an excited level as time evolved. In particular, it would not apply in processes with pair production when $|\,\langle out|in\rangle\,| < 1$.

4.3. General Formalism

We shall keep the initial discussion completely general and consider the situation in which the Heisenberg picture Hamiltonian is the sum of two pieces,

$$H(t) = H_0(t) + H'(t)$$

where H_0 can be exactly dealt with and H' is "small" in an appropriate sense. Note that even when H is independent of time, H_0 and H' still depend on time through the time dependence of the Heisenberg operators that enter it. The Heisenberg equations for the dynamical variables have the form

$$i\frac{\partial\Omega}{\partial t} = [\Omega, H(t)].$$

The goal of time dependent perturbation theory is to expand the evolution operator $U(t,t_0)$ which carries the time dependence of the Heisenberg picture operators $\Omega = U^{-1}\Omega_S U$, or alternatively the time dependence of

the Schrödinger picture system states $|\psi, t\rangle = U\,|\psi, t_0\rangle$, in a power series in H'. A complication is that there is time dependence in U even when the perturbation vanishes. To systematically deal with this complication a new Interaction Picture (sometimes called the Dirac Picture) has been devised in which the operators carry the (known) time dependence due to H_0 and the perturbation only enters the modified evolution operator U_I, which is constructed to be the identity in the absence of H'.

Thus for each Heisenberg picture operator with no explicit time dependence, we define an interaction picture operator by

$$\Omega_I(t) \equiv U_I(t)\Omega(t)U_I^{-1}(t) \tag{4.1}$$

and require that Ω_I satisfies the Heisenberg equation with Hamiltonian $H_{0I} = U_I H_0(t) U_I^{-1}$ ⋆:

$$i\frac{\partial \Omega_I}{\partial t} = [\Omega_I, H_{0I}(t)].$$

differentiating (4.1) we find the requirement

$$\begin{aligned}[\Omega_I, H_{0I}(t)] =& i\dot{U}_I\Omega U_I^{-1} + iU_I\Omega\dot{U}_I^{-1} + U_I[\Omega, H(t)]U_I^{-1} \\ =& i\dot{U}_I U_I^{-1}\Omega_I + i\Omega_I U_I\dot{U}_I^{-1} + [\Omega_I, H_{0I}(t)] + [\Omega_I, H_I'(t)] \\ =& [\Omega_I, H_{0I}(t)] + [\Omega_I, H_I'(t) - i\dot{U}_I U_I^{-1}].\end{aligned}$$

Thus the equation for U_I is just

$$i\dot{U}_I(t) = H_I'(t)U_I(t) = U_I(t)H'(t).$$

We choose the initial condition $U_I(t_0) = I$, in which case it is a good idea to display two time arguments $U_I(t, t_0)$ as we did for U. Notice that since U relates Heisenberg and Schrödinger pictures, the equation ( ) for U can be also written

$$i\dot{U} = UH(t) \tag{4.2}$$

from which it is clear that we can express $U = U_0 U_I$ where

$$i\dot{U}_0 = U_0 H_{0I}(t)$$

To expand U_I in powers of H' it is convenient first to incorporate initial condition information by writing the integral equation

$$U_I(t, t_0) = I - i\int_{t_0}^{t} dt' H_I'(t')U_I(t', t_0), \tag{4.3}$$

and then to generate the perturbation series by iteration

$$\begin{aligned}U_I(t, t_0) =& I - i\int_{t_0}^{t} dt' H_I'(t')(I - i\int_{t_0}^{t'} dt'' H_I'(t'')U_I(t'', t_0) \\ =& I - i\int_{t_0}^{t} dt' H_I'(t') + (-i)^2\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H_I'(t')H_I'(t'') + \cdots.\end{aligned}$$

There is a useful way to summarize the entire perturbation series, which employs the concept of the time ordered product of operators. Consider a set of operators each associated with a different time,

⋆ Note that if $H_0 = H_0(\Omega_k(t), t)$, then $H_{0I} = H_0(\Omega_{Ik}(t), t)$.

$A_1(t_1), A_2(t_2), \ldots, A_N(t_N)$. The time ordered product of these operators is defined as the ordinary product with the operators ordered according to the time argument: the operator $A_k(t_k)$ to the left of $A_l(t_l)$ if $t_k > t_l$. If there are any anticommuting operators in the set, there is also an overall -1 if one achieves the time ordering by an odd permutation of fermionic operators. Thus, for example,

$$T[A(t_1)B(t_2)] = \begin{cases} A(t_1)B(t_2) & t_1 > t_2 \\ \pm B(t_2)A(t_1) & t_2 > t_1 \end{cases}$$

with the $-$ for both operators fermionic. Now the factors of H_I' in the series for U_I are all time ordered due to the limits of integration. If we make use of the time ordering symbol, we can extend all integrations to the full range $t_0 < t' < t$ provided we divide the nth term by $n!$ to account for the overcounting due to the $n!$ orderings of the t's. Thus the entire series becomes

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \frac{1}{n!}(-i)^n \int_{t_0}^{t} dt_1 dt_2 \cdots dt_n T[H_I'(t_1)H_I'(t_2)\cdots H_I'(t_n)].$$

If it weren't for the time ordering symbol this would be just the exponential series. It is therefore a useful mnemonic to write

$$U_I(t, t_0) = Te^{-i\int_{t_0}^{t} dt' H_I'(t')} \tag{4.4}$$

where it is understood that $t > t_0$. This equation is known as the Dyson Formula. Since the formula just reflects the equation U_I satisfies, we can write a similar formula for the full U:

$$U(t, t_0) = Te^{-i\int_{t_0}^{t} dt' H_S(t')} = \bar{T}e^{-i\int_{t_0}^{t} dt' H(t')} = \bar{T}e^{-i\int_{t_0}^{t} dt' H_{0,I}(t')} Te^{-i\int_{t_0}^{t} dt' H_I'(t')}$$

where $\bar{T}$ denotes anti-time ordering (later times to the *right*).

It will be useful to extend the definition of $U_I(t, t_0)$ to times earlier than t_0. We shall do this in a way to preserve the *closure* property

$$U_I(t, t_1)U_I(t_1, t_0) = U_I(t, t_0), \tag{4.5}$$

which follows from the differential equation and initial condition for $t > t_1 > t_0$. If we set $t = t_0$ in (4.5), the r.h.s. is just I so we have to define

$$U_I(t, t_0) \equiv U_I^{-1}(t_0, t) = U_I^{\dagger}(t_0, t) \qquad \text{for } \ t < t_0.$$

(Note that $U_I^{\dagger}U_I = I$ is a simple consequence of the differential equation and the hermiticity of H_I'.) It is then simple to check that (4.5) holds for all time orderings.

Next let us show how to express various physical quantities in the Interaction picture. One interesting quantity is the so called vacuum persistence amplitude $\langle out|in\rangle$, given by

$$\langle out|in\rangle = \langle in|\, U(\infty, -\infty)\, |in\rangle = \langle in|\, U_0(\infty, -\infty)U_I(\infty, -\infty)\, |in\rangle\,.$$

We shall identify all pictures at $t = -\infty$. In the usual situation where external fields vanish at early times the state $|in\rangle$ will be the ground state of $H_0(-\infty) = H_{0I}$. Furthermore, with no external fields in $H_{0I}(t)$

the latter will be time independent for all time (since its time evolution is governed by H_{0I} itself). In this situation

$$U_0(t,t_0) = e^{-iH_{0I}(t-t_0)}$$

and $\langle in|$ is an eigenstate of $U_0(t,t_0)$ with eigenvalue $e^{-iE_0(t-t_0)}$, where E_0 is the ground state energy of H_{0I}. By **convention** we can choose our zero of energy so that $E_0 = 0$, in which case we have

$$\langle out|in\rangle = \langle 0,I|\, U_I(\infty,-\infty)\,|0,I\rangle \qquad\qquad E_0 = 0.$$

The persistence amplitude carries a lot of information, because it can be defined for any choice of external fields. Its dependence on these external fields can then be exploited to obtain numerous matrix elements relevant to the zero field situation. We shall see many applications of this remark in the course of our studies. One can also get the energy spectrum for static external fields from this amplitude by switching them on at some early time keeping them constant for a long time $2T$ and then switching them off. The T dependence of $\langle out|in\rangle$ will then display the dependence $e^{-iE_k 2T}$ from which the energy eigenvalues can be read off. The states that are probed by this device will depend on the manner of the switching on procedure. For adiabatic switching on, only the ground state in the presence of the static field will contribute.

4.4. Scattering in an External Field: Born Approximation

One can equally well choose initial and final states that contain particles. For example, the amplitude for a scalar particle with momentum $\mathbf{p}$ initially making a transition to $\mathbf{p}'$ at very late times is

$$\langle 0,I|\, a(\mathbf{p}')U_I(\infty,-\infty)a^\dagger(\mathbf{p})\,|0,I\rangle \approx \\ \delta(\mathbf{p}'-\mathbf{p}) - i\int_{-\infty}^{\infty} dt\,\langle 0,I|\, a(\mathbf{p}')H_I'(t)a^\dagger(\mathbf{p})\,|0,I\rangle\,,$$

where we kept only terms to first order. As a concrete example, consider the interaction $\int dt H_I'(t) = -\int d^4x\phi_I^2(x)B(x)/2$, where $B(x)$ is a fixed external scalar field. Since the ϕ_I are free fields, they can be expressed in terms of creation and annihilation operators and the matrix element evaluated:

$$\begin{aligned}\int_{-\infty}^{\infty} dt\,\langle 0,I|\, a(\mathbf{p}')H_I'(t)a^\dagger(\mathbf{p})\,|0,I\rangle \\ &= -\int d^4x B(x)e^{i(p-p')\cdot x}\frac{1}{(2\pi)^3\sqrt{2\omega(\mathbf{p})}\sqrt{2\omega(\mathbf{p}')}} \\ &= -\frac{1}{(2\pi)^3\sqrt{2\omega(\mathbf{p})}\sqrt{2\omega(\mathbf{p}')}}\tilde{B}(p'-p),\end{aligned}$$

where $\tilde{B}(q) = \int d^4x B(x)e^{-iq\cdot x}$ is the Fourier transform of the external field. Note that in the case B is static the time integral gives a factor of $2\pi\delta(\omega'-\omega)$.

Recall from basic scattering theory that if the scattering matrix for a particle from a static potential is

written

$$\langle \mathbf{q}, out|\mathbf{p}, in\rangle = \delta(\mathbf{q}-\mathbf{p}) - 2\pi i\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))T(\mathbf{q},\mathbf{p})$$

then the differential scattering cross section is given by

$$\begin{aligned}\frac{d\sigma}{d\Omega} =& \frac{d^3q}{d\Omega}\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))\frac{(2\pi)^4}{v}|T(\mathbf{q},\mathbf{p})|^2,\\ =& q^2dq\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))\frac{(2\pi)^4}{v}|T(\mathbf{q},\mathbf{p})|^2,\\ =& p\omega(\mathbf{p})\frac{(2\pi)^4}{v}|T(\mathbf{q},\mathbf{p})|^2,\\ =& \omega(\mathbf{p})^2(2\pi)^4|T(\mathbf{q},\mathbf{p})|^2,\end{aligned}$$

where v is the speed of the incident particle. Defining the *spatial* Fourier transform $\tilde{B}(\mathbf{k}) \equiv \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}}B(\mathbf{x})$, we then obtain

$$T^{Born}(\mathbf{q},\mathbf{p}) = -\frac{1}{(2\pi)^3 2\omega(\mathbf{p})}\tilde{B}(\mathbf{q}-\mathbf{p}),$$

giving the cross section

$$\frac{d\sigma}{d\Omega}^{Born} = \frac{1}{16\pi^2}|\tilde{B}(\mathbf{q}-\mathbf{p})|^2.$$

Another scattering example, the amplitude for a Dirac particle with momentum and helicity $\mathbf{p}, \lambda$ initially making a transition to $\mathbf{p}', \lambda'$ at very late times is

$$\begin{aligned}&\langle 0, I|\, b_{\lambda'}(\mathbf{p}')U_I(\infty,-\infty)b^\dagger_\lambda(\mathbf{p})\,|0,I\rangle \approx\\ &\qquad \delta_{\lambda'\lambda}\delta(\mathbf{p}'-\mathbf{p}) - i\int_{-\infty}^{\infty} dt\,\langle 0, I|\, b_{\lambda'}(\mathbf{p}')H'_I(t)b^\dagger_\lambda(\mathbf{p})\,|0,I\rangle\,,\end{aligned}$$

where we kept only terms to first order. In the case of a weak external electromagnetic field, $\int dtH'_I(t) = -\int d^4x j^\mu_I(x)A_\mu(x)$, where $j^\mu_I = \frac{q}{2}[\bar{\psi}_I, \gamma^\mu\psi_I]$. Since the ψ_I are free fields, they can be expressed in terms of creation and annihilation operators and the matrix element evaluated:

$$\begin{aligned}&\int_{-\infty}^{\infty} dt\,\langle 0, I|\, b_{\lambda'}(\mathbf{p}')H'_I(t)b^\dagger_\lambda(\mathbf{p})\,|0,I\rangle\\ &\qquad = -q\int d^4x A_\mu(x)e^{i(p-p')\cdot x}\frac{1}{(2\pi)^3\sqrt{2\omega(\mathbf{p})}\sqrt{2\omega(\mathbf{p}')}}\bar{u}_{\lambda'}(\mathbf{p}')\gamma^\mu u_\lambda(\mathbf{p})\\ &\qquad = -q\frac{1}{(2\pi)^3\sqrt{2\omega(\mathbf{p})}\sqrt{2\omega(\mathbf{p}')}}\bar{u}_{\lambda'}(\mathbf{p}')\gamma^\mu\tilde{A}_\mu(p'-p)u_\lambda(\mathbf{p}),\end{aligned}$$

where $\tilde{A}(q) = \int d^4x A(x)e^{-iq\cdot x}$ is the Fourier transform of the potential. Note that in the case A is static the time integral gives a factor of $2\pi\delta(\omega'-\omega)$.

Recall from basic scattering theory that if the scattering matrix for a particle from a static potential is written

$$\langle \mathbf{q}, out|\mathbf{p}, in\rangle = \delta_{\lambda'\lambda}\delta(\mathbf{q}-\mathbf{p}) - 2\pi i\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})$$

then the differential scattering cross section is given by

$$\begin{aligned}\frac{d\sigma}{d\Omega} =& \frac{d^3q}{d\Omega}\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))\frac{(2\pi)^4}{v}|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\\ =& q^2dq\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))\frac{(2\pi)^4}{v}|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\\ =& p\omega(\mathbf{p})\frac{(2\pi)^4}{v}|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\\ =& \omega(\mathbf{p})^2(2\pi)^4|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\end{aligned}$$

where v is the speed of the incident particle.

Defining $\tilde{A}(\mathbf{k}) \equiv \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}}A(\mathbf{x})$, we then obtain

$$T^{Born}_{\lambda'\lambda}(\mathbf{q},\mathbf{p}) = -\frac{1}{(2\pi)^3 2\omega(\mathbf{p})}\bar{u}_{\lambda'}q\gamma\cdot\tilde{A}(\mathbf{q}-\mathbf{p})u_\lambda(\mathbf{p}),$$

giving the cross section

$$\frac{d\sigma}{d\Omega}^{Born} = \frac{1}{16\pi^2}|\bar{u}_{\lambda'}q\gamma\cdot\tilde{A}(\mathbf{q}-\mathbf{p})u_\lambda(\mathbf{p})|^2.$$

Consider the example of the scattering of an electron with $q=-e$ from the Coulomb potential of a nucleus of atomic number Z (Mott Scattering), $A^0 = Ze/4\pi r$, $\mathbf{A}=0$. Then $\tilde{A}^0(\mathbf{k}) = Ze/k^2$ and

$$\frac{d\sigma}{d\Omega}^{Born} = \frac{e^4Z^2}{16\pi^2(\mathbf{q}-\mathbf{p})^4}|\bar{u}_{\lambda'}(\mathbf{q})\gamma^0u_\lambda(\mathbf{p})|^2 = \frac{\alpha^2Z^2}{(\mathbf{q}-\mathbf{p})^4}|\bar{u}_{\lambda'}(\mathbf{q})\gamma^0u_\lambda(\mathbf{p})|^2.$$

Here we have introduced the fine structure constant $\alpha = e^2/4\pi \approx 1/137$. Evaluating the spinor matrix element in terms of two component helicity spinors leads to (using $|\mathbf{q}|=|\mathbf{p}|$)

$$\bar{u}_{\lambda'}(\mathbf{q})\gamma^0u_\lambda(\mathbf{p}) = \left[\omega(\mathbf{p})+m+\frac{4\lambda\lambda'\mathbf{p}^2}{\omega+m}\right]\chi^\dagger_{\lambda'}(\mathbf{q})\chi_\lambda(\mathbf{p}).$$

The absolute square of $\chi'^\dagger\chi$ can be evaluating by noting that the 2×2 matrix $\chi\chi^\dagger$ is a projector onto the spin state of definite helicity:

$$\chi_\lambda(\mathbf{p})\chi^\dagger_\lambda(\mathbf{p}) = \frac{1+2\lambda\hat{\mathbf{p}}\cdot\boldsymbol{\sigma}}{2}.$$

Thus we have

$$\begin{aligned}|\chi^\dagger_{\lambda'}(\mathbf{q})\chi_\lambda(\mathbf{p})|^2 =& \frac{1}{4}\,\mathrm{tr}[(1+2\lambda\hat{\mathbf{p}}\cdot\boldsymbol{\sigma})(1+2\lambda'\hat{\mathbf{q}}\cdot\boldsymbol{\sigma})]\\ =& \frac{1}{2}(1+4\lambda\lambda'\hat{\mathbf{p}}\cdot\hat{\mathbf{q}})\end{aligned}$$

Inserting all this into the formula for the differential cross section, we obtain after simplifying

$$\frac{d\sigma}{d\Omega}^{Born} = \frac{\alpha^2Z^2}{(\mathbf{q}-\mathbf{p})^4}[\omega^2+m^2+\mathbf{q}\cdot\mathbf{p}+4\lambda\lambda'(\mathbf{p}^2+(\omega^2+m^2)\hat{\mathbf{p}}\cdot\hat{\mathbf{q}})]$$

To compare all of the details of this formula with experiment we would have to prepare a polarized beam of electrons with definite helicity and also measure the spin of the final electron. A noteworthy feature of such a complete experiment is that at high energies there is an overall factor of $(1+4\lambda\lambda') = 2\delta_{\lambda\lambda'}$, which means

that helicity is conserved at high energy. If we don't measure the final spin we should sum over $\lambda' = \pm 1/2$ to obtain

$$\frac{d\sigma}{d\Omega}^{Born}_{Unobserved\ spin} = 2\frac{\alpha^2 Z^2}{(\mathbf{q}-\mathbf{p})^4}[\omega^2 + m^2 + \mathbf{q}\cdot\mathbf{p}].$$

Similarly, if we have a completely unpolarized beam, we need to average over λ to obtain

$$\frac{d\sigma}{d\Omega}^{Born}_{unpol} = \frac{\alpha^2 Z^2}{(\mathbf{q}-\mathbf{p})^4}[\omega^2 + m^2 + \mathbf{q}\cdot\mathbf{p}],$$

independent of the final spin.

Two simplifying limits can be considered. The *nonrelativistic* or *low energy* limit $\mathbf{p}^2 << m^2$ is

$$\frac{d\sigma}{d\Omega}^{NR} \sim 2m^2\frac{\alpha^2 Z^2}{(\mathbf{q}-\mathbf{p})^4}[1 + 4\lambda\lambda'\hat{\mathbf{p}}\cdot\hat{\mathbf{q}}]$$

Apart from the helicity dependence due to the spin of the electrons this is just the Rutherford formula. The opposite limit, the *ultrarelativistic* or *high energy* limit $\mathbf{p}^2 >> m^2$ is (assume $\hat{\mathbf{p}}\cdot\hat{\mathbf{q}} \neq -1$)

$$\frac{d\sigma}{d\Omega}^{UR} \sim \frac{\alpha^2 Z^2}{(\mathbf{q}-\mathbf{p})^4}[\mathbf{p}^2 + \mathbf{q}\cdot\mathbf{p}][1 + 4\lambda\lambda'].$$

where the high energy helicity conservation is transparent.

4.5. Pair production in a time varying external field

Let us return to the scalar field external field perturbation $H_I'(t) = -\int d^3x\phi_I^2 B(x)/2$ in the case where the initial state is the ground state of $H_{0,I}$, $|0\rangle_I$, and the final state contains two particles: $\langle 0, I|\, a(\vec{p}_1)a(\vec{p}_2)$. The transition amplitude to first order is

$$T_{fi} = \frac{i}{2}\int d^4x\, \langle 0, I|\, a(\vec{k}_1)a(\vec{k}_2)\phi_I^2(x)B(x)/2\, |0, I\rangle = \frac{i}{(2\pi)^3 2\sqrt{\omega_1\omega_2}}\int d^4x e^{-i(k_1+k_2)\cdot x}B(x) \equiv \frac{i\tilde{B}(k_1+k_2)}{(2\pi)^3 2\sqrt{\omega_1\omega_2}}$$

Squaring and integrating over final momenta we get the toal pair production probability

$$P_{\text{pair}} = \frac{1}{2}\int \frac{d^3k_1 d^3k_2}{(2\pi)^6 4\omega_1\omega_2}|\tilde{B}(k_1+k_2)|^2 = \frac{1}{4(2\pi)^5}\int d^4K\theta(-K^2-4m^2)\sqrt{1+\frac{4m^2}{K^2}}|\tilde{B}(K)|^2$$

since $\omega_1 + \omega_2 > 2m$, $P_{\text{pair}} \neq 0$ only when B oscillates with frequencies greater than $2m$. In particular a static external field will not produce pairs in perturbation theory.

4.6. Perturbation theory for Time Ordered Products

Another class of quantities that will be very useful to us is the matrix element of the time ordered product of a finite number of Heisenberg picture fields between asymptotic states:

$$\langle out|\, T[A_1(t_1)A_2(t_2)\cdots A_N(t_N)]\, |in\rangle\,.$$

The simplest way to transcribe this matrix element to interaction picture is to first assume the ordering

$t_1 > t_2 \cdots > t_N$ so that the T symbol can be removed. Then

$$\begin{aligned}
&\langle out|\, A_1(t_1)A_2(t_2)\cdots A_N(t_N)\,|in\rangle \\
&\quad = \langle 0,I|\, U_I(\infty,-\infty)U_I^{-1}(t_1,-\infty)A_{I1}(t_1)U_I(t_1,-\infty)U_I^{-1}(t_2,-\infty)A_{I2}(t_2) \\
&\qquad U_I(t_2,-\infty)\cdots U_I^{-1}(t_N,-\infty)A_{IN}(t_N)U(t_N,-\infty)\,|0,I\rangle \\
&\quad = \langle 0,I|\, U_I(\infty,t_1)A_{I1}(t_1)U_I(t_1,t_2)A_{I2}(t_2)U_I(t_2,t_3)\cdots U_I(t_{N-1},t_N)A_{IN}(t_N) \\
&\qquad U_I(t_N,-\infty)\,|0,I\rangle\,,
\end{aligned}$$

where use has been made of the closure property of U_I. Now we notice that all of the interaction picture operators that appear in the final matrix element, including those in the Dyson formula for each U_I are time ordered. Thus if we insert the time ordering symbol in front of all the operators we can combine all of the U_I's into a single $U_I(\infty,-\infty)$ arriving at

$$\begin{aligned}
&\langle out|\, T[A_1(t_1)A_2(t_2)\cdots A_N(t_N)]\,|in\rangle \\
&\qquad = \langle 0,I|\, T[U_I(\infty,-\infty)A_{I1}(t_1)A_{I2}(t_2)\cdots A_{IN}(t_N)]\,|0,I\rangle\,.
\end{aligned}$$

Finally, we simply note that had the time ordering been any other, the same steps would have led to the same final result.

In the usual situation where H_{0I} is the Hamiltonian for free fields, all of the interaction picture operators are free fields, and to evaluate each finite order in perturbation theory one only needs to master the computation of the vacuum expectation values of the time ordered product of a finite number of free fields. Free fields can always be expressed as a linear functional of creation and annihilation operators. Thus if $\phi_k(x)$ is a free field, it can be written

$$\phi_k(x) = \phi_k^+(x) + \phi_k^-(x)$$

where $\phi_k^+(x)$ annihilates $|0\rangle$ and $\phi_k^-(x)$ annihilates $\langle 0|$. Thus we have

$$\begin{aligned}
&\langle 0|T[\phi_k(x)\phi_l(x')]\,|0\rangle \\
&=\theta(t-t')\,\langle 0|\,\phi_k^+(x)\phi_l^-(x')\,|0\rangle \pm \theta(t'-t)\,\langle 0|\,\phi_l^+(x')\phi_k^-(x)\,|0\rangle \\
&\equiv\theta(t-t')C_{kl}(x-x') \pm \theta(t'-t)C_{lk}(x'-x)
\end{aligned}$$

where $C_{kl}(x-x') = [\phi_k^+(x),\phi_l^-(x')]_\pm$ is a c number since the fields are free.

Now consider a general time ordered product of N free fields

$$\langle 0|\, T[\phi_1(x_1)\cdots\phi_N(x_N)]\,|0\rangle$$

and first assume $t_1 > t_2 > \cdots > t_N$. Then the leftmost field is ϕ_1 and it can be replaced by its annihilation part ϕ_1^+, which is then moved via the commutation relations all the way to the right where it kills the vacuum. The (anti)commutators

$$[\phi_1^+(x_1),\phi_k(x_k)]_\pm = [\phi_1^+(x_1),\phi_k^-(x_k)]_\pm = \langle 0|\, T[\phi_1(x_1)\phi_k(x_k)]\,|0\rangle$$

since the (anti)commutators are c numbers and $t_1 > t_k$ by assumption. Thus

$$\begin{aligned}
\langle 0|T[\phi_1(x_1)\cdots\phi_N(x_N)]\,|0\rangle &= \langle 0|\, T[\phi_1(x_1)\phi_2(x_2)]\,|0\rangle\,\langle 0|\, T[\phi_3(x_3)\cdots\phi_N(x_N)]\,|0\rangle \\
&\pm \langle 0|\, T[\phi_1(x_1)\phi_3(x_3)]\,|0\rangle\,\langle 0|\, T[\phi_2(x_2)\phi_4(x_4)\cdots\phi_N(x_N)]\,|0\rangle \pm\cdots
\end{aligned} \tag{4.6}$$

where the sign in front of each term is dictated by the number of times the order of fermionic operators is switched. The time ordering symbol is not needed with our assumed ordering of times. But now we notice

that if the time ordering had been any other the same steps would have led to the same result *provided* we keep the T symbol in place. Thus we have related the vacuum expectation value of the time ordered product of N free fields to those of 2 free fields and $N-2$ free fields. By induction we can therefore express the vacuum expectation value of the time ordered product of N free fields as sums of products of the vacuum expectation values of the time ordered product of pairs of free fields.

The result, known as Wick's Theorem, can be expressed as follows. First note that the answer is 0 unless N is even. Then the vacuum expectation value of the time ordered product of N free fields is the sum of terms, one for each distinct pairing off of all the N fields. The term for each such pairing off is simply $\pm$ the product of the vacuum expectation values of the time ordered product of each pair of fields in the given pairing off. The sign is determined by comparing the ordering of the N operators in the original time ordered product with the order they appear in the given term after being paired off. If the latter ordering is achieved by an odd permutation of fermionic operators the sign is $-$; otherwise it is $+$. It doesn't matter what order we display the factors within a given term, since switching their order would always be an even permutation: a pairing of a boson field with a fermion field would always contribute zero!

4.7. A Technical Comment on Time Derivatives in Time Dependent Perturbation Theory

It is important to appreciate some subtle differences between time derivatives of operators in different pictures. For example, interaction picture depends on a specific breakup of the Heisenberg picture Hamiltonian $H = H_0 + H'$, so in Heisenberg picture

$$\dot{\Omega} = \frac{1}{i}[\Omega, H_0] + \frac{1}{i}[\Omega, H'].$$

The transformation to interaction picture, being a purely algebraic similarity transformation shows that

$$\begin{aligned}(\dot{\Omega})_I =& \frac{1}{i}[\Omega_I, H_{0I}] + \frac{1}{i}[\Omega_I, H'_I] \\ =& \dot{\Omega}_I + \frac{1}{i}[\Omega_I, H'_I],\end{aligned}$$

so there is in general a discrepancy between the interaction picture operator corresponding to $\dot{\Omega}$ and the time derivative of the operator Ω_I. When we use the Dyson formula for time dependent perturbation theory to calculate a matrix element involving $\dot{\Omega}$, care must be taken about this difference. However there is a very simple prescription to keep things straight. This is to always think of time derivatives of operators in the Dyson formula as acting *outside* the time ordering symbol. Note the following identity

$$\begin{aligned}\frac{\partial}{\partial t} T[e^{-i\int_{-\infty}^{\infty} dt' H'_I(t')} \Omega_I(t)] =& T[e^{-i\int_t^{\infty} dt' H'_I(t')}](\dot{\Omega}_I(t) + \frac{1}{i}[\Omega_I(t), H'_I(t))])T[e^{-i\int_{-\infty}^{t} dt' H'_I(t')}] \\ =& T[e^{-i\int_{-\infty}^{\infty} dt' H'_I(t')}(\dot{\Omega}(t))_I].\end{aligned}$$

This comment becomes particularly useful in cases such as scalar electrodynamics where the relation between $\pi^\dagger(x)$ and $\dot{\phi}(x) = \pi^\dagger(x) + iQA_0\phi(x)$ involves the interaction. Since the interaction picture fields are free, the relationship in that picture is $\dot{\phi}_I(x) = \pi^\dagger_I(x)$. Thus the exponent in the Dyson formula shows a disquieting asymmetry between space and time:

$$\int dt H'_I(t) = \int d^4x \left(iQ\mathbf{A} \cdot (\phi^\dagger_I \nabla \phi_I - (\nabla \phi^\dagger_I)\phi_I) + Q^2 \mathbf{A}^2 \phi^\dagger_I \phi_I - iQA_0(\phi^\dagger_I \dot{\phi}_I - \dot{\phi}^\dagger \phi_I) \right).$$

However, it is possible to prove a "reshuffling theorem" that if all time derivatives in the Dyson formula are understood to be taken outside the time ordering symbol, covariance is restored. In other words there

are two sources of apparent non -covariance: the form of H_I' and the time ordering operation itself. To present the results of the reshuffling theorem, we introduce the symbol T^* to signify time ordering in which all time derivatives are taken outside the time ordering symbol. Then the reshuffling theorem for scalar electrodynamics can be stated

$$T[e^{-i\int_{-\infty}^{\infty} dt' H_I'(t')}] = T^*\left[e^{-i\int d^4x (iQA^\mu(\phi_I^\dagger \partial_\mu \phi_I - (\partial_\mu \phi_I^\dagger)\phi_I) + Q^2 A^\mu A_\mu \phi_I^\dagger \phi_I)}\right].$$

When employing the Wick expansion to the r.h.s. one simply needs to remember that one never uses a quantity such as $\langle 0| T[\partial_\mu \phi(x)\phi^\dagger(y)]|0\rangle$, namely all derivatives occur outside not inside the time ordering symbols.

As an illuminating example of these ideas we quote the improved Dyson formula for the outin matrix element of the current operator:

$$\langle out| j_\mu(x) |in\rangle = \langle 0, I| T^*\left[e^{-i\int d^4x (iQA^\mu \cdot (\phi_I^\dagger \partial_\mu \phi_I - (\partial_\mu \phi_I^\dagger)\phi_I) + Q^2 A^\mu A_\mu \phi_I^\dagger \phi_I)}\right.$$
$$\left.(-iQ(\phi_I^\dagger \partial_\mu \phi_I - (\partial_\mu \phi_I^\dagger)\phi_I) - 2Q^2 A_\mu \phi_I^\dagger \phi_I)\right] |0, I\rangle .$$

Take particular note of the manifest covariance of the r.h.s. of this formula.

4.8. Propagators for Scalar and Dirac Fields

Wick's Theorem assures us that to obtain a general time ordered product of free fields, we only need to know the two field case, $\langle 0| T[\phi_1(x_1)\phi_2(x_2)]|0\rangle$, which is also called the two point function and sometimes the propagator.

Let us first work out the propagator for a free scalar field which has the representation

$$\phi(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \left(a(\mathbf{p})e^{ip\cdot x} + b^\dagger(\mathbf{p})e^{-ip\cdot x}\right), \tag{4.7}$$

where a annihilates a particle and $b^\dagger$ creates an antiparticle. These operators satisfy the commutation relations

$$[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = [b(\mathbf{p}), b^\dagger(\mathbf{p}')] = \delta(\mathbf{p}' - \mathbf{p}),$$

with all other commutators vanishing[⋆]. The Hamiltonian for the free scalar field is easy to write down

$$H = \int d^3p\omega(\mathbf{p})(a^\dagger(\mathbf{p})a(\mathbf{p}) + b^\dagger(\mathbf{p})b(\mathbf{p}))$$
$$= \int d^3x : (\dot\phi^\dagger\dot\phi + \nabla\phi^\dagger \cdot \nabla\phi + m^2\phi^\dagger\phi) :$$

where the double colons $: (\cdots) :$ denotes normal ordering, *i.e.* all creation operators to the left of all annihilation operators.

⋆ Notice that the commutator

$$[\phi(x_1), \phi^\dagger(x_2)] = \int \frac{d^3p}{(2\pi)^3 2\omega(\mathbf{p})}(e^{ip\cdot(x_1-x_2)} - e^{ip\cdot(x_2-x_1)})$$

vanishes for space-like separations $(x_2 - x_1)^2 > 0$. To see this go to a Lorentz frame where $t_2 = t_1$ (always possible for space-like separations). Then the second term cancels the first after the variable change $\mathbf{p} \to -\mathbf{p}$. If a, b satisfied anticommutation relations, the anticommutator would not have this locality property since the two terms would then add. This is the spin-statistics connection for scalar fields. Also notice that if the $\omega(\mathbf{p})$ were absent anticommutation relations *would* be local.

Clearly the vacuum expectation of the time ordered product of two ϕ's or two $\phi^\dagger$'s vanishes, and

$$\begin{aligned}\langle 0|\, T[\phi(x_1)\phi^\dagger(x_2)]\,|0\rangle =&\theta(t_1-t_2)\int \frac{d^3p}{(2\pi)^3 2\omega(\mathbf{p})}e^{ip\cdot(x_1-x_2)}\\ &+\theta(t_2-t_1)\int \frac{d^3p}{(2\pi)^3 2\omega(\mathbf{p})}e^{ip\cdot(x_2-x_1)},\end{aligned} \tag{4.8}$$

where we recall that $p\cdot x=\mathbf{p}\cdot\mathbf{x}-\omega t$. To make this expression less unwieldy, it is helpful to use the following integral representation for the step function

$$\theta(t)=\int_{-\infty}^{\infty}\frac{dp^0}{2\pi i}e^{-ip^0t}\frac{1}{-p^0-i\epsilon}\qquad\qquad \epsilon\to 0_+.$$

Including the factor $e^{-i\omega t}$ gives

$$\theta(t)e^{-i\omega t}=\int_{-\infty}^{\infty}\frac{dp^0}{2\pi i}e^{-ip^0t}\frac{1}{\omega-p^0-i\epsilon},$$

after a shift of p^0. Inserting this representation into (4.8), gives

$$\langle 0|\, T[\phi(x_1)\phi^\dagger(x_2)]\,|0\rangle=\int\frac{d^4p}{(2\pi)^4 2\omega(\mathbf{p})}e^{ip\cdot(x_1-x_2)}\frac{-i}{\omega-p^0-i\epsilon}+e^{ip\cdot(x_2-x_1)}\frac{-i}{\omega-p^0-i\epsilon},$$

where now $p\cdot x=\mathbf{p}\cdot\mathbf{x}-p^0t$. Thus we can change $p\to -p$ in the second term and then combine it with the first to obtain finally

$$\langle 0|\, T[\phi(x_1)\phi^\dagger(x_2)]\,|0\rangle=\int\frac{d^4p}{(2\pi)^4}e^{ip\cdot(x_1-x_2)}\frac{-i}{p^2+m^2-i\epsilon}\equiv\Delta_F(x_1-x_2). \tag{4.9}$$

From its definition the propagator should have the property that only positive frequency components should be present as $t_k\to+\infty$ and negative frequency components as $t_k\to-\infty$. This property is assured in (4.9) by the $-i\epsilon$ in the denominator. The propagator is a Green function for the Klein-Gordon differential $-\partial^2+m^2$:

$$(-\partial^2+m^2)\Delta_F(x,y)=-i\delta(x-y).$$

The $i\epsilon$ prescription tells us which boundary conditions to impose. This prescription also follows if we define Δ_F by analytically continuing the Euclidean space Green function to Minkowski space. To continue from Minkowski space to Euclidean space, one rotates the p^0 integration contour to the imaginary axis in the counterclockwise direction (to avoid the poles at $\pm(\omega-i\epsilon)$). In order to preserve convergence at infinity, x^0 must be simultaneously rotated in the opposite (clockwise) direction. Changing variables $p^0=-ip^4$ and calling $x^0=-ix^4$ (positive p^0 rotates to negative p^4 but positive x^0 rotates to positive x^4) gives the Euclidean Green function

$$\Delta_F\to\Delta_E\equiv\int\frac{d^4p}{(2\pi)^4}e^{i(\mathbf{x}\cdot\mathbf{p}+x^4p^4)}\frac{1}{\mathbf{p}^2+(p^4)^2+m^2}.$$

Clearly Δ_E satisfies

$$(-\partial_4^2-\nabla^2+m^2)\Delta_E(x-y)=\delta^4(x-y).$$

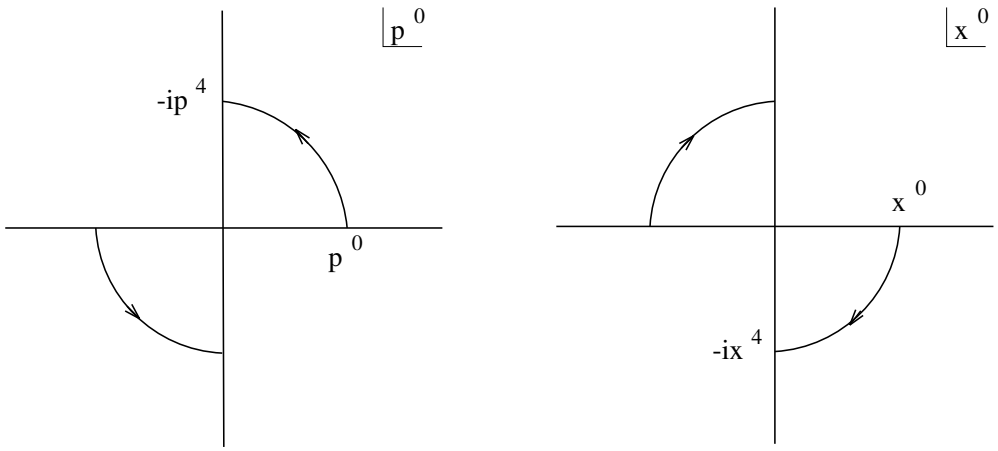


FIG. 1. Wick Rotations.

Next we turn to the evaluation of the propagator for the Dirac field. Remembering that ψ is fermionic we have

$$\begin{aligned} S_F(x_1 - x_2)_{ab} &\equiv \langle 0| T[\psi_a(x_1)\bar{\psi}_b(x_2)] |0\rangle \\ &= \theta(t_1 - t_2) \langle 0| \psi_a(x_1)\bar{\psi}_b(x_2) |0\rangle - \theta(t_2 - t_1) \langle 0| \bar{\psi}_b(x_2)\psi_a(x_1) |0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3 2\omega(\mathbf{p})}[\theta(t_1 - t_2)e^{ip\cdot(x_1-x_2)} \sum_\lambda u^a_\lambda(\mathbf{p})\bar{u}^b_\lambda(\mathbf{p}) \\ &\quad - \theta(t_2 - t_1)e^{ip\cdot(x_2-x_1)} \sum_\lambda v^a_\lambda(\mathbf{p})\bar{v}^b_\lambda(\mathbf{p})], \end{aligned} \tag{4.10}$$

where in the integrand we have $p^0 = \omega(\mathbf{p})$.

To simplify the expression for S_F we need to evaluate $\sum_\lambda u^a_\lambda(\mathbf{p})\bar{u}^b_\lambda(\mathbf{p})$ and $\sum_\lambda v^a_\lambda(\mathbf{p})\bar{v}^b_\lambda(\mathbf{p})$. If we regard them as matrices with indices a, b, we know we can write each as a linear combination of the 16 matrices Γ we used in constructing the bilinears. By virtue of the summ over helicity, conjugation by the Lorentz transformation matrices simply does the corresponding Lorentz transformation on $p^\mu = (\mathbf{p}, \omega)$. Thus they must be scalars formed from p and the matrices Γ. The only possibilities are I and $p \cdot \gamma$, so

$$\sum_\lambda u^a_\lambda(\mathbf{p})\bar{u}^b_\lambda(\mathbf{p}) = A\delta_{ab} + Bp \cdot \gamma_{ab},$$

and we only need to determine A, B. First notice that multiplying by the matrix $m + p \cdot \gamma$ must give 0. which determines $A = -mB$. Then, multiplying both sides by γ^0 gives $2\omega \sum_\lambda B = -4\omega$ or $B = -1$. Thus

$$\sum_\lambda u^a_\lambda(\mathbf{p})\bar{u}^b_\lambda(\mathbf{p}) = (m - p \cdot \gamma)_{ab}.$$

The definition $v = i\gamma^2 u^*$ then determines

$$\sum_\lambda v^a_\lambda(\mathbf{p})\bar{v}^b_\lambda(\mathbf{p}) = [i\gamma^2(m - p\cdot\gamma^*)(-i\gamma^2)]_{ab} = -(m + p\cdot\gamma)_{ab}.$$

Inserting these relations into (4.10) yields

$$S_F(x_1 - x_2)_{ab} = \int \frac{d^3p}{(2\pi)^3 2\omega(\mathbf{p})}[\theta(t_1 - t_2)e^{ip\cdot(x_1-x_2)}(m - p\cdot\gamma)_{ab} + \theta(t_2 - t_1)e^{ip\cdot(x_2-x_1)}(m + p\cdot\gamma)_{ab}].$$

The final step is to employ the integral representation for the step functions as we did for the scalar propagator. This process results in a four dimentional momentum integral with p^0 substituted for ω in the exponents but not in front of γ^0.Then the two terms involving m and those involving the spatial γ^k combine, after the change of variable $p \to -p$ in the second term, exactly as in the scalar case. The two terms involving γ^0 have a factor of ω which cancels that in the denominator, but then they combine with the opposite relative sign to produce a $2p^0$ in the numerator. Thus the net result is simply

$$S_F(x_1 - x_2)_{ab} = -i\int \frac{d^4p}{(2\pi)^4}e^{ip\cdot(x_1-x_2)}\left(\frac{m - p\cdot\gamma}{m^2 + p^2 - i\epsilon}\right)_{ab}. \tag{4.11}$$

Just as with the scalar propagator S_F may be recognized as a Green function for the differential dirac operator $\frac{1}{i}\gamma\cdot\partial + m$:

$$\left(\frac{1}{i}\gamma\cdot\partial + m\right)S_F(x - y) = -i\delta^4(x - y),$$

with boundary condition dictated by the $i\epsilon$ prescription. As before this boundary condition can be enforced by defining S_F as the continuation of the Euclidean Green function S_E. The continuation from S_F to S_E proceeds by rotating the p^0 integration contour to the imaginary axis in the counterclockwise direction (of course rotating x^0 in the opposite direction), changing variables $p^0 = -ip^4$, and defining $x^0 = -ix^4$, $\gamma^0 = -i\gamma^4_E$:

$$S_F(x - y) \to S_E(x - y) = \int \frac{d^4p}{(2\pi)^4}e^{ip\cdot(x_1-x_2)}\left(\frac{m - p\cdot\gamma_E}{m^2 + p^2}\right)_{ab}. \tag{4.12}$$

S_E satisfies the equation

$$\left(\frac{1}{i}\gamma\cdot\partial + m\right)S_E(x - y) = \delta^4(x - y).$$

4.9. Vacuum expectations from large time limits of general transition amplitudes.

One obstacle to formulating an efficient perturbation theory for systems with interacting quantum fields is that one can't "turn off" the interactions at early and late times as is possible with externally applied fields. Thus *out* and *in* states are eigenstates of complicated interacting Hamiltonians. One approach to this difficulty is to artificially make the coupling constants time dependent and force them to vanish at early and late times. Another approach, which we shall favor, is to relax the requirement that the initial and final states be eigenstates of the Hamiltonian with vanishing external fields. Then one calculates in first instance a quantity that is not of immediate interest, but which can be simply related to such quantities.

A quantity of more or less direct physical interest is the vacuum expectation value of the time ordered product of several quantum fields. More generally the *outin* matrix element of such a time ordered product is relevant if time varying external fields are present. So let us consider how to obtain this quantity in perturbation theory by first calculating with general initial and final states. Using the evolution operator and assuming $t_1 > t_2 > \cdots > t_n$, we therefore consider

$$\langle f| U(\infty,-\infty)T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|i\rangle = \\ \langle f| U(\infty,t_1)\Omega_{S1}U(t_1,t_2)\cdots U(t_{n-1},t_n)\Omega_{Sn}U(t_n,-\infty)\,|i\rangle\,.$$

Choose the time T so that all external fields vanish for times earlier than $-T$ and later than T. Then

$$\begin{aligned} U(t_n,-\infty)\,|i\rangle =& U(t_n,-T)e^{-i(\infty-T)H_S}\,|i\rangle \\ =& U(t_n,-T)e^{-i(\infty-T)E_G}\sum_r e^{-i(\infty-T)(E_r-E_G)}\,|r\rangle\,\langle r|i\rangle\,. \end{aligned}$$

We would now like to argue that the infinite oscillations wash out all contributions but the (assumed nondegenerate[⋆]) ground state. In a field theory this is quite plausible since the excited states correspond to particles so the sum over r is really an integral over a range of continuous energies. But even without this smearing, we can make the washing out rigorous by calculating with imaginary time: $it = \beta > 0$. Then $i\infty$ is really $+\infty$ and all excited states are damped exponentially. Massless particle states could introduce a subtlety here, but the part of phase space that is not exponentially damped is infinitesimal: this has the effect of changing exponential damping to a power law damping. If we buy this argument, then we can assert quite generally that $U(t_n,-\infty)\,|i\rangle = U(t_n,-\infty)\,|0\rangle\,\langle 0|i\rangle$ and similarly $\langle f|\,U(\infty,t_1) = \langle f|0\rangle\,\langle 0|\,U(\infty,t_1)$.

Since we take (as usual) the Heisenberg and Schrödinger pictures to coincide at $t=-\infty$, then $|in\rangle = |0\rangle$ and $\langle out| = \langle 0|\,U(\infty,-\infty)$. Thus we have obtained the relation

$$\langle f|\,U(\infty,-\infty)T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|i\rangle = \\ \langle f|0\rangle\,\langle 0|i\rangle\,\langle out|\,T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|in\rangle\,.$$

In other words calculating with any initial and final states that have finite overlap[†] with the true ground state gives us a constant times the desired matrix element. We can easily evaluate the multiplicative constant by considering by the same reasoning

$$\begin{aligned} \langle f|\,U(\infty,-\infty)\,|i\rangle &= \langle f|0\rangle\,\langle 0|i\rangle\,\langle out|in\rangle \\ &\to e^{-2i\infty E_G}\,\langle f|0\rangle\,\langle 0|i\rangle \qquad \text{External Fields} = 0. \end{aligned}$$

Putting this into our relation we obtain

$$\langle out|\,T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|in\rangle = \\ e^{-2i\infty E_G}\,\frac{\langle f|\,U(\infty,-\infty)T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|i\rangle}{\langle f|\,U(\infty,-\infty)_{Ext=0}\,|i\rangle}, \tag{4.13}$$

where the subscript on U in the denominator denotes vanishing external fields. In field theory applications E_G is the energy of the vacuum, which is zero if we measure all energies relative to that of the vacuum. In the

⋆ There are interesting cases of degenerate vacua, when there is "spontaneous symmetry breakdown." In such cases the choice of initial and final states determines which of the degenerate vacua is picked out.

† The infinite number of degrees of freedom in quantum field theory requires care here: the overlap between different states in a theory with n degrees of freedom is typically f^n with $f<1$. Since $n=\infty$, we should expect $\langle f|0\rangle\,\langle 0|i\rangle \sim e^{-\infty}$. In field theory $n=\infty$ because the volume of space is infinite and because space is continuous. Thus strict application of the above relation should be done in the presence of both an infrared and ultraviolet cutoff, which can then be removed after extracting the desired amplitude.

absence of gravity all physical quantities depend only on energy differences, so we lose nothing by doing this. Gravity couples directly to the energy density and therefore is sensitive to the energy as opposed to energy differences, but then E_G only appears in the combination $\Lambda \equiv E_G + \Lambda_0$, with Λ_0 the "bare" cosmological constant. Replacing Λ_0 by Λ in effect sets $E_G = 0$.

The formula (4.13) is a convenient starting point for developing perturbation theory. Any breakup

$$H_S(t) = H_0(t) + H'(t)$$

determines an interaction picture defined by

$$\Omega_I(t) = U_0^{-1}(t,-\infty)\Omega_S U_0(t,-\infty) = U_I(t,-\infty)\Omega(t)U_I^{-1}(t,-\infty),$$

where

$$\begin{aligned} i\dot{U} =& H_S(t)U = UH(t) \\ i\dot{U}_0 =& U_0 H_{0I}(t) \\ i\dot{U}_I =& H_I(t)U_I \end{aligned}$$

and all U's are the identity at $t = -\infty$. Then the evolution operator satisfies

$$U(t_1,t_2) = U(t_1,-\infty)U^{-1}(t_2,-\infty) = U_0(t_1,-\infty)U_I(t_1,t_2)U_0^{-1}(t_2,-\infty).$$

Plugging these relations into (4.13) then gives

$$\begin{aligned} &\langle out|\, T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|in\rangle = \\ &\qquad e^{-2i\infty E_G}\frac{\langle f|\, U_0(\infty,-\infty)T[U_I(\infty,-\infty)\Omega_{1I}(t_1)\cdots\Omega_{nI}(t_n)]\,|i\rangle}{\langle f|\, U_0(\infty,-\infty)_{Ext=0}U_I(\infty,-\infty)_{Ext=0}\,|i\rangle}. \end{aligned} \tag{4.14}$$

This formula is completely general: we have even allowed H_0 to contain time varying external fields, which is hardly ever done in practice. Since all operators in this formula are in interaction picture, it is most convenient to choose $|i\rangle, |f\rangle$ to have simple properties with respect to $H_{0I}(-\infty)$. Let us call the ground state of this operator $|in,0\rangle$. Then $\langle in,0|\, U_0(\infty,-\infty)$ is the ground state of $H_{0I}(+\infty)$ and therefore deserves the name $\langle out,0|$. When all external fields vanish, H_{0I} is time indepenent and we call its ground state $|0,I\rangle \equiv |in,0\rangle$ and its ground state energy E_0. Then $\langle in,0|\, U_0(\infty,-\infty)_{Ext=0} = e^{-2i\infty E_0}\,\langle 0,I|$. Thus choosing $|i\rangle = |f\rangle = |0,I\rangle = |in,0\rangle$ we obtain the useful formula

$$\begin{aligned} &\langle out|\, T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|in\rangle = \\ &\qquad e^{-2i\infty(E_G-E_0)}\frac{\langle out,0|\, T[U_I(\infty,-\infty)\Omega_{1I}(t_1)\cdots\Omega_{nI}(t_n)]\,|in,0\rangle}{\langle 0,I|\, U_I(\infty,-\infty)_{Ext=0}\,|0,I\rangle}. \end{aligned} \tag{4.15}$$

In the usual case where we do not include external fields in H_0, the formula simplifies further

$$\begin{aligned} &\langle out|\, T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|in\rangle = \\ &\qquad e^{-2i\infty E_G}\frac{\langle 0,I|\, T[U_I(\infty,-\infty)\Omega_{1I}(t_1)\cdots\Omega_{nI}(t_n)]\,|0,I\rangle}{\langle 0,I|\, U_I(\infty,-\infty)_{Ext=0}\,|0,I\rangle}. \end{aligned} \tag{4.16}$$

Using the Wick expansion one can describe the perturbation series for the numerators and denominators of these formulas using Feynman diagrams. The diagrams contributing to the denominator are all those completely disconnected from either external fields or from the points assigned to the operators in the numerator. The numerator contains this same sum of diagrams as a multiplicative factor. Thus the division by the denominator is achieved by simply deleting all such disconnected "vacuum" diagrams from the expansion of the numerator.

5. Path History Quantization

One of the principal drawbacks of the canonical operator formulation of quantum mechanics is that it obscures symmetries that bring in time in an essential way. Lorentz boosts are of this type, so the operator approach inevitably hides the full symmetries of relativity. This is of course also true in the Hamilton equation form of classical mechanics. In classical mechanics one can work with the Lagrangian and Action Principle which keep dynamical symmetries like Poincaré transparent. The path integral approach to quantum mechanics is the quantum analogue of this alternative and, as we shall see, is a much more convenient formulation of quantum field theory than the operator approach. Even in the operator approach we have seen the advantage of expressing results in terms of finite time evolutions, since it is these that reflect the true symmetries of the system. The central object in the path integral approach then is not a state but an amplitude for the evolution of one state into another.

To keep notation simple, we shall suppress indices in dealing with a general quantum mechanical system the coordinates of which are collectively denoted q and the conjugate momenta of which are p. Then we seek an alternative scheme for calculating, for example, the amplitude

$$\langle q''|\, U(t_1,t_2)\, |q'\rangle\,.$$

U is of course a very complicated operator for finite time differences, but as $t_1 \to t_2$, it is simply related to the Hamiltonian.

Thus we are led to break up the time interval into infinitesimal pieces $t_1 - t_2 = (N+1)a$ and employ the closure property of U to write it as a product of $N+1$ infinitesimal evolutions.

$$\begin{aligned}\langle q''|\, U(t_1,t_2)\, |q'\rangle &= \langle q''|\, U(t_1,t_1-a)U(t_1-a,t_1-2a)\cdots U(t_2+a,t_2)\, |q'\rangle \\ &= \int \prod_{k=1}^{N} dq_k\, \langle q''|\, U(t_1,t_1-a)\, |q_N\rangle\, \langle q_N|\, U(t_1-a,t_1-2a)\, |q_{N-1}\rangle \\ &\qquad\qquad \cdots \langle q_1|\, U(t_2+a,t_2)\, |q'\rangle\,.\end{aligned}$$

Next, we write, assuming H_S is constant over a time interval a,

$$\begin{aligned}&\langle q_k|\, U(t_2+ka,t_2+(k-1)a)\, |q_{k-1}\rangle \\ &\quad\approx \int dp_{k-1}\, \langle q_k|\, e^{-iaH(t_2+(k-1/2)a)/2}\, |p_{k-1}\rangle\, \langle p_{k-1}|\, e^{-iaH(t_2+(k-1/2)a)/2}\, |q_{k-1}\rangle \\ &\quad\equiv \int dp_{k-1} \exp\{-\frac{i}{2\hbar}a(\mathcal{H}_k(q_k,p_{k-1},ia)+\mathcal{H}_k(q_{k-1},p_{k-1},-ia^*)^*)\} \\ &\qquad\qquad \langle q_k|p_{k-1}\rangle\, \langle p_{k-1}|q_{k-1}\rangle \\ &\quad= \int \frac{dp_{k-1}}{2\pi\hbar} \exp\{\frac{i}{\hbar}[p_{k-1}(q_k-q_{k-1}) \\ &\qquad\qquad -\frac{a}{2}(\mathcal{H}_k(q_k,p_{k-1},ia)+\mathcal{H}_k(q_{k-1},p_{k-1},-ia^*)^*)]\},\end{aligned}$$

where for the moment $\mathcal{H}_k$ is defined by these equations. Putting everything together we obtain

$$\begin{aligned}\langle q''|\, U(t_1,t_2)\, |q'\rangle = \int \prod_{k=1}^{N} \frac{dq_k dp_k}{2\pi\hbar}\frac{dp_0}{2\pi\hbar} \exp\{\frac{i}{\hbar}\sum_{k=1}^{N+1}[p_{k-1}(q_k-q_{k-1}) \\ -\frac{a}{2}(\mathcal{H}_k(q_k,p_{k-1},ia)+\mathcal{H}_k(q_{k-1},p_{k-1},-ia^*)^*)]\},\end{aligned} \tag{5.1}$$

where $q_{N+1} \equiv q''$ and $q_0 \equiv q'$. Apart from the assumption that external fields are constant over the time interval a this formula is exact. But it is not useful until we get a simple approximation for $\mathcal{H}_k$. For $a \to 0$

we should be able to approximate

$$\begin{aligned}\langle q_k| e^{-iaH(t_2+(k-1/2)a)/2} |p_{k-1}\rangle &\approx \langle q_k| (1 - i\frac{a}{2}H(t_2 + (k-1/2)a)) |p_{k-1}\rangle \\ &\approx (1 - i\frac{a}{2}H_k^W(q_k, p_{k-1})) \langle q_k|p_{k-1}\rangle \\ &\approx \exp\{-i\frac{a}{2}H_k^W(q_k, p_{k-1})\} \langle q_k|p_{k-1}\rangle\end{aligned}$$

where $H_k^W(q,p)$ is the operator $H(t_2 + (k-1/2)a)$ rewritten through use of the canonical commutation relations with all p's on the right and all q's on the left. After this is done q can then be replaced with the eigenvalue q_k and p with the eigenvalue p_{k-1}. In the limit $a \to 0$ with $t_1 - t_2 = (N+1)a$ fixed it should be valid to replace $\mathcal{H}_k(q_k, p_{k-1})$ with $H_k^W(q_k, p_{k-1})$. We define the quantity

$$H_k^Q(q_k, q_{k-1}, p_{k-1}) = \frac{1}{2}(H_k^W(q_k, p_{k-1}) + H_k^W(q_{k-1}, p_{k-1})^*) \tag{5.2}$$

which appears in the path integral. Up to the reordering terms H_k^W is just the classical Hamiltonian for the system. In the common case where the Hamiltonian is a function of p's plus a function of q's it is nothing more nor less than the classical Hamiltonian. But notice that the object appearing in the path integral is (5.2) which depends on the *two* coordinates describing the initial and final states of the basic unit of propagation even in the case where there are no reordering terms, when it is simply the average of the classical Hamiltonian over the two coordinates. In this continuum limit we can think of the sum in the exponent in (5.1) as an integral

$$\frac{i}{\hbar}\int_{t_2}^{t_1} dt[p(t)\dot{q}(t) - \frac{1}{2}(H^W(q_>(t), p(t), t) + H^W(q_<(t), p(t), t)^*)]. \tag{5.3}$$

The coefficient of $\frac{i}{\hbar}$ is just the classical action $\int dtL$ plus terms of order $\hbar^2$ expressed as a Legendre transform of the Hamiltonian. In the continuum limit, the number of integration variables tends to infinity and the limit gives the definition of the path integral representation of $\langle q_f| U(t_1, t_2) |q_i\rangle$.

<u>The Classical Limit</u> For a general Hamiltonian, this is as far as one can go without further approximations. One such approximation one can always try is the limit $\hbar \to 0$, the classical limit. Such a limit is dominated by the functions $q(t), p(t)$ for which the coefficient of $i/\hbar$ is stationary. Since, in this limit, this coefficient is just the classical action, the stationarity conditions are simply the classical Hamilton equations: $\dot{q} = \partial H/\partial p$ and $\dot{p} = -\partial H/\partial q$. Thus the path history version of the quantum principle is that for $\hbar \neq 0$ transition amplitudes are computed by evaluating $e^{i(Action)/\hbar}$ for all possible histories and averaging this expression over all such histories. The classical limit is understood as the situation in which this average is dominated by solutions of the classical equations of motion.

<u>Imaginary Time.</u> In working with the path integral it is technically advantageous to work with actually damped integrands rather than the oscillating integrand occurring in the quantum path integral. This can be achieved with the Wick rotation $it = \tau$ where real positive t is rotated to real positive τ. Considering the basic unit of the path integral, the matrix element of the operator, $e^{-iaH/\hbar}$, we see that this rotation is mathematically justified when H is an operator bounded below, *i.e.* its eigenvalue spectrum is bounded below. It obviously should not be attempted if H has eigenvalues down to $-\infty$. Fortunately, most reasonable physical systems have this property, and for these the Wick rotated path integral is the superior one to work with, especially for applications outside of perturbation theory. For a constant Hamiltonian (no external

fields) this path integral calculates $\langle q_f | e^{-\beta H} | q_i \rangle$ where $\beta = (\tau_1 - \tau_2)/\hbar$. If we identify $q_f = q_i = q$ and integrate over q, it calculates $\mathrm{Tr}\, e^{-\beta H}$, the statistical mechanical partition function for temperature $1/\beta$. In this way the Wick rotated path integral is related to a quantity of direct physical interest in another context. In the limit $\beta \to \infty$ (low temperature) one obtains information about the energy levels and degeneracies of the system. However, for applications to quantum mechanics it is necessary to continue back to real time at the end of the calculation of physical transition amplitudes.

Technically the Wick rotation amounts to replacing ia by $\delta > 0$ in (5.1). Thus one obtains, after approximating $\mathcal{H}_k$ by H_k^W,

$$\begin{aligned}\langle q'' | U(-i\tau_1, -i\tau_2) | q' \rangle \approx \int \prod_{k=1}^{N} \frac{dq_k dp_k}{2\pi\hbar} \frac{dp_0}{2\pi\hbar} \exp\{\frac{1}{\hbar} \sum_{k=1}^{N+1} [ip_{k-1}(q_k - q_{k-1}) \\ - \frac{\delta}{2}(H_k^W(q_k, p_{k-1}) + H_k^W(q_{k-1}, p_{k-1})^*)]\}.\end{aligned} \tag{5.4}$$

The Wick rotated version of the quantum action (5.3) is of course

$$\frac{1}{\hbar} \int_{\tau_2}^{\tau_1} d\tau [ip(\tau)\dot{q}(\tau) - \frac{1}{2}(H^W(q_>(\tau), p(\tau), \tau) + H^{W*}(q_<(\tau), p(\tau), \tau))]. \tag{5.5}$$

Here we have identified $q(t) \equiv q(\tau)$ and the complex conjugation in H^{W*} ignores the i's coming from the Wick rotation.

<u>Matrix Elements of Time Ordered Products.</u> For $T > t_1 > \cdots > t_n > -T$

$$\begin{aligned}\langle q_f | U(T, t_1) q_1 U(t_1, t_2) q_2 \cdots q_n U(t_n, -T) | q_i \rangle = \\ \langle q_f | U(T, -T) T[q_1(t_1) q_2(t_2) \cdots q_n(t_n)] | q_i \rangle\end{aligned}$$

where the Heisenberg picture operators $q_k(t_k) \equiv U^{-1}(t_k, -T) q_k U(t_k, -T)$ are defined so that Heisenberg and Schrödinger pictures agree at $t = -T$. Working with the l.h.s. of this relation we can insert a complete set of coordinate basis states between each pair of U's and then replace each operator q_k by its eigenvalue and each matrix element of U by its path integral representation. The integrals over the basis labels q'_k then simply extend the sum over piecewise histories $q_i \to q_n \cdots \to q_1 \to q_f$ to the sum over all histories $q_i \to q_f$. Thus we arrive at

$$\begin{aligned}&\langle q_f | U(T, -T) T[q_1(t_1) q_2(t_2) \cdots q_n(t_n)] | q_i \rangle \\ &= \int \mathcal{D}q \mathcal{D}p q(\tau_1) \cdots q(\tau_n) \exp\{\frac{1}{\hbar} \int_{-\mathcal{T}}^{\mathcal{T}} d\tau [ip(\tau)\dot{q}(\tau) - H^Q(q(\tau), p(\tau), \tau)]\}.\end{aligned}$$

where we have used the Wick rotated version with $iT = \mathcal{T}$. Written in this way with the time ordering symbol on the l.h.s. this formula is valid for any time ordering.

In field theory we are really interested in the vacuum (ground state) expectation value of time ordered products. These can be obtained by taking the limit $\mathcal{T} \to \infty$. Then, inserting energy eigenstates at the left

and right, all states but the ground states are exponentially suppressed so we have

$$\begin{aligned}
&\langle q_f|U(\infty,-\infty)T[q_1(t_1)q_2(t_2)\cdots q_n(t_n)]\,|q_i\rangle \\
&\quad=\langle q_f|0\rangle\,\langle 0|\,U(\infty,-\infty)T[q_1(t_1)q_2(t_2)\cdots q_n(t_n)]\,|0\rangle\,\langle 0|q_i\rangle \\
&\quad=\langle q_f|0\rangle\,\langle 0|q_i\rangle\,\langle out|\,T[q_1(t_1)q_2(t_2)\cdots q_n(t_n)]\,|in\rangle\,.
\end{aligned}$$

The wave functions that multiply the desired result can be obtained from

$$\langle q_f|\,U_{Ext=0}(\infty,-\infty)\,|q_i\rangle = e^{-2\infty E_G}\,\langle q_f|0\rangle\,\langle 0|q_i\rangle\,,$$

which of course has its own path integral representation. Thus by division we obtain, defining energy so that $E_G=0$,

$$\begin{aligned}
&\langle out|T[q_1(t_1)q_2(t_2)\cdots q_n(t_n)]\,|in\rangle \\
&\quad=\frac{\int \mathcal{D}q\mathcal{D}p\;q(\tau_1)\cdots q(\tau_n)\exp\{\frac{1}{\hbar}\int_{-\infty}^{\infty}d\tau[ip(\tau)\dot{q}(\tau)-H^Q(q(\tau),p(\tau),\tau)]\}}{\int \mathcal{D}q\mathcal{D}p\;\exp\{\frac{1}{\hbar}\int_{-\infty}^{\infty}d\tau[ip(\tau)\dot{q}(\tau)-H^Q_{Ext=0}(q(\tau),p(\tau))]\}}.
\end{aligned}$$

<u>Coordinate Space Path Integral.</u> A strong motivation for using the path integral formulation for quantum field theory is that it makes possible a more symmetrical treatment of space and time. This symmetry is even more striking after the Wick rotation when the Lorentz group $O(3,1)$ becomes simply $O(4)$ the group of rotations in four dimensions. To achieve the full force of this benefit though we would like to be able to use the configuration space action $\int dtL(q,\dot{q},t)$ rather than the phase space one. We can do this provided it is possible to "integrate out" the conjugate momenta p. This is generally possible in quantum field theory because field theoretic Hamiltonians typically only depend on the conjugate momenta Π through an additive term $\frac{1}{2}\int d^3x\Pi^2$. Thus the integral over the Π is gaussian and can be explicitly performed. In the language of quantum mechanics, the field theoretic Hamiltonian is always of the form $p^2/2+V(q)+[f(q)p+pf(q)]/2=p^2/2+V(q)+f(q)p-i\hbar f'(q)/2$. In this case we can integrate out the p's even before the continuum limit which converts ordinary integrals to path integrals:

$$\begin{aligned}
&\int\frac{dp_{k-1}}{\sqrt{2\pi\hbar}}\exp\{\frac{1}{\hbar}[ip_{k-1}(q_k-q_{k-1}) \\
&\quad-\delta(\frac{1}{2}p_{k-1}^2+\frac{1}{2}(V(q_k)+V(q_{k-1})+p_{k-1}(f(q_k)+f(q_{k-1})-\frac{i\hbar}{4}(f'(q_k)-f'(q_{k-1})))]\}. \\
&=\frac{1}{\delta^{1/2}}\exp\{-\frac{1}{\hbar}[\frac{(q_k-q_{k-1}+\frac{i\delta}{2}(f(q_k)+f(q_{k-1})))^2}{2\delta} \\
&\quad+\frac{\delta}{2}(V(q_k)+V(q_{k-1})-\frac{i\hbar}{2}(f'(q_k)-f'(q_{k-1})))]\}.
\end{aligned}$$

So that the path integral expression (5.4) becomes

$$\begin{aligned}
\langle q''|\,U\,|q'\rangle &\approx\left(\frac{1}{2\pi\hbar\delta}\right)^{\frac{(N+1)}{2}}\int\prod_{k=1}^{N}dq_k\exp\{-\frac{1}{\hbar}\sum_{k=1}^{N+1}[\frac{(q_k-q_{k-1}+\frac{i\delta}{2}(f(q_k)+f(q_{k-1})))^2}{2\delta} \\
&\quad+\frac{\delta}{2}(V(q_k)+V(q_{k-1})-\frac{i\hbar}{2}(f'(q_k)-f'(q_{k-1}))]\} \\
&\to\int\mathcal{D}q\exp\{-\frac{1}{\hbar}\int_{\tau_2}^{\tau_1}d\tau[\frac{1}{2}\left(\frac{dq}{d\tau}+\frac{i}{2}(f(q_k)+f(q_{k-1}))\right)^2+V(q(\tau))]\}.
\end{aligned}$$

Note that the (divergent) prefactor is necessary to obtain the same result for the evolution amplitude as with the usual operator formalism, and it naturally appears when we start the path history formulation in phase

space. However, notice also that the dependence on the evolution time $(N+1)\delta$ is exactly of the form that would come from adding a constant $-(1/2\delta)\ln(2\pi\hbar\delta)$ to the overall energy of the system. Thus since only energy differences are measurable, the physics is insensitive to the presence of this factor. In quantum field theory we could lump this contribution into the zero point vacuum energy we are supposed to subtract in any case. We recognize the coefficient of $1/\hbar$ in the exponent as the (imaginary time) classical action for the system. It is this configuration space path integral that gives quantum field theory its neatest expression. Then the classical action has the form $\int d^4x\mathcal{L}(\phi,\partial_\mu\phi)$ with $\mathcal{L}$ a Lorentz scalar field. The configuration variables of quantum field theory are the fields $\phi(\mathbf{x},t)$ one for each point in space time. To define the sum over histories of fields one therefore needs a lattice in space-time. In the Wick-rotated version this could be taken, for example, to be a hypercubic lattice in 4 dimensional Euclidean space. The path integral defined via such a lattice can be taken as the *definition* of quantum field theory. We shall find that perturbation theory can be developed directly from the continuum path integrals, essentially because one can avoid the actual evaluation of the integrals by various tricks.

Gaussian Integrals When evaluating ground state averages of physical quantities in perturbation theory, one can manage to avoid ever having to do a functional integral. This is because the free field functional integral will cancel between numerator and denominator after extracting the source dependence. However there are cases where one needs to know the numerator (or denominator) separately, for example, when one uses the path integral representation of the partition function. Since the free field integral is simply gaussian, we can in fact calculate it.

We start by noting that the general multivariable gaussian ordinary integral is given by

$$\int_{-\infty}^{\infty}\prod_{i=1}^{N}\left(\frac{dx_i}{\sqrt{\pi}}\right)e^{-\frac{1}{2}\sum_{km}x_kM_{km}x_m}=\prod_i m_i^{-1/2}=\det{}^{-1/2}M,$$

where m_i are the eigenvalues of the real symmetric matrix M. If we always define Euclidean functional integrals in terms of a lattice, this result can be directly applied. Then after taking the continuum limit, we can write for the neutral scalar field

$$\int\mathcal{D}\phi\exp\left\{-\int d^4x\left[\frac{1}{2}(\partial\phi)^2+\frac{m^2}{2}\phi^2\right]\right\}\equiv\det{}^{-1/2}[m^2-\partial^2].$$

A charged scalar field can be decomposed $\phi=(\phi_1+i\phi_2)/\sqrt{2}$ in terms of two real scalar fields so the corresponding formula is

$$\int\mathcal{D}\phi\mathcal{D}\phi^*\exp\left\{-\int d^4x_E\left[(\partial\phi^\dagger\partial\phi)+m^2\phi^\dagger\phi\right]\right\}\equiv\det{}^{-1}[m^2-\partial^2].$$

We have already encountered determinants of differential operators in our study of external field problems, for example, the *outin* matrix element for a charged scalar field in the presence of an external gauge field is proportional to $\det^{-1}[m^2-(\partial-iQA)^2]$. Thus we can immediately write the path history representation for this matrix element:

$$\langle out|in\rangle_A=\frac{\int\mathcal{D}\phi\mathcal{D}\phi^*\exp\left\{-\int d^4x_E\left[(\partial+iQA)\phi^*(\partial-iQA)\phi+m^2\phi^*\phi\right]\right\}}{\int\mathcal{D}\phi\mathcal{D}\phi^*\exp\left\{-\int d^4x_E\left[\partial\phi^*\cdot\partial\phi+m^2\phi^*\phi\right]\right\}}.$$

Since this is what must be inserted into the gauge field path integral to couple gauge fields to charged fields, this completes the process for converting to path integration language the qft of scalar fields interacting with

gauge fields. To get *outin* matrix elements of time ordered products of fields we use the generating functional

$$\begin{aligned}
&\langle out|T[e^{i\int d^4x(J^*\phi+\phi^\dagger J)}]\,|in\rangle \\
&=\frac{\int \mathcal{D}\phi\mathcal{D}\phi^* \exp\left\{-\int d^4x_E\left[(\partial+iQA)\phi^*(\partial-iQA)\phi+m^2\phi^*\phi-J^*\phi-\phi^*J\right]\right\}}{\int \mathcal{D}\phi\mathcal{D}\phi^* \exp\left\{-\int d^4x_E\left[\partial\phi^*\cdot\partial\phi+m^2\phi^*\phi\right]\right\}} \\
&=\frac{\det{}^{-1}(m^2-D^2)}{\det{}^{-1}(m^2-\partial^2)}e^{\int d^4xd^4yiJ^*(x)\Delta_F(x,y;A)iJ(y)}.
\end{aligned}$$

We have written the functional integrals in Euclidean space, which we indicate with the E subscript. To express the results in terms of Minkowski space, just use $d^4x_E = id^4x$. Remember that $\Delta_F(x,y;A)$ is precisely the continuation back to Minkowski space of the Euclidean Green function $\Delta_E(x_E,y_E;A)$. To confirm the last equality one simply completes the square and changes variables in the by now familiar way. Doing this in Euclidean space leads to the exponent $\int d^4x_Ed^4y_EJ^*(x)\Delta_E(x_E,y_E;A)J(y)$, which, when continued back to Minkowski space, gives the result shown.

We know from our experience with the operator formulation that the boundary conditions that fix Δ_F in Minkowski space are the requirements that it contains only positive frequencies at very late times and only negative frequencies at very early times. In Euclidean space, the boundary condition on Δ_E is simply that it vanish at infinity. Since Δ_F, Δ_E are analytic continuations of one another, either statement of the boundary conditions is satisfactory. From the path integral point of view it is more natural to choose the Eucldean version.

<u>Anticommuting Fields</u> To convert the persistence amplitude for the Dirac field in the presence of external gauge fields to a path integral formalism, we must be able to produce a factor of $\det\{m-i\gamma\cdot D\}$ in the *numerator* unlike the denominator as gaussian integrals tend to produce. This requires the introduction of "anti-commuting" numbers or a numbers in contrast to ordinary c numbers. Any two a numbers e, f satisfy $ef+fe=0$. In particular the square of an a number vanishes! Thus the most general function of a single a number a is the linear one c_1+c_2a. The theory of functions of a numbers is quite trivial. A function of N a numbers is at most linear in each variable, but that involves terms with up to N factors. With $N\to\infty$ one can of course have any number of factors as long as each factor is a different a number.

How do we integrate over a numbers? To define this we define $\int daf(a)$ to be a linear operation that assigns a unique c number to each function f. We also require the fundamental translation property $\int daf(a+e)=\int daf(a)$. But there are only two linearly independent functions of a namely 1 and a itself. So we only need to specify $\int da1$ and $\int daa$. The translation property for the second of these holds only if $\int da1=0$. Thus we only need to specify $\int daa$ to be some fixed c number, and then integration is completely defined! By definition we take $\int daa=1$. Then the integral of the arbitrary function c_1+c_2a is simply c_2, the coefficient of the linear term in a.

Now consider integration of a gaussian over 2M a numbers, $a_k, \bar{a}_k$:

$$\begin{aligned}
&\int da_1d\bar{a}_1\cdots da_Md\bar{a}_Me^{\bar{a}^TCa} \\
&=\frac{1}{M!}\int da_1d\bar{a}_1\cdots\int da_Md\bar{a}_M(\bar{a}^TCa)^M \\
&=\int da_1d\bar{a}_1\cdots\int da_Md\bar{a}_M\bar{a}_1C_{1k_1}a_{k_1}\bar{a}_2C_{2k_2}a_{k_2}\cdots\bar{a}_MC_{1k_M}a_{k_M} \\
&=C_{1k_1}C_{2k_2}\cdots C_{1k_M}\epsilon^{k_1k_2\cdots k_M}=\det C.
\end{aligned}$$

This shows that integrating gaussians over anticommuting numbers yields determinants with positive powers.

Next consider the case where the exponent of the gaussian is a bilinear in M a numbers b_k: $\frac{1}{2}b^T Ab$ where A is an antisymmetric matrix. It can be reduced to the previous case by considering its square:

$$\left(\int db_1 \cdots db_M e^{b^T Ab/2}\right)^2 = \int db_1 \cdots db_M db'_1 \cdots db'_M e^{(b^T Ab + b'^T Ab')/2}$$
$$=(-)^{\frac{M^2}{2}} \int da_1 d\bar{a}_1 \cdots da_M d\bar{a}_M e^{\bar{a}^T Aa}$$
$$=(-)^{\frac{M^2}{2}} \det A$$

where the change of variables $a = (b+ib')/\sqrt{2}, \bar{a} = (b-ib')/\sqrt{2}$ has been used. The phase out front is never relevant, since when M is odd, both sides vanish. We conclude that

$$\int db_1 \cdots db_M e^{b^T Ab/2} = det^{1/2} A.$$

Since the l.h.s. is a polynomial in the matrix elements of A we have proved an interesting corollary that the determinant of an antisymmetric matrix is the square of a polynomial in the matrix elements. That polynomial is sometimes known as the Pfaffian, $Pf(A) = det^{1/2}[A]$ for antisymmetric matrices A.

Now we can repeat our discussion of bosonic functional integrals for the fermionic case:

$$\langle out|in\rangle_A = \frac{\det(m - i\gamma \cdot D)}{\det(m - i\gamma \cdot \partial)}$$
$$= \frac{\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp\left\{-i \int d^4x (\bar{\psi}(m - i\gamma \cdot D)\psi\right\}}{\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp\left\{-i \int d^4x (\bar{\psi}(m - i\gamma \cdot \partial)\psi\right\}}.$$

Introducing anticommuting sources, $\eta, \bar{\eta}$ in the combination $i \int d^4x[\bar{\eta}\psi + \bar{\psi}\eta]$ in the exponent and competing the square, one can easily see that the source dependence is just a factor

$$e^{i \int \bar{\eta}(m - i\gamma \cdot D)^{-1}\eta} = e^{\int d^4x d^4y i\bar{\eta}(x) S_F(x,y;A) i\eta(y)}.$$

By differentiating with respect to the sources, one can show that this expression is just the generating function for the outin matrix element of time ordered products of fields:

$$\langle out| Te^{i \int d^4x[\bar{\eta}\psi + \bar{\psi}\eta]} |in\rangle$$
$$= \frac{\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp\left\{-i \int d^4x \left[(\bar{\psi}(m - i\gamma \cdot D - \bar{\eta}\psi - \bar{\psi}\eta)\psi\right]\right\}}{\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp\left\{-i \int d^4x (\bar{\psi}(m - i\gamma \cdot \partial)\psi\right\}}$$
$$= \frac{\det(m - i\gamma \cdot D)}{\det(m - i\gamma \cdot \partial)} e^{\int d^4x d^4y i\bar{\eta}(x) S_F(x,y;A) i\eta(y)}.$$

Gauge Fields The path integral for gauge theories presents special problems because of the gauge invariance of the action. We shall establish the proper formulation for Quantum Electrodynamics by first obtaining the path integral in Coulomb gauge $\nabla \cdot \mathbf{A} = 0$, following the general procedure sketched in this chapter. Once that has been done, we can discuss within the new path integral formalism more general gauges, including covariant ones.

The Hamiltonian operator for charged fields interacting with the quantized e.m. field $\mathbf{A}$ in Coulomb gauge is (as we shall see in the next chapter)

$$\begin{aligned} H_{eff} = \int d^3x \ \left(\frac{1}{2}\mathbf{\Pi_T}^2 + \frac{1}{2}\partial_k \mathbf{A}_T \cdot \partial_k \mathbf{A}_T - \mathbf{A}_T \cdot \mathbf{J}_e\right) \\ + \int (J_e^0 + j^0)\left(-\frac{1}{2\nabla^2}\right)(J_e^0 + j^0) + H_{fields}(\mathbf{A})|_{A_0=0}, \end{aligned} \tag{5.6}$$

Here we understand all operators to be in Schrödinger picture. We see that this is a case where the conjugate momentum appears only quadratically so we can immediately transcribe an infinitesimal evolution between eigenstates of $\mathbf{A}$ near imaginary time τ

$$\begin{aligned} &\langle \mathbf{A}''|e^{-d\tau H_{eff}}|\mathbf{A}'\rangle \\ &\quad \approx \delta(\nabla \cdot \mathbf{A}'') \exp\{d\tau \int d^3x[\frac{1}{2}(\dot{\mathbf{A}})^2(\tau) + \frac{1}{2}(\partial \mathbf{A})^2(\tau) - \mathbf{A}(\tau) \cdot \mathbf{J}_e)] \\ &\quad + d\tau \int (J_e^0 + j^0)\left(-\frac{1}{2\nabla^2}\right)(J_e^0 + j^0) + d\tau H_{fields}(\mathbf{A}(\tau)|_{A_0=0}\}, \end{aligned}$$

where the terms involving $\mathbf{A}(\tau)$ with no time derivative are averaged over $\mathbf{A}(\tau) = \mathbf{A}'', \mathbf{A}'$ and $\dot{\mathbf{A}}(\tau)$ symbolizes the continuum limit of $(\mathbf{A}'' - \mathbf{A}')/d\tau$. The delta function multiplying the r.h.s. symbolizes the condition $\nabla \cdot \mathbf{A} = 0$. It is necesaary if we wish to formally integrate over all three components of $\mathbf{A}$. Notice that for the moment we are only changing the e.m. field into path integral language, leaving the charged fields as operators.

Next we employ a widely used trick for simplifying the term describing the instantaneous Coulomb interaction. This involves introducing an auxiliary variable $A^4(\mathbf{x}, t)$ and writing the identity

$$e^{-d\tau \int (J_e^0 + j^0)(-\frac{1}{2\nabla^2})(J_e^0 + j^0)} = \frac{\int \mathcal{D}A^4(\tau) \exp\left\{-d\tau \int d^3x[\frac{1}{2}(\nabla A^4)^2 - A^4(J_e^4 + j^4)]\right\}}{\int \mathcal{D}A^4(\tau) \exp\left\{-d\tau \int d^3x \frac{1}{2}(\nabla A^4)^2\right\}}$$

where $j^4 = ij^0$ and $J_e^4 = iJ_e^0$ as appropriate after the Wick rotation. This identity is proved by completing the square by a shift of integration variable $A^4 \to A^4 - (1/\nabla^2)j^4$. Here we use the fact that ultimately $d\tau \to 0$ to neglect any commutators that might arise due to the fact that j^0 is really an operator. (This latter approximation is in exactly the same spirit as the replacement of $\mathcal{H}_k$ by H_k^W in our general derivation of the path integral.) The variable A^4 enters now exactly as the imaginary time component of the vector potential would have entered before it was eliminated by solving the Gauss' law constraint. For example, using $\nabla \cdot \mathbf{A} = 0$ we can write

$$\int d^3x[\frac{1}{2}\dot{\mathbf{A}}^2 + \frac{1}{2}(\nabla \mathbf{A}_k)^2 - \mathbf{A} \cdot \mathbf{J}_e) + \frac{1}{2}(\nabla A^4)^2 - A^4 J_e^4] = \int d^3x[\frac{1}{4}F_{\mu\nu}F_{\mu\nu} - A_\mu J_e^\mu].$$

Furthermore $H_{fields}|_{A^4=0} - \int d^3x A^4(\tau) j^4 = H_{fields}$ with $A^4 \neq 0$ playing the role of the imaginary time component of the vector potential.

Composing many infinitesimal evolutions, we arrive at the path integral representation

$$\langle \mathbf{A}_f | U(\infty, -\infty) | \mathbf{A}_i \rangle =$$

$$\frac{\int \mathcal{D}A_\mu \delta(\nabla \cdot \mathbf{A}) e^{-\int d^4x (F_{\mu\nu}F_{\mu\nu}/4 - A_\mu J_e^\mu)} \langle f | T e^{-\int d\tau H_{fields}(A_\mu(\tau))} | i \rangle}{\int \mathcal{D}A^4 e^{-\int d\tau (\nabla A^4)^2/2}}.$$

As discussed before, by taking the evolution over an infinite time interval we effectively project onto the ground states (vacuum) at early and late times. We can therefore pick convenient initial and final states, *e.g.* we can let them be eigenstates of $\mathbf{A}$ with vanishing eigenvalue and ground states of $H_{fields}(A_\mu(\tau = \mp\infty), \mp\infty)$ respectively. We then obtain the following general formula for the *outin* matrix element of time ordered products of fields

$$\langle out | T[\psi(y_1) \cdots \psi(y_m) \bar{\psi}(z_m) \cdots \bar{\psi}(z_1)] | in \rangle_{J_e, A_e} =$$

$$\frac{\int \mathcal{D}A_\mu \delta(\nabla \cdot \mathbf{A}) e^{-\int d^4x (F_{\mu\nu}F_{\mu\nu}/4 - A_\mu J_e^\mu)} \langle out | T[\psi(y_1) \cdots \bar{\psi}(z_1)] | in \rangle_{A(\tau)+A_e(\tau)}}{\int \mathcal{D}A_\mu \delta(\nabla \cdot \mathbf{A}) e^{-\int d^4x F_{\mu\nu}F_{\mu\nu}/4} \langle out | in \rangle_{A(\tau)}}.$$

An important observation is that the set of operators in the time ordered product on the l.h.s. can be expanded to include any number of vector potentials by functionally differentiating w.r.t. the external current J_e^μ. This is clear from the pure exponential dependence on J_e on the r.h.s. Thus in addition to describing the presence of real external sources, the J_e dependence provides a generating function for all correlation functions of any number of vector potentials in the source free case: simply set $J_e = 0$ *after* differentiating the appropriate number of times w.r.t. J_e. To illustrate this point, consider the path integral for the free e.m. field (no charged fields).

$$\langle out | in \rangle_{J_e} = \frac{\int \mathcal{D}A_\mu \delta(\nabla \cdot \mathbf{A}) e^{-\int d^4x (F_{\mu\nu}F_{\mu\nu}/4 - A_\mu J_e^\mu)}}{\int \mathcal{D}A_\mu \delta(\nabla \cdot \mathbf{A}) e^{-\int d^4x F_{\mu\nu}F_{\mu\nu}/4}}.$$

We can extract the dependence on J_e by shifting the integration variable in the numerator by $A_\mu \to A_\mu + C_\mu$, where C is chosen so that the linear term in A is cancelled. It must of course also be restricted by the Coulomb gauge condition $\nabla \cdot \mathbf{C} = 0$. After the shift the coefficient of the linear term in A becomes after an integration by parts

$$\partial_\mu (\partial_\mu C_\nu - \partial_\nu C_\mu) + J_\nu = 0.$$

This equation is only consistent if the external current is conserved $\partial_\mu J^\mu = 0$ which we are assuming. For $\nu = 4$, remembering the Coulomb gauge condition, this equation determines $C^4 = (-1/\nabla^2) J^4$. After using current conservation to write $\partial_4 C^4 = (1/\nabla^2) \nabla \cdot \mathbf{J}_e$ the spatial components are determined to be $\mathbf{C} = (-1/\partial^2)(\mathbf{J} - \nabla(1/\nabla^2)\nabla \cdot \mathbf{J})$. The question of boundary conditions is settled in Euclidean space by requiring that C vanish in all four directions at infinity. As we have discussed this prescription becomes the familiar $i\epsilon$ one when continued back to Minkowski space. Inserting these results into the path integral we find

$$\begin{aligned}\langle out | in \rangle_{J_e} =& e^{-\int d^4x (F^C_{\mu\nu}F^C_{\mu\nu}/4 - C_\mu J_e^\mu)} \\ =& e^{\int d^4x C_\mu J_e^\mu/2} \\ =& e^{\frac{1}{2}\int d^4x d^4y J_e^\mu(x) D_{\mu\nu}(x-y) J_e^\nu(y)}.\end{aligned}$$

where

$$D_{jk}(x) = \int \frac{d^4p}{(2\pi)^4} e^{ix\cdot p} \frac{\delta_{jk} - p^j p^k/\mathbf{p}^2}{p^2}$$
$$D_{44}(x) = \int \frac{d^4p}{(2\pi)^4} e^{ix\cdot p} \frac{1}{\mathbf{p}^2} = \delta(t)\frac{1}{4\pi|\mathbf{x}|}$$
$$D_{4k} = D_{k4} = 0.$$

To get correlation functions of any number of vector potentials we differentiate $\langle out|in\rangle_{J_e}$ with respect to the current any number of times. Because of current conservation we may only differentiate w.r.t. three components of J^μ say the two transverse components $\mathbf{J}_T$ and J^4. But that is sufficient since the longitudinal component of $\mathbf{A}$ is zero by the gauge condition. If we set $J_e = 0$ after differentiating, it is easy to check that the Wick expansion follows with the two point function

$$\langle A_\mu(x) A_\nu(y)\rangle = D_{\mu\nu}(x-y).$$

Note that with path integrals it is more precise to speak of correlation functions, which are functional averages of some number of fields, than of the vacuum expectations of time ordered products of field operators. They are of course numerically equal to each other.

General Gauges The path integral formulation of gauge theories is particularly suited to the discussion of general gauges. To keep the formalism covariant we would like to be able to replace the noncovariant Coulomb gauge condition with a covariant one such as $\partial_\mu A_\mu = 0$. The procedure for gauge fixing that follows is due to Fadeev and Popov, and is quite general. The method starts by selecting some general gauge fixing condition $F(A, \partial A) = 0$. This condition should have the property that for any value of the gauge potential it is possible to find a gauge transformation to a potential for which $F = 0$, and further that if $F(A) = 0$ then $F(A_\Omega) \neq 0$ with Ω any nontrivial gauge transformation which vanishes at infinity. We don't require this property for more general gauge transformations because that would rule out Coulomb gauge which seems to be perfectly adequate. The path integrand in such a gauge should contain a factor of $\delta(F(A))$.

The F-P procedure is to define a functional $\Delta_F(A)$ by the requirement

$$1 = \Delta_F(A) \int \mathcal{D}\Omega \delta(F(A_\Omega))$$

where A_Ω is the transformation of A by the gauge group element $\Omega(x)$, and the measure $\mathcal{D}\Omega$ is gauge invariant. By this definition Δ_F is clearly gauge invariant. Now insert this representation for 1 in the "unfixed" gauge field path integrand. Next change functional integration variables so that $A_\Omega \to A$. Here we implicitly assume that the unfixed measure $\mathcal{D}A_\mu$ is invariant under changes of variables which are gauge transformations*. Then the infinite volume of the gauge group $\int \mathcal{D}\Omega$ comes out as a common factor in both the numerator and denominator of the functional average and so cancels. We are left with the factors

$$\Delta_F(A)\delta(F(A))$$

in the functional integrand. The delta function fixes the gauge and the factor Δ_A is in general needed to guarantee that different choices for F yield the same answer for gauge invariant quantities. (It is only for functional averages of gauge invariant quantities that the rest of the integrand stays invariant under the variable change that removes Ω from the delta function.)

$\star$ In a completely general context this assumption might clash with the more basic translational invariance of the measure. In such a case there would be a Jacobian accompanying the variable change. This is not needed for the usual abelian and nonabelian gauge theories however.

Our Coulomb gauge path integral did not include a factor of $\Delta_{Coul}(A)$ so we need to confirm that it is inconsequential for the abelian case. For this we have to compute

$$\frac{1}{\Delta_{Coul}} = \int \mathcal{D}\Lambda \delta(\nabla \cdot \mathbf{A} + \nabla^2 \Lambda) = \frac{1}{\det(-\nabla^2)}.$$

We see that Δ_{Coul} is independent of A so that it will cancel between the numerator and denominator of functional averages. Thus our failure to put it in gives no differences in physical quantities. Thus all gauge choices are formally equivalent to Coulomb gauge provided the Fadeev-Popov factor $\Delta_F(A)$ is included along with the gauge fixing delta functional.

In practice $\Delta_F(A)$ always multiplies $\delta(F(A))$ so the former is only needed for A satisfying the gauge condition. For such A, $\delta(F(A_\Omega))$ only contributes for infinitesimal Ω:

$$F(A(x) + \delta A(x)) \approx \int d^4y \frac{\delta F(A(x))}{\delta A_{\mu ab}(y)} \delta A_{\mu ab}(y).$$

Denoting the infinitesimal generators of the gauge group by $G(z)$, δA is linear in the matrix elements of G: $\delta A_{\mu ab}(y) = \int d^4z \mathcal{L}_{\mu ab,cd}(y,z;A) G_{cd}(z)$ where $\mathcal{L}$ is a linear differential operator depending on A in general. For example, the infiniesimal nonabelian gauge transformation reads $\delta A = \partial G - ig[A,G]$. Thus $\Delta_F = \det(\frac{\delta F}{\delta A_\mu}\mathcal{L}_\mu)$ where the determinant is that of a linear differential operator which is also a matrix in the internal group space with matrix elements labelled (ab, ef):

$$\int d^4y \frac{\delta F_{ab}(A(x))}{\delta A_{\mu cd}(y)} \mathcal{L}_{\mu cd,ef}(y,z;A).$$

An important class of covariant gauges consists of the Lorentz gauges $\partial \cdot A = f$ where f is some fixed function. The F-P determinant for this gauge is $\det(-\partial^2)$ in the abelian case of QED. Notice that it is independent of both f and A. In the nonabelian case, it would depend upon A of course, but is still independent of f. So for QED the F-P determinant cancels between numerator and denominator. and it is safe to ignore it†. Functional averages of gauge invariant quantities will be independent of f. We can exploit this to get rid of the functional delta function in the path integrand by averaging over f with a gaussian weight function $e^{-\int d^4x f^2(x)/2\alpha}$. This just multiplies the numerator and denominator by the same constant so it won't alter gauge invariant quantities. The net effect of this is to remove the delta function and instead add a non gauge invariant term $\int d^4x (\partial \cdot A)^2(x)/2\alpha$ to the Euclidean action. For $\alpha = 1$ (Feynman Gauge) the effect of this term is to simplify the kinetic term for A:

$$\int d^4x \left(\frac{1}{4}F_{\mu\nu}F_{\mu\nu} + \frac{1}{2\alpha}(\partial \cdot A)^2\right) = \int d^4x \left(\frac{1}{2}\partial_\mu A_\nu \partial_\mu A_\nu + \frac{1-\alpha}{2\alpha}(\partial \cdot A)^2\right)$$
$$\underset{\alpha \to 1}{\sim} \int d^4x \frac{1}{2}\partial_\mu A_\nu \partial_\mu A_\nu$$

Just as with the Coulomb gauge, the photon propagator for the "α" gauges is obtained by evaluating the path integral for the gauge field in the presence of an external source. The source dependence is easily

† There are some applications where it is nonethless important to keep it. For example, it contributes a constant to the zero point energy which just subtracts the spurious contribution of the time-like and longitudinal components of the vector potential. Also when one computes the finite temperature partition function in a covariant gauge, the F-P factor removes two photon degrees of freedom so the total number is $4 - 2 = 2$. Without the F-P factor the Stefan-Boltzmann law would be off by a factor of 2!

obtained by shifting $A \to A + C$ with C satisfying

$$\partial^2 C_\mu + \frac{1-\alpha}{\alpha}\partial\partial\cdot C = -J^\mu$$

which is solved first for $\partial \cdot C = -\alpha \cdot J$ and then

$$C_\mu = -\frac{1}{\partial^2}(J_\mu - (1-\alpha)\frac{\partial_\mu \partial_\nu}{\partial^2}J^\nu).$$

Thus the propagator is

$$D^\alpha_{\mu\nu}(x) = \int \frac{d^4p}{(2\pi)^4} e^{ip\cdot x}\frac{\delta_{\mu\nu} - (1-\alpha)p_\mu p_\nu/p^2}{p^2}.$$

In the continuation to Minkowski space $\delta_{\mu\nu}$ is replaced by $\eta_{\mu\nu}$, p^2 by $p^2 - i\epsilon$, and there is an additional factor of $-i$. We note the great simplification for Feynman gauge. The case $\alpha = 0$, known as Landau gauge is effectively the gauge $\partial \cdot A = 0$ because the coefficient of the gauge breaking term in the action blows up damping out all contributions to the integral not satisfying this condition. We see that in this case the divergence of the propagator vanishes, $\partial_\mu D^0_{\mu\nu} = 0$. The fundamental reason that QED is independent of α is that A always couples to a conserved current so that the terms involving α in the propagator decouple.

To get a bit more insight into the role of current conservation consider a photon propagator attached to two conserved vertices in momentum space

$$Amp \sim \frac{R^1_\mu R^{2\mu}}{p^2} = \frac{\mathbf{R}^1\cdot\mathbf{R}^2 - R^1_0 R^2_0}{p^2}.$$

It would seem that this coupling corresponds to the propagation of four photon states, 3 "space" components and 1 time component, the last one coupling with the "wrong" sign. These wrong sign states are sometimes called "ghosts." But by current conservation $R^a_0 = \mathbf{p}\cdot\mathbf{R}^a/p^0$. In the limit of physical photon momentum,*i.e.*for which $p^2 = 0$ the residue of the pole is just

$$\mathbf{R}^1\cdot\mathbf{R}^2 - \frac{\mathbf{p}\cdot\mathbf{R}^1\mathbf{p}\cdot\mathbf{R}^2}{\mathbf{p}^2}$$

which is to say only the transverse states (perpendicular to $\mathbf{p}$) truly propagate.

6. The Dirac Equation

In our development of relativistic quantum mechanics we shall study intensively the case of spin 1/2 particles solved by Dirac. After this we shall return to the problem of higher spin. As a practical matter, the Dirac equation contains a tremendous amount of the present understanding of elementary particles. In fact, all of the known fundamental constituents of matter are either spin 1/2 (quarks and leptons), or spin 1 gauge particles (photon, gluons, weak vector bosons). To these we may have to add a fundamental scalar (Higgs particle) and of course there is always the spin 2 graviton that must be understood eventually. But we can go a tremendous distance to understanding the physics of the standard model through study of the Dirac equation interacting with a gauge field. Indeed, for many applications the gauge field doesn't even need to be quantized!

Dirac's original motivation for his equation is still useful and inspiring to recall. The most straightforward attempt to write down a relativistic version of the Schrödinger equation

$$\begin{aligned} i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} &= \sqrt{m^2c^4-(\hbar c)^2\nabla^2}\psi(\mathbf{x},t) \\ &\equiv m\sum_{k=0}^{\infty}\binom{1/2}{n}\left(\frac{-\hbar^2\nabla^2}{m^2}\right)^n\psi(\mathbf{x},t) \end{aligned} \tag{6.1}$$

is horribly nonlocal, involving an arbitrarily high number of spatial derivatives. Dirac proposed to get around this by making ψ a multicomponent wave function and defining the square root in a local way by using matrices, in analogy with the properties of the Pauli spin matrices

$$(\boldsymbol{\sigma}\cdot\mathbf{v})^2 = \mathbf{v}^2.$$

To achieve this he required the introduction of four anticommuting matrices γ^μ with $\mu = 0,1,2,3$:

$$\{\gamma^\mu,\gamma^\nu\} = -2\eta^{\mu\nu}. \tag{6.2}$$

Using (6.2) it is simple to show that

$$(\gamma^\mu\partial_\mu)^2 = \frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu = -\partial_\mu\partial^\mu = \frac{\partial^2}{\partial t^2}-\nabla^2,$$

where we have chosen units for which $c = 1$. We shall also choose units so that $\hbar = 1$. Then the Dirac equation is

$$\frac{1}{i}\gamma^\mu\partial_\mu\psi + m\psi = 0. \tag{6.3}$$

To cast this equation as a relativistic Schrödinger equation we rewrite it as

$$i\frac{\partial\psi(\mathbf{x},t)}{\partial t} = \left(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla+\beta m\right)\psi(\mathbf{x},t)$$

where we have multipied through by $\gamma^0\equiv\beta$ using $\beta^2 = I$ and have defined $\boldsymbol{\alpha}\equiv\gamma^0\boldsymbol{\gamma}$.

So far we have not specified the gamma matrices. It is simple to show that they must be at least 4×4 (In D space-time dimensions the minimum size is $2^{D/2}\times 2^{D/2}$ for even D.)

Let σ^k, $k = 1, 2, 3$ be the 2×2 Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then we have two popular representations:

Standard Representation of γ^μ

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

$$\boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}$$

$$\gamma_5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

$$\sigma_{km} = \epsilon_{kmn} \begin{pmatrix} \sigma^n & 0 \\ 0 & \sigma^n \end{pmatrix}$$

Chiral (Natural) Representation of γ^μ

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

$$\boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}$$

$$\gamma_5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$$

$$\sigma_{km} = \epsilon_{kmn} \begin{pmatrix} \sigma^n & 0 \\ 0 & \sigma^n \end{pmatrix}$$

where ϵ_{kmn} is the completely antisymmetric three tensor with $\epsilon_{123} = +1$. The standard representation is more convenient for slowly moving particles, whereas the chiral one is more conveneient for massless fermions that move at the speed of light.

Clearly $\beta = \gamma^0$ is hermitian, $\boldsymbol{\gamma}$ is anti-hermitian, and $\boldsymbol{\alpha}$ is hermitian. Thus the Hamiltonian

$$H = \frac{1}{i}\boldsymbol{\alpha} \cdot \nabla + \beta m \tag{6.4}$$

is a hermitian operator as it should be.

To justify the interpretation of the Dirac particle as spin 1/2, we need to construct an angular momentum operator for which $\boldsymbol{\alpha}$ transforms as a vector. This is clearly accomplished by choosing the spin operator to be $\boldsymbol{\Sigma}/2$ with

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}.$$

Then the Dirac Hamiltonian obviously commutes with the total angular momentum

$$\mathbf{J} = \mathbf{r} \times \frac{1}{i}\nabla + \frac{\boldsymbol{\Sigma}}{2}.$$

The Dirac wave function transforms under the $\frac{1}{2} \oplus \frac{1}{2}$ representation of the rotation group. In particular it describes a spin 1/2 particle.

To get the energy spectrum of the Dirac particle we must find all the eigenstates of the Hamiltonian

(6.4). Clearly, it is best to work with momentum eigenstates

$$\psi_{\mathbf{p}} = u(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}},$$

since the momentum operator commutes with H. Then the coefficient spinor must satisfy

$$(\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m)u(\mathbf{p}) = Eu(\mathbf{p}). \tag{6.5}$$

Since the square of the matrix on the l.h.s. is just $m^2 + \mathbf{p}^2$, we immediately learn that

$$E(\mathbf{p}) = \pm\sqrt{m^2 + \mathbf{p}^2}.$$

It is also fairly simple to find the eigenfunctions in the Standard Representation. It is sufficient to find the positive energy eigenfunctions, because the negative energy ones can be obtained from them by a simple operation: Complex conjugating (6.5) and multiplying both sides by $i\gamma^2$, we learn that $i\gamma^2 u^*(-\mathbf{p})$ is an eigenstate of H with eigenvalue $-E$ if $u(\mathbf{p})$ is an eigenstate with eigenvalue $+E$. An explicit solution for u, with $E = +\omega(\mathbf{p}) > 0$, is easily shown to be

$$u(\mathbf{p}) = \begin{pmatrix} \phi \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{m+\omega(\mathbf{p})}\phi \end{pmatrix} \tag{6.6}$$

where $\omega = \sqrt{m^2 + \mathbf{p}^2}$ and ϕ is any two component spinor. The two independent components just represent the two spin states of a spin 1/2 particle. As we have explained the negative energy eigenfunctions are then

$$u_-(\mathbf{p}) = i\gamma^2 u^*(-\mathbf{p}). \tag{6.7}$$

The positive energy states of the Dirac particle give the desired relativistic description of a relativistic spin 1/2 particle. The negative energy states are a disaster for the single particle interpretation of the Dirac wave function. As soon as the Dirac particle in a positive energy state is subjected to external forces transitions to negative energy states will be induced and the system is unstable. Dirac himself proposed the remedy for this problem, which exploits the Pauli Exclusion Principle for fermions: If one postulates that the Dirac particle is a fermion, then one can consider the "vacuum" to be the state in which *all* of the negative energy levels are filled, then any further particles added to this state must by the exclusion principle occupy only positive energy states! This vacuum state is sometimes called the "negative energy sea." Although we call this state the vacuum, it is clear from its construction that it is far from empty, since it contains an infinite number of particles. The presence of the sea will make itself felt as soon as one considers interactions of the Dirac particles with themselves or with independent force fields. Strictly speaking, there is an enormous (infinitely) negative energy associated with the sea. However in quantum mechanics, only energy differences are measurable (in the absence of gravity) and we might as well measure the energy of all states relative to that of the sea, *i.e.* we take $E_{sea} = 0$. The total momentum contained in the sea is automatically zero because momentum states in all directions are occupied.

In addition if the Dirac particle is charged the sea possesses an infinite uniform positive charge density. It is convenient in this case to postulate a compensating negative background charge density, so our vacuum will be neutral

To understand the dynamical significance of the presence of the sea, it is instructive to consider the excitation spectrum of the whole system. We can first of all add N particles with momenta $\mathbf{p_1}, \mathbf{p_2}, \ldots, \mathbf{p_N}$ to the sea. The resulting state will have energy

$$E_{N\ particles} = \sum_k \sqrt{m^2 + \mathbf{p}_k^2}$$

and N units of charge.

Next we can *remove* N particles with momenta $\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_N}$ from the sea. This state, which we describe as a state of N holes, will also have a positive energy

$$E_{N\ holes} = \sum_k \sqrt{m^2 + \mathbf{q}_k^2}$$

since we have subtracted negative energy. However the momentum of this state is

$$\mathbf{P}_{N\ holes} = -\sum_k \mathbf{q}_k$$

for the same reason, and the charge is $-N$ units. Thus this state appears to be a state of N particles of opposite charge to the ordinary (positive energy) Dirac particle and with momenta $-\mathbf{q_1}, -\mathbf{q_2}, \ldots, -\mathbf{q_N}$. This is how antiparticles appear in the theory.

Finally, one can simply excite N particles from the sea to positive energy states. This excitation does not change the number of particles so it is neutral. The energy of this state is the sum of the positive energy eigenvalues occupied minus the sum of the negative energy levels vacated. It is a state with N particles and N antiparticles (holes), with corresponding energy. In principle, this state can be prepared from the sea by delivering sufficient energy, at least $2Nm$, to the system.

If the original Dirac particle is the electron with charge $-e$, then the theory predicts the existence of its antiparticle, the positron, with charge $+e$. The positron was unknown when Dirac discovered his equation, and the idea of identifying the positron with the proton failed because the latter does not have the electron's mass nor could it annihilate with an electron into photons. Thus when the positron was eventually discovered, it was a spectacular confirmation of Dirac's theory.

7. Second Quantization

We have seen that the consistent interpretation of the Dirac theory requires the presence of the sea of an infinite number of negative energy particles. In practice, however, at least if interactions are weak, all but a finite number of particles in the sea are spectators in any calculation. Thus we need a formalism that allows us to concentrate only on the part of the sea that is active in a given process. The formalism which achieves this has been developed, and for historical reasons goes under the name of second quantization. It should be stressed that it is completely equivalent to a description in terms of the many body Schrödinger wave function for the sea, but it is much less cumbersome, and almost indispensable to efficient calculation. The formalism is applicable to any system of identical particles, bosons or fermions. We shall present only the fermion case in lecture since we will only be applying it to the Dirac theory. The changes needed for the boson case will be indicated in a series of exercises.

We begin with the concept of the occupation number basis. Consider a system of fermions, and let an orthonormal complete set of states for a single fermion be labelled by an index α:

$$\psi_\alpha(\mathbf{x}).$$

Then a multi-fermion antisymmetrized tensor product state is competely specified by the set of numbers $\{n_\alpha\}$ which give the number of fermions (0 or 1) occupying each state α.

Now the crucial idea of second quantization is to define a creation operator $b_\alpha^\dagger$ by

$$b_\alpha^\dagger \left|n_1 \ldots n_\alpha \ldots\right\rangle = (-)^{\sum_{\gamma<\alpha} n_\gamma} \left|n_1 \ldots n_\alpha + 1 \ldots\right\rangle .$$

The state dependence of the prefactor is necessary for consistency with the antisymmetry of the state under fermion interchange. Since $b_\alpha^\dagger$ adds a fermion in state α its square must vanish. It is a direct consequence of this fact and its definition that

$$\{b_\alpha^\dagger, b_\beta^\dagger\} = 0$$

By considering

$$\langle\{n\}| b_\alpha^\dagger |\{n'\}\rangle^* = \langle\{n'\}| b_\alpha |\{n\}\rangle$$

it is clear that b_α removes a fermion from the state α and so is an annihilation operator. It is then simple to check that

$$\{b_\alpha, b_\beta^\dagger\} = \delta_{\alpha\beta}.$$

For our application, the state label will include a continuous momentum as well as a discrete label. In that case the Kronecker delta will of course include a factor of a Dirac delta function $\delta(\mathbf{p} - \mathbf{p}')$.

Having introduced the creation and annihilation operators, we now have a very efficient notation for the occupation number basis. Call the state with no states occupied $|0'\rangle$. Then the state with

$$n_{\alpha_1} = n_{\alpha_2} = \cdots n_{\alpha_N} = 1$$

and all other occupation numbers zero is just

$$b_{\alpha_N}^\dagger \cdots b_{\alpha_2}^\dagger b_{\alpha_1}^\dagger \left|0'\right\rangle .$$

We are now in a position to relate all this formalism to the standard many-body Schrödinger wave function description. The wave function describing the state in which the single particle states $\alpha_1 \cdots \alpha_N$ are

occupied is just

$$\Psi_{\alpha_1\cdots\alpha_N}(\mathbf{x_1},\cdots,\mathbf{x_2}) = \frac{1}{\sqrt{N!}}\sum_P \delta_P \psi_{\alpha_1}(\mathbf{x}_{P(1)})\psi_{\alpha_2}(\mathbf{x}_{P(2)})\cdots\psi_{\alpha_N}(\mathbf{x}_{P(N)})$$

where $\delta_P = -1$ if P is an odd permutation and $+1$ if P is an even permutation. If we define the field operator

$$\psi(\mathbf{x}) \equiv \sum_\alpha b_\alpha \psi_\alpha(\mathbf{x}), \tag{7.1}$$

with anticommutation relations

$$\{\psi(\mathbf{x}),\psi(\mathbf{y})\} = \{\psi^\dagger(\mathbf{x}),\psi^\dagger(\mathbf{y})\} = 0$$
$$\{\psi(\mathbf{x}),\psi^\dagger(\mathbf{y})\} = \sum_\alpha \psi_\alpha(\mathbf{x})\psi^*_\alpha(\mathbf{y}) = \delta(\mathbf{x}-\mathbf{y}),$$

Then we have

$$\Psi_{\alpha_1\cdots\alpha_N}(\mathbf{x_1},\cdots,\mathbf{x_N}) = \frac{1}{\sqrt{N!}}\langle 0'|\,\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\cdots\psi(\mathbf{x}_N)b^\dagger_{\alpha_N}\cdots b^\dagger_{\alpha_2}b^\dagger_{\alpha_1}\,|0'\rangle\,.$$

A completely general N body wave function is of course a general superposition of such states

$$\Phi^N = \sum_{\alpha_1\cdots\alpha_N} c_{\alpha_1\cdots\alpha_N}\Psi_{\alpha_1\cdots\alpha_N} = \frac{1}{\sqrt{N!}}\langle 0'|\,\psi(\mathbf{x}_1)\psi(\mathbf{x}_2)\cdots\psi(\mathbf{x}_N)\,|\Phi\rangle$$

where

$$|\Phi\rangle = \sum_{\alpha_1\cdots\alpha_N} c_{\alpha_1\cdots\alpha_N} b^\dagger_{\alpha_N}\cdots b^\dagger_{\alpha_2}b^\dagger_{\alpha_1}\,|0'\rangle$$

The state $|\Phi\rangle$ thus can describe the most general many body Schrödinger wave function. Indeed, the second quantization formalism allows one to even superpose states of different numbers of particles, so it is the more general description!

We have shown how to describe general quantum states, it remains to show how linear operators are related between the two formalisms. All possible observables in the Schrödinger wave function description are completely symmetric in the operators acting on each particle. Such operators can be classified according to the number of particles involved in each term of the operator. For example, a one body operator acting on N identical particles has the form

$$\Omega^{(1)} = \sum_k \omega_k$$

where ω_k acts only on the variables of the kth particle. A two body operator would be a sum of terms, each acting on a pair of particles:

$$\Omega^{(2)} = \sum_{k<m} \omega_{km}.$$

A K body operator is obviously a sum of terms each acting on K particles. In a typical nonrelativistic system with N identical particles, *e.g.* the electrons of a $Z = N$ atom, the kinetic energy of the electrons is a 1 body operator, while the potential energy is a two body operator (since the electrons interact in pairs). The interaction of the electrons with an external field would be described by another one body operator.

The second quantized description of these operators is straightforward. Starting with the one body case we first represent the single particle operator ω by its matrix elements in the single particle basis we have introduced

$$\omega_{\alpha\beta} = \langle\alpha|\,\omega\,|\beta\rangle\,.$$

Then the one body operator $\Omega^{(1)}$ is just

$$\Omega^{(1)} = \sum_{\alpha\beta} b^\dagger_\alpha \omega_{\alpha\beta} b_\beta.$$

To see that this definition has the correct action on multi-particle states, one simply applies it to one of the basis states

$$\Omega^{(1)} b^\dagger_{\alpha_N} \cdots b^\dagger_{\alpha_2} b^\dagger_{\alpha_1} |0'\rangle = \sum_k b^\dagger_{\alpha_N} \cdots \sum_{\gamma_k} (b^\dagger_{\gamma_k} \langle\gamma_k|\,\omega\,|\alpha_k\rangle) \cdots b^\dagger_{\alpha_1} |0'\rangle$$

and we see that the r.h.s. is a sum of terms for the kth of which ω has the correct action on the kth single particle state label.

Multi-body operators have the obvious analogous second quantized description. We limit the discussion here to the two-body case. Let $|\beta_1\rangle\,|\beta_2\rangle$ denote the standard unsymmetrized two particle tensor product state. It is convenient to reverse the order of factors in the corresponding bra $\langle\alpha_2|\,\langle\alpha_1|$ so that

$$(\langle\alpha_2|\,\langle\alpha_1|)\,|\beta_1\rangle\,|\beta_2\rangle = \delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}.$$

Then we define

$$\omega_{\alpha_2\alpha_1,\beta_1\beta_2} = \langle\alpha_2|\,\langle\alpha_1|\,\omega\,|\beta_1\rangle\,|\beta_2\rangle$$

and the two-body operator $\Omega^{(2)}$ is then just

$$\Omega^{(2)} = \frac{1}{2} \sum_{\alpha_1\alpha_2,\beta_1\beta_2} b^\dagger_{\alpha_2} b^\dagger_{\alpha_1} \omega_{\alpha_2\alpha_1,\beta_1\beta_2} b_{\beta_1} b_{\beta_2}.$$

A typical example which requires both one and two body operators is the hamiltonian for a system of N nonrelativistic fermions with an interaction potential energy $V(\mathbf{r}_k, \mathbf{r}_m) = V(\mathbf{r}_m, \mathbf{r}_k)$ between each pair (k, m). The Schrödinger picture Hamiltonian for this system is just

$$H = \sum_k \frac{-\hbar^2}{2m}\nabla_k^2 + \sum_{k<m} V(\mathbf{r}_k, \mathbf{r}_m).$$

According to the procedure just outlined, the second quantized version of the kinetic term is

$$\sum_{\alpha\beta} b^\dagger_\alpha \langle\alpha|\left(\frac{-\hbar^2}{2m}\nabla_k^2\right)|\beta\rangle\, b_\beta = \int d^3r \psi^\dagger(\mathbf{r})\left(\frac{-\hbar^2}{2m}\nabla^2\right)\psi(\mathbf{r}) = \int d^3r \frac{\hbar^2}{2m}\nabla\psi^\dagger(\mathbf{r})\cdot\nabla\psi(\mathbf{r}),$$

where in the second form we have gone to the coordinate basis, using the definition of the second quantized field operator (7.1). The second quantized version of the potential term is the two-body operator

$$\frac{1}{2}\sum_{\alpha_1\alpha_2,\beta_1\beta_2} b^\dagger_{\alpha_2} b^\dagger_{\alpha_1} \langle\alpha_2|\,\langle\alpha_1|\,V\,|\beta_1\rangle\,|\beta_2\rangle\, b_{\beta_1} b_{\beta_2} = \frac{1}{2}\int d^3x d^3y V(\mathbf{x},\mathbf{y})\psi^\dagger(\mathbf{y})\psi^\dagger(\mathbf{x})\psi(\mathbf{x})\psi(\mathbf{y}).$$

Two details to note about this expression are the overall factor of 1/2 and the order of operators. These are necessary to arrange that the potential energy has the correct sign and normalization. The complete

hamiltonian for this system can now be compactly written as

$$H = \int d^3r \frac{\hbar^2}{2m} \nabla\psi^\dagger(\mathbf{r}) \cdot \nabla\psi(\mathbf{r}) + \frac{1}{2}\int d^3x d^3y V(\mathbf{x},\mathbf{y})\psi^\dagger(\mathbf{y})\psi^\dagger(\mathbf{x})\psi(\mathbf{x})\psi(\mathbf{y}).$$

This is the hamiltonian for a quantum field theory. The fundamental quantum operators are the local fields $\psi(\mathbf{x})$ or the corresponding creation and annihilation operators. The operators x, p etc. of the Schrodinger description have been demoted to c number labels and derivatives with respect to them.

One final feature of the formalism to explain is the role of the field equations. In the Schrödinger picture the quantum dynamics is given by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\Phi, t\rangle = H_S(t)|\Phi, t\rangle$$

where we stress that we allow time varying external forces to be present (hence the time dependence of H). The field equations arise in the Heisenberg picture wherein the time dependence resides in the operators rather than in the system states which are constant in time. To pass to the Heisenberg picture we write

$$|\Phi, t\rangle = U(t)|\Phi, 0\rangle$$

where

$$i\hbar\frac{\partial}{\partial t}U = H_S(t)U \qquad\qquad U(0) = I,$$

and give the time independent Schrödinger picture operators Ω time dependence according to

$$\Omega(t) \equiv U^\dagger \Omega U.$$

The Heisenberg picture Hamiltonian is similarly related to the Schrödinger picture one by

$$H(t) \equiv U^\dagger H_S(t) U.$$

Then the Heisenberg picture operators corresponding to constant Schrödinger picture operators satisfy the Heisenberg equations

$$i\hbar\dot{\Omega}(t) = [\Omega(t), H(t)].$$

Returning to our system of nonrelativistic fermions we find that the Heisenberg equation for the field operator ψ implies

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t) = \frac{-\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t) + \int d^3z V(\mathbf{x},\mathbf{z})\psi^\dagger(\mathbf{z})\psi(\mathbf{z})\psi(\mathbf{x}). \tag{7.2}$$

This is the quantum field equation. It is a nonlinear differential equation for a quantum operator, and because of its operator nature it has much more information packed in it than is immediately apparent. The origin of the name "second quantization" for this formalism is that (7.2) looks like a nonlinear version of the Schrödinger equation. Indeed if the particles did not mutually interact, it would be exactly the Schrödinger equation. If "first quantization" produced the Schrödinger equation, we have now reinterpreted the latter as a classical field equation, which is then "second quantized". Of course we know that all we have really done is given a clever reformulation of ordinary many body quantum mechanics, and the name is really a misnomer, which has stuck.

We should mention that (7.2) is an elegant point of departure for the Hartree-Fock approximation, in which one approximates the nonlinear term by a one body term

$$\int d^3y K(\mathbf{x},\mathbf{y})\psi(\mathbf{y}) \tag{7.3}$$

in which K is chosen to make this term as close as possible to the original two body term.

How much of the quantum dynamics is captured in the quantum field equation? The time dependent many body wave function in Schrödinger picture, which summarizes the complete quantum dynamics of the system can be recast

$$\frac{1}{\sqrt{N!}}\langle 0'|\,\psi(\mathbf{x}_1)\cdots\psi(\mathbf{x}_N)\,|\Phi,t\rangle = \frac{1}{\sqrt{N!}}\langle 0'|\,U(t)\psi(\mathbf{x}_1,t)\cdots\psi(\mathbf{x}_N,t)\,|\Phi,0\rangle\,.$$

So in addition to the time dependence of $\psi(\mathbf{x},t)$ which we could get from the field equation, we would also need to find the time dependence of the state

$$\langle 0'|\,U(t).$$

In our example, the state $\langle 0'|$ is really dynamically inert *i.e.*

$$\langle 0'|\,U(t) = \langle 0'|\,,$$

because each term of $H_S(t)$ has a $b^\dagger$ on the left so that $\langle 0'|\,H_S(t) = 0$. Thus knowledge of $\psi(\mathbf{x},t)$ at all times allows us to reconstruct the time dependence of the many body Schrödinger wave function completely. If we were transforming a different matrix element, in which $\langle 0'|$ were replaced by a nonempty state, to the Heisenberg picture, the time dependence of U acting on this state would have to be found in addition to the time dependence of the Heisenberg operators. The time dependence of $\langle A|\,U(t)$ has a simple interpretation if $\langle A|$ is characterized as an eigenstate, say, of some definite Schrödinger picture observable Ω. Then $\langle A|\,U(t)$ is the corresponding eigenstate of the corresponding Heisenberg picture operator $\Omega(t) = U^\dagger\Omega U$. In particular if the time dependent terms in the Schrödinger picture Hamiltonian $H_S(t) = H_0 + H'(t)$ vanish initially and finally,

$$H'(T) = H'(0) = 0,$$

and $|A\rangle$ is an eigenstate of $H_0 = H_S(0) = H(0)$, then $\langle A|\,U(T)$ is an eigenstate of the Heisenberg picture operator $H_0(T) = H(T)$.

Let us consider what is gained and lost in the alternative formulations of many body quantum physics. The second quantization formalism contains all the information contained in the (anti)symmetrized wave functions. The (unphysical) non-symmetrized wave functions are of course lost, but that is desirable. The second quantization machinery allows a broader range of dynamical options. For example, particle number conservation is built into the wave function description. This conservation law is reflected in the existence of the number operator

$$N = \sum_\alpha b_\alpha^\dagger b_\alpha = \int d^3x\psi^\dagger(\mathbf{x})\psi(\mathbf{x}) \tag{7.4}$$

which counts the number of particles in a given state. The conservation law is the statement that N commutes with the Hamiltonian: $[N,H] = 0$. We can also identify $\psi^\dagger(\mathbf{x})\psi(\mathbf{x})$ as the number density operator. There is, however, no principle which excludes considering a Hamiltonian with terms that don't commute with N. For example, a term

$$\int d^3x d^3x' A(\mathbf{x},\mathbf{x}')(\psi(\mathbf{x})\psi(\mathbf{x}') - \psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{x}'))$$

would not commute with N.

8. Second Quantized Dirac Equation

The formalism developed in the last chapter can now be fruitfully applied to the Dirac equation, cast as a Schrödinger equation,

$$i\frac{\partial\psi(\mathbf{x},t)}{\partial t} = \left(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m\right)\psi(\mathbf{x},t).$$

Let us regard the component label a as an additional coordinate: $\psi_a(\mathbf{x}) = \psi(\mathbf{x},a)$. For definiteness let us choose our single particle basis to be momentum eigenstates, with an additional label λ for spin and $\pm$ to distinguish positive and negative energy states. Thus the role of $\psi_\alpha(\mathbf{x})$ of the previous chapter will be played by

$$\psi^{(\pm)}_{\lambda,\mathbf{p}}(\mathbf{x},a) = \frac{1}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}u^a_{\lambda\pm}(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}.$$

The prefactors are conventional and with them in place the condition of orthonormality implies that

$$\sum_a u^{a*}_{\lambda'\pm'}u^a_{\lambda\pm} = 2\omega(\mathbf{p})\delta_{\lambda'\lambda}\delta_{\pm'\pm}.$$

We therefore write the explicit solution (6.6) for $u = u^+$ in the rescaled form

$$u_\lambda(\mathbf{p}) = \sqrt{\omega(\mathbf{p})+m}\begin{pmatrix}\phi_\lambda\\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{m+\omega(\mathbf{p})}\phi_\lambda\end{pmatrix}$$

so that the normalization condition on the two spinor ϕ is

$$\phi^\dagger_{\lambda'}\phi_\lambda = \delta_{\lambda'\lambda}.$$

There are two widely used choices for ϕ_λ. One is to simply choose the two orthogonal spinors

$$\begin{pmatrix}1\\0\end{pmatrix} \qquad \begin{pmatrix}0\\1\end{pmatrix}.$$

In the rest frame $\mathbf{p} = 0$ these are just eigenstates of Σ_3 with eigenvalues $+1, -1$ respectively. The other choice is to pick them to be eigenstates of helicity $h = \mathbf{p}\cdot\boldsymbol{\sigma}/2|\mathbf{p}|$ denoted $\chi_\lambda(\mathbf{p})$:

$$h\chi_\lambda(\mathbf{p}) = \lambda\chi_\lambda(\mathbf{p}) \qquad \lambda = \pm\frac{1}{2}.$$

Explicit forms for χ_λ are developed in the exercises. Notice that for the helicity basis the expression for u simplifies to

$$u_\lambda(\mathbf{p}) = \sqrt{\omega(\mathbf{p})+m}\begin{pmatrix}\chi_\lambda\\ \frac{2\lambda|\mathbf{p}|}{m+\omega(\mathbf{p})}\chi_\lambda\end{pmatrix},$$

and furthermore u_λ is itself an eigenstate of helicity $\mathbf{p}\cdot\boldsymbol{\Sigma}/2|\mathbf{p}|$ with eigenvalue λ.

In all cases we maintain our choice (6.7)

$$u_-(\mathbf{p}) = i\gamma^2 u^*(-\mathbf{p})$$

for the negative energy basis functions. For the helicity basis choice for u_λ, this construction gives a negative energy spinor with the same helicity, as can easily be shown by applying h to both sides. For the rest frame Σ_3 basis this construction reverses the sign of Σ_3 in the rest frame.

The properties of $\psi^{(\pm)}_{\lambda,\mathbf{p}}(\mathbf{x},a)$ needed for the second quantization formalism are

$$\text{Orthonormality:} \qquad \int d^3x \sum_a \psi^{(\pm)*}_{\lambda,\mathbf{p}}(\mathbf{x},a)\psi^{(\pm)'}_{\lambda',\mathbf{p}'}(\mathbf{x},a) = \delta_{\lambda'\lambda}\delta_{(\pm)'\pm}\delta(\mathbf{p}'-\mathbf{p}) \tag{8.1}$$

and

$$\text{Completeness:} \qquad \int d^3p \sum_{\lambda\pm} \psi^{(\pm)}_{\lambda,\mathbf{p}}(\mathbf{x},a)\psi^{(\pm)*}_{\lambda,\mathbf{p}}(\mathbf{x}',a') = \delta_{aa'}\delta(\mathbf{x}'-\mathbf{x}). \tag{8.2}$$

To pass to the second quantized formalism, we simply introduce creation and annihilation operators $b^\dagger_{\lambda\pm}(\mathbf{p}), b_{\lambda\pm}(\mathbf{p})$ with anticommutation relations

$$\{b_{\lambda\pm}(\mathbf{p}), b^\dagger_{\lambda'\pm'}(\mathbf{p}')\} = \delta_{\lambda'\lambda}\delta_{(\pm)'\pm}\delta(\mathbf{p}'-\mathbf{p}),$$

and define the Dirac quantum field operator

$$\psi^a(\mathbf{x}) = \int d^3p \sum_{\lambda\pm} b_{\lambda\pm}(\mathbf{p})\psi^{(\pm)}_{\lambda,\mathbf{p}}(\mathbf{x},a).$$

By virtue of (8.2) the field operators satisfy

$$\{\psi^a(\mathbf{x}), \psi^{a'\dagger}(\mathbf{x}')\} = \delta_{aa'}\delta(\mathbf{x}-\mathbf{x}').$$

A principal virtue of second quantization is the efficiency with which we can construct the state describing the negative energy sea. We simply apply to the empty state all of the creation operators for negative energy states:

$$|sea\rangle \equiv |0\rangle = \mathcal{N}\prod_{\mathbf{p},\lambda} b^\dagger_{\lambda-}(\mathbf{p})\,|0'\rangle\,. \tag{8.3}$$

This looks terribly complicated, but we can uniquely characterize this state very simply: It is annihilated by all of the positive energy annihilation operators and by all the negative energy creation operators

$$b_{\lambda+}(\mathbf{p})\,|0\rangle = b^\dagger_{\lambda-}(\mathbf{p})\,|0\rangle = 0.$$

These conditions tell us everything we need to know about the sea. The annihilation operator for a negative energy electron creates a hole in the sea. Thus the construction of the sea is completely equivalent to interchanging the role of the creation and annihilation operators for the negative energy Dirac particles.

To see the consequences of this interchange of roles, let us consider a few of the observables of the theory. The Hamiltonian is just

$$\begin{aligned} H &= \int d^3x \sum_a \psi^{a\dagger}(\mathbf{x})(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi^a(\mathbf{x}) \\ &= \int d^3p\omega(\mathbf{p}) \sum_\lambda (b^\dagger_{\lambda+}(\mathbf{p})b_{\lambda+}(\mathbf{p}) - b^\dagger_{\lambda-}(\mathbf{p})b_{\lambda-}(\mathbf{p})) \\ &= \int d^3p\omega(\mathbf{p}) \sum_\lambda (b^\dagger_{\lambda+}(\mathbf{p})b_{\lambda+}(\mathbf{p}) + b_{\lambda-}(\mathbf{p})b^\dagger_{\lambda-}(\mathbf{p})) - 2\int d^3p\omega(\mathbf{p})\delta(\mathbf{0}) \end{aligned}$$

where in the last form we have reordered the creation and annihilation operators of the negative energy contributions, the nonzero anticommutator producing the negative infinite constant term. Notice that thanks

to the Fermi statistics both contributions to the energy are positive. This constant is just the energy of the sea. the factor of $\delta(\mathbf{0})$ can be identified with $Volume/(2\pi)^3$ so the sea has an infinite negative energy density. As we have already stressed we can and will choose to measure all energies relative to that of the sea which amounts to dropping this constant*, so henceforth we shall take the free Dirac second quantized Hamiltonian to be

$$H_{Dirac} = \int d^3p\, \omega(\mathbf{p}) \sum_\lambda (b^\dagger_{\lambda+}(\mathbf{p}) b_{\lambda+}(\mathbf{p}) + b_{\lambda-}(\mathbf{p}) b^\dagger_{\lambda-}(\mathbf{p})).$$

When we pass to the Heisenberg picture we find the field equation for ψ to be nothing other than the Dirac equation

$$(\frac{1}{i}\gamma\cdot\partial + m)\psi = 0.$$

Since we have selected our single particle basis to be eigenstates of $\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m$ the time dependence of ψ in Heisenberg picture is simply

$$\psi^a(\mathbf{x},t) = \int d^3p \sum_{\lambda\pm} b_{\lambda\pm}(\mathbf{p}) \psi^{(\pm)}_{\lambda,\mathbf{p}}(\mathbf{x},a) e^{\mp i\omega(\mathbf{p})t} \tag{8.4}$$

The annihilation operators for the positive and negative energy Dirac particles are thus identified with the positive and negative frequency components of the Dirac field in Heisenberg picture. This is a useful observation because when we introduce time dependent external fields which are switched off at early and late times, it will allow us to easily relate the operators that characterize the sea at late times to the ones that characterize the sea at early times.

Returning to our survey of observables, the momentum operator is just

$$\begin{aligned}
\mathbf{P} &= \int d^3x \sum_a \psi^{a\dagger}(\mathbf{x}) \frac{1}{i}\nabla\psi^a(\mathbf{x}) \\
&= \int d^3p\mathbf{p} \sum_\lambda (b^\dagger_{\lambda+}(\mathbf{p}) b_{\lambda+}(\mathbf{p}) + b^\dagger_{\lambda-}(\mathbf{p}) b_{\lambda-}(\mathbf{p})) \\
&= \int d^3p\mathbf{p} \sum_\lambda (b^\dagger_{\lambda+}(\mathbf{p}) b_{\lambda+}(\mathbf{p}) + b_{\lambda-}(-\mathbf{p}) b^\dagger_{\lambda-}(-\mathbf{p}))
\end{aligned}$$

The term $2\int d^3p\mathbf{p}\delta(\mathbf{0})$ arising from reordering the negative energy operators automatically vanishes and need not be dropped. We see from the explicit form of the momentum operator that $b_{\lambda-}(-\mathbf{p})$ creates from the sea a particle of momentum $+\mathbf{p}$.

The charge operator is just $Q = qN$ where q is the unit of charge carried by the Dirac particle and N is the number operator

$$\begin{aligned}
Q =& q\int d^3x \sum_a \psi^{a\dagger}(\mathbf{x})\psi^a(\mathbf{x}) \\
=& q\int d^3p \sum_\lambda (b^\dagger_{\lambda+}(\mathbf{p}) b_{\lambda+}(\mathbf{p}) - b_{\lambda-}(-\mathbf{p}) b^\dagger_{\lambda-}(-\mathbf{p})) + 2q\int d^3p\delta(\mathbf{0})
\end{aligned}$$

from which we see that $b_{\lambda-}(-\mathbf{p})$ creates a state of charge $-q$. We shall also in future drop the constant term

$\star$ Alternatively we could introduce a bare cosmological constant to cancel it.

in Q so the charge of the sea is then zero† There is a convenient way to make this subtraction. Instead of taking the charge density to be $\rho(\mathbf{x}) = q\psi^\dagger\psi$, take it to be the symmetrized form

$$\rho(\mathbf{x}) = \frac{q}{2}\sum_a [\psi_a^\dagger(\mathbf{x})\psi_a(\mathbf{x}) - \psi_a(\mathbf{x})\psi_a^\dagger(\mathbf{x})] \tag{8.5}$$

Then when the operators in $Q = \int d^3x\rho$ are suitably reordered the piece coming from the positive energy term exactly cancels that from the negative energy term. At this point we can also identify the current operator from local current conservation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0. \tag{8.6}$$

Inserting (8.5) into (8.6) and using the Dirac equation, we identify

$$\mathbf{j} = \frac{q}{2}\sum_a [\psi_a^\dagger(\mathbf{x})(\boldsymbol{\alpha}\psi)_a(\mathbf{x}) - (\boldsymbol{\alpha}\psi)_a(\mathbf{x})\psi_a^\dagger(\mathbf{x})].$$

We can assemble $(\rho,\mathbf{j})$ in a four vector j^μ:

$$\begin{aligned} j^\mu(\mathbf{x},t) =& \frac{q}{2}\sum_a [\psi_a^\dagger(\mathbf{x},t)(\beta\gamma^\mu\psi)_a(\mathbf{x},t) - (\beta\gamma^\mu\psi)_a(\mathbf{x},t)\psi_a^\dagger(\mathbf{x},t)] \\ =& \frac{q}{2}\sum_a [\psi_a^\dagger(\mathbf{x},t), (\beta\gamma^\mu\psi)_a(\mathbf{x},t)] \\ =& \frac{q}{2}\sum_a [\bar\psi_a(\mathbf{x},t), (\gamma^\mu\psi)_a(\mathbf{x},t)] \end{aligned}$$

We have made use of the Dirac adjoint

$$\bar\psi_a \equiv \sum_b \psi_b^\dagger\beta_{ba}$$

in the last form, which we will also sometimes shorten even more by suppressing the spinor indices, $\frac{q}{2}[\bar\psi, \gamma^\mu\psi]$. Current conservation $\partial_\mu j^\mu = 0$ is an immediate consequence of the Dirac equation and its Dirac adjoint

$$\bar\psi(-i\gamma\cdot\overleftarrow{\partial} - m) = 0.$$

The final observable we mention is the angular momentum

$$\mathbf{J} = \int d^3x \sum_{a,b} \psi^{a\dagger}(\mathbf{x})\left(\frac{1}{i}(\mathbf{x}\times\nabla) + \frac{1}{2}\boldsymbol{\Sigma}\right)_{ab}\psi^b(\mathbf{x}).$$

The sea is of course rotationally invariant

$$\mathbf{J}\,|0\rangle = 0,$$

as will be shown in an exercise. Of particular interest is the action of the helicity on the single particle states.

† If there is more than one species of fermion in the universe, the coefficient of this term is $\sum_f q_f$. One way of getting rid of the sea charge is to insist that this sum of charges vanishes. As it happens the standard model of strong weak and electromagnetic interactions has this property, which is required for cancellation of the axial anomaly. Thanks to Mr. Yan-Bo Xie for drawing my attention to this circumstance. In a similar vein supersymmetry is often proposed so that the zero point energy of bosons exactly cancels the sea energy of fermions.

On the particle states we find

$$\mathbf{p}\cdot\mathbf{J}b^\dagger_{\lambda+}(\mathbf{p})\,|0\rangle = \frac{1}{2\omega}\sum_a u^{a*}_{\lambda'+}(\mathbf{p})\frac{1}{2}\mathbf{p}\cdot\mathbf{\Sigma}_{ab}u^b_{\lambda+}(\mathbf{p})b^\dagger_{\lambda'+}(\mathbf{p})\,|0\rangle = \lambda|\mathbf{p}|b^\dagger_{\lambda+}(\mathbf{p})\,|0\rangle\,,$$

confirming that this state carries helicity λ. The helicity of the one hole state $b_{\lambda-}(-\mathbf{p})\,|0\rangle$ which possesses momentum $+\mathbf{p}$ is also λ but the reason is slightly subtle. First of all

$$\mathbf{p}\cdot\mathbf{J}b_{\lambda-}(-\mathbf{p})\,|0\rangle = -\frac{1}{2\omega}\sum_a u^{a*}_{\lambda-}(-\mathbf{p})\frac{1}{2}\mathbf{p}\cdot\mathbf{\Sigma}_{ab}u^b_{\lambda'-}(-\mathbf{p})b_{\lambda'-}(-\mathbf{p})\,|0\rangle$$

where the minus sign arises because $b^\dagger_{\lambda-}$ occurs in $\psi^\dagger$ and must anticommute with ψ before it can contract against $b_{\lambda-}$. But then

$$\frac{1}{2}(-\mathbf{p})\cdot\mathbf{\Sigma}u_{\lambda-}(-\mathbf{p}) = \lambda|\mathbf{p}|u_{\lambda-}(-\mathbf{p}).$$

Thus

$$\frac{\mathbf{p}\cdot\mathbf{J}}{|\mathbf{p}|}b_{\lambda-}(-\mathbf{p})\,|0\rangle = \lambda b_{\lambda-}(-\mathbf{p})\,|0\rangle$$

as we claimed.

This survey of single particle observables has established:

6. The state $b^\dagger_{\lambda+}(\mathbf{p})\,|0\rangle$ is a one particle state of momentum $\mathbf{p}$, energy $\omega = \sqrt{\mathbf{p}^2+m^2}$, charge q, and helicity λ.
7. The state $b_{\lambda-}(-\mathbf{p})\,|0\rangle$ is a one particle state of momentum $\mathbf{p}$, energy $\omega = \sqrt{\mathbf{p}^2+m^2}$, charge $-q$, and helicity λ.

In particular if the first state is an electron of charge $-e$, then the second is a positron of charge $+e$. To emphasize these facts it is traditional to rename the creation and annihilation operators for negative energy particles. So define

$$b_{\lambda-}(-\mathbf{p}) \equiv d^\dagger_\lambda(\mathbf{p}) \qquad\qquad b^\dagger_{\lambda+}(\mathbf{p}) \equiv b^\dagger_\lambda(\mathbf{p}).$$

So $b^\dagger_\lambda(\mathbf{p})$ creates a particle and $d^\dagger_\lambda(\mathbf{p})$ creates an antiparticle. Similarly it is useful to define the Dirac spinor

$$v_\lambda(\mathbf{p}) \equiv u_{\lambda-}(-\mathbf{p}) = i\gamma^2 u^*_\lambda(\mathbf{p}).$$

Note that u and v satisfy

$$\begin{aligned}(\gamma\cdot p+m)u_\lambda(\mathbf{p}) &= 0 \qquad & (h-\lambda)u_\lambda(\mathbf{p}) &= 0\\ (\gamma\cdot p-m)v_\lambda(\mathbf{p}) &= 0 \qquad & (h+\lambda)v_\lambda(\mathbf{p}) &= 0\end{aligned}$$

with opposite signs in front of the mass and opposite helicities.

With these definitions the free Dirac field operator in Heisenberg picture has the representation

$$\psi(x) = \int\frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega}}\sum_\lambda\left(b_\lambda(\mathbf{p})u_\lambda(\mathbf{p})e^{ix\cdot p} + d^\dagger_\lambda(\mathbf{p})v_\lambda(\mathbf{p})e^{-ip\cdot x}\right),$$

where $p\cdot x = \mathbf{p}\cdot\mathbf{x}-\omega(\mathbf{p})t$ is the Minkowski scalar product. A point to bear in mind with this new interpretation is that one body operators will generally contain terms like $b^\dagger d^\dagger$ which create a particle antiparticle pair and terms like bd which destroy such a pair. Thus when we couple couple currents to the electromagnetic field we will have charge conservation, but *not* particle number conservation.

9. The Discrete Symmetries of the Dirac Equation

9.1. Parity

The parity transformation $\mathbf{x} \to -\mathbf{x}$ can be extended to a symmetry of the Hamiltonian. Consider the following transformation on the field:

$$\psi(\mathbf{x},t) \to P^{-1}\psi(\mathbf{x},t)P = e^{i\phi}\beta\psi(-\mathbf{x},t), \tag{9.1}$$

where we have allowed a multiplicative phase. Then the Hamiltonian transforms to

$$\begin{aligned} P^{-1}HP &= \int d^3x\psi^\dagger(-\mathbf{x})\beta(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\beta\psi(-\mathbf{x}) \\ &= \int d^3x\psi^\dagger(-\mathbf{x})(-\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi(-\mathbf{x}) = H \end{aligned} \tag{9.2}$$

after changing integration variables, so it is parity invariant. From the parity transformation (9.1) we can infer how Parity acts on the particle states:

$$e^{i\phi}\beta\psi(-\mathbf{x},t) = e^{i\phi}\int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega}}\sum_\lambda \left(b_\lambda(-\mathbf{p})\beta u_\lambda(-\mathbf{p})e^{ix\cdot p} + d^\dagger_\lambda(-\mathbf{p})\beta v_\lambda(-\mathbf{p})e^{-ip\cdot x}\right),$$

where we have reversed the sign of $\mathbf{p}$ by a change of variables. Next we note that depending on the spin basis we choose,

$$\text{Helicity Basis:} \qquad \begin{aligned} \beta u_\lambda(\mathbf{p}) &= -\,ie^{i\lambda(\pi+2\phi_\mathbf{p})}u_{-\lambda}(-\mathbf{p}) \\ \beta v_\lambda(\mathbf{p}) &= -\,ie^{-i\lambda(\pi+2\phi_\mathbf{p})}v_{-\lambda}(-\mathbf{p}) \end{aligned}$$

where we have used the formula

$$\chi_\lambda(\mathbf{p}) = -ie^{i\lambda(\pi+2\phi_\mathbf{p})}\chi_{-\lambda}(-\mathbf{p})$$

obtained in the exercises, or

$$\text{Rest Frame } \Sigma_3: \qquad \begin{aligned} \beta u_\mu(\mathbf{p}) &= u_\mu(-\mathbf{p}) \\ \beta v_\mu(\mathbf{p}) &= -\,v_\mu(-\mathbf{p}). \end{aligned}$$

Using these spinor properties, we learn that

$$\text{Helicity Basis:} \qquad \begin{aligned} P^{-1}b_\lambda(\mathbf{p})P &= e^{i\phi}ie^{-i\lambda(\pi+2\phi_\mathbf{p})}b_{-\lambda}(-\mathbf{p}) \\ P^{-1}d^\dagger_\lambda(\mathbf{p})P &= e^{i\phi}ie^{i\lambda(\pi+2\phi_\mathbf{p})}d^\dagger_{-\lambda}(-\mathbf{p}) \\ P^{-1}b^\dagger_\lambda(\mathbf{p})P &= e^{-i\phi}(-i)e^{i\lambda(\pi+2\phi_\mathbf{p})}b^\dagger_{-\lambda}(-\mathbf{p}) \end{aligned}$$

where the last equation is just the hermitian conjugate of the first. And

$$\text{Rest Frame } \Sigma_3: \qquad \begin{aligned} P^{-1}d^\dagger_\mu(\mathbf{p})P &= -\,e^{i\phi}d^\dagger_\mu(-\mathbf{p}) \\ P^{-1}b^\dagger_\mu(\mathbf{p})P &= e^{-i\phi}b^\dagger_\mu(-\mathbf{p}). \end{aligned}$$

It should be noted that the arbitrary phase we allowed in the definition of parity cancels out for neutral states, *i.e.* those with an equal number of $b^\dagger$'s and $d^\dagger$'s acting on the sea, which we can take to be parity invariant. This means that whereas the intrinsic parity of a single particle is conventional, that of a particle antiparticle pair is not. For example the above formulae imply that the parity of the ground state s wave of positronium is odd, *i.e.* the ground states are 0^- and 1^-.

It is apparent that a Dirac particle and its antiparticle are closely related to each other. They have identical mass and spin, but exactly opposite charges. In fact this relationship is a reflection of a symmetry in the dynamics under interchange of a particle with its antiparticle. To explore this symmetry, first define a unitary transformation C, $CC^\dagger = I$ by the rules

$$C^{-1}b^\dagger C = d^\dagger \qquad C^{-1}d^\dagger C = b^\dagger \qquad C\,|0\rangle = |0\rangle\,. \tag{9.3}$$

(The last of these equations implies an extremely complicated transformation of the empty state, but we shall never see those complications.) From the definition of C we can work out how the field transforms

$$C^{-1}\psi(x)C = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega}} \sum_\lambda \left(d_\lambda(\mathbf{p})u_\lambda(\mathbf{p})e^{ix\cdot p} + b^\dagger_\lambda(\mathbf{p})v_\lambda(\mathbf{p})e^{-ip\cdot x}\right)$$

which we can relate to $\psi^\dagger$:

$$\psi^\dagger_a(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega}} \sum_\lambda \left(b^\dagger_\lambda(\mathbf{p})u^*_{\lambda a}(\mathbf{p})e^{-ix\cdot p} + d_\lambda(\mathbf{p})v^*_{\lambda a}(\mathbf{p})e^{ip\cdot x}\right).$$

But $v^* = i\gamma^2 u$ and $u^* = i\gamma^2 v$ so we can infer

$$C^{-1}\psi_a(x)C = i(\gamma^2)_{ab}\psi^\dagger_b(x). \tag{9.4}$$

When we suppress spinor indices, it is usually convenient to think of $\psi^\dagger$ as a row vector and ψ as a column vector. To write (9.4) with suppressed indices we want the r.h.s. to be a column vector which we could indicate by $(\psi^\dagger)^T$ meaning the transpose on spinor indices. In that case (9.4) can be written

$$C^{-1}\psi(x)C = i\gamma^2(\psi^\dagger)^T(x). \tag{9.5}$$

It should also be noted that the occurrence of γ^2 in the charge conjugation transformation law is specific to the standard representation; with other representations a different matrix would appear.

The invariance of H is obvious from its expression in terms of creation and annihilation operators, but it is also instructive to see it using the local definition

$$\begin{aligned} C^{-1}HC =& \int d^3x\psi^T i\gamma^2(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)i\gamma^2(\psi^\dagger)^T \\ =& -\int d^3x\psi^T(-\frac{1}{i}\boldsymbol{\alpha}^T\cdot\nabla + \beta^T m)(\psi^\dagger)^T \\ =& H \end{aligned}$$

where the last step involves an integration by parts, a transposition of Dirac indices, and a reordering of the order of ψ and $\psi^\dagger$ giving a minus sign which cancels the overall minus sign in the second line.

The transformation of the charge and current densities under C should simply change their signs. This is not hard to see:

$$\begin{aligned} C^{-1}j^\mu C =& \frac{q}{2}(\psi^T i\gamma^2\beta\gamma^\mu i\gamma^2(\psi^\dagger)^T - \psi^\dagger i\gamma^2\gamma^{\mu T}\beta i\gamma^2\psi \\ =& -j^\mu \end{aligned}$$

where use is made of

$$i\gamma^2\gamma^\mu i\gamma^2 = -\gamma^{\mu *} \qquad \beta\gamma^\mu\beta = \gamma^{\mu\dagger}.$$

Note that because we have used the symmetrized definition of the current, there is no reordering of operators necessary in arriving at this result.

9.3. Majorana Fermions

In our discussion so far it has seemed inevitable that the Dirac particle carries charge. More precisely, it carries a conserved fermion number N which could be identified with charge. If there were several species f of Dirac particle the fermion number N_f of each species might be separately conserved. Including terms of the form

$$\bar{\psi}_f \Gamma \psi_{f'}$$

with $f' \neq f$ in the Hamiltonian could violate the individual N_f but $\sum_f N_f$ would still be conserved. Majorana pointed out that even with only one species of fermion it is possible to make it totally neutral, *i.e.* carry no conserved quantum number at all.

Starting with the Dirac theory we can see that this is possible, because we can consider redefining creation and annihilation operators to be eigenoperators under charge conjugation:

$$b_{\lambda\pm}(\mathbf{p}) = \frac{1}{\sqrt{2}}(b_\lambda(\mathbf{p}) \pm d_\lambda(\mathbf{p})) \qquad \text{with } C^{-1} b_{\lambda\pm}(\mathbf{p}) C = \mp b_{\lambda\pm}(\mathbf{p}).$$

Then the Hamiltonian is the sum of two commuting pieces

$$H = H_+ + H_- \qquad \text{with } H_\pm = \int d^3p\, \omega(\mathbf{p}) \sum_\lambda b^\dagger_{\lambda\pm}(\mathbf{p}) b_{\lambda\pm}(\mathbf{p}).$$

Clearly, it is perfectly consistent to consider the quantum system defined by H_+ (or H_-) alone. The number operator of the Dirac theory

$$N = \int d^3p \sum_\lambda (b^\dagger_{\lambda-}(\mathbf{p}) b_{\lambda+}(\mathbf{p}) + b^\dagger_{\lambda+}(\mathbf{p}) b_{\lambda-}(\mathbf{p})),$$

clearly has no meaning in the truncated theory, but that is to be expected.

One might worry that truncating the theory in this way might spoil locality, but this is not the case. We can just as easily redefine the local fields to be eigenoperators of charge conjugation:

$$\psi_\pm(\mathbf{x}) = \frac{1}{\sqrt{2}}(\psi(\mathbf{x}) \pm i\gamma^2 (\psi^\dagger)^T(\mathbf{x})) \qquad \text{with } C^{-1}\psi_\pm C = \pm\psi_\pm, \tag{9.6}$$

Which satisfy anticommutation relations

$$\{\psi^a_\pm(\mathbf{x}), \psi^b_\pm(\mathbf{x}')\} = \pm(i\gamma^2)_{ab}\delta(\mathbf{x}-\mathbf{x}').$$

Clearly

$$\psi_\pm(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega}} \sum_\lambda \left(b_{\lambda\pm}(\mathbf{p}) u_\lambda(\mathbf{p}) e^{ix\cdot p} \pm b^\dagger_{\lambda\pm}(\mathbf{p}) v_\lambda(\mathbf{p}) e^{-ip\cdot x} \right),$$

and in terms of these fields

$$\begin{aligned} H_\pm =& \frac{1}{2}\int d^3x \psi^\dagger_\pm (\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi_\pm \\ =& \frac{1}{2}\int d^3x (\pm\psi^T_\pm) i\gamma^2 (\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi_\pm. \end{aligned}$$

In the second line we have used $\psi^\dagger_\pm = \pm\psi^T_\pm i\gamma^2$, a consequence of (9.6).

The appearance of $i\gamma^2$ in the above discussion is due to our choice of the standard representation for the gamma matrices. The Majorana representation is characterized by the condition that the gamma matrices be pure imaginary $\gamma^{\mu *} = -\gamma^\mu$. In that case the charge conjugation transformation does not involve a matrix at all and all of the $i\gamma^2$'s disappear.

In the case of massless fermions $m = 0$, it is possible to describe relativistic spin 1/2 particles with only one helicity. In the Dirac theory the easiest way to see this is to consider the matrix

$$\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3,$$

which anticommutes with the γ^μ. It therefore commutes with the Lorentz matrices

$$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu].$$

γ_5 commutes with the $\boldsymbol{\alpha} \cdot \nabla$ term of the Dirac Hamiltonian but not with the $m\beta$ term. So when $m = 0$, and only then, the energy eigenstates can be simultaneously eigenstates of γ_5. Since $\gamma_5^2 = 1$ the eigenvalues of γ_5, called chirality, are ± 1 and

$$\frac{I \pm \gamma_5}{2}$$

are projectors onto orthogonal two dimensional subspaces with chirality ± 1 respectively. Defining

$$R = \frac{I + \gamma_5}{2}\psi \qquad\qquad L = \frac{I - \gamma_5}{2}\psi,$$

R for "right-handed" and L for "left-handed," the Dirac hamiltonian for $m = 0$ decomposes into two commuting terms

$$H = \int d^3x (R^\dagger \frac{1}{i}\boldsymbol{\alpha} \cdot \nabla R + L^\dagger \frac{1}{i}\boldsymbol{\alpha} \cdot \nabla L)$$

either of which could define a consistent dynamics of the corresponding subsystem. These subsystems are called Weyl fermions. The corresponding momentum space spinors of definite chirality are

$$\frac{I \pm \gamma_5}{2} u_\lambda(\mathbf{p}) = \frac{1}{2}\sqrt{|\mathbf{p}|}(1 \pm 2\lambda)\begin{pmatrix} \chi_\lambda \\ \pm\chi_\lambda \end{pmatrix},$$

from which it is clear that helicity is identical to $Chirality/2$. In other words right-handed Weyl fermions have helicity $+1/2$ and left-handed ones have helicity $-1/2$. Since γ_5 is real and γ^2 anticommutes with γ_5, the antiparticle spinors

$$\frac{I \pm \gamma_5}{2} v_\lambda(\mathbf{p}) = i\gamma^2 (\frac{I \pm \gamma_5}{2} u_\lambda(\mathbf{p}))^*$$

have the opposite correlation between chirality and helicity. So if the Weyl particle is right handed the particle has helicity $+1/2$ and the antiparticle has helicity $-1/2$. Since charge conjugation interchanges the role of particle and antiparticle one can even choose by convention all Weyl particles to be left(right)-handed. One can do this, for example, by writing $R^\dagger = L'^T i\gamma^2$. Of course, such an L' has charge opposite to R.

The Weyl particle with helicity $\pm 1/2$ comes along with its antiparticle with helicity $\mp 1/2$. Thus the Weyl system has the same helicity content as the massless Majorana system. In fact one can describe the Majorana system using Weyl fields. First, separate the Majorana field $\psi^\dagger = \psi^T i\gamma^2$ into two fields of definite

chirality

$$\psi_{R,L} = \frac{I \pm \gamma_5}{2}\psi.$$

Then notice that

$$\begin{aligned}\psi_R^\dagger =& \psi^\dagger \frac{I+\gamma_5}{2} \\ =& \psi^T i\gamma^2 \frac{I+\gamma_5}{2} \\ =& \psi_L^T i\gamma^2\end{aligned}$$

so that the right-handed component of the Majorana field can be eliminated in favor of the hermitian conjugate of the left-handed component. When this is done the Majorana hamiltonian simplifies to

$$H_{Maj} = \int d^3x \psi_L^\dagger \frac{1}{i}\boldsymbol{\alpha}\cdot\nabla\psi_L + \frac{m}{2}\int d^3x [\underbrace{\psi_L^T i\gamma^2\beta\psi_L}_{\Delta F=-2} + \underbrace{(\psi_L^T i\gamma^2\beta\psi_L)^\dagger}_{\Delta F=+2}]$$

which reduces to the Weyl hamiltonian for $m = 0$. Notice that the massless limit conserves fermion number but the mass term violates fermion number conservation by ± 2 units. This is the so-called Majorana mass term for a Weyl fermion. To construct the Dirac mass term which conserves fermion number, one must add a right-handed Weyl fermion to the theory.

The Weyl theory violates parity invariance, essentially because β fails to commute with γ_5:

$$P^{-1}L(\mathbf{x})P = \frac{I-\gamma_5}{2}\beta\psi(-\mathbf{x}) = \beta R(-\mathbf{x}).$$

How then have we managed to show that it is equivalent to the parity conserving Majorana theory? The answer is that although parity is violated in the Weyl theory, CP the product of parity times charge conjugation remains a symmetry. The Majorana field is inert under charge conjugation, and has only parity as a nontrivial symmetry. It is the CP symmetry of the Weyl theory that corresponds to the parity of the Majorana theory.

Notice that it is impossible to have a fermion that is simultaneously Majorana and Weyl: even at zero mass one always must have both helicities. This is true in four space-time dimensions but in other dimensions it need not be so. For example in 10 dimensions one can define Majorana-Weyl fermions. A Dirac fermion in $D = 2k$ dimensions has 2^k degrees of freedom, 2^{k-1} states for the particle and the same number for the antiparticle. In some dimensions (including 4 and 10) one can have Majorana fermions with only 2^{k-1} degrees of freedom. In the massless case one can define Weyl fermions in all even dimensions giving 2^{k-1} degrees of freedom. In $2 + 8n$ dimensions one can have Majorana-Weyl fermions with only 2^{k-2} degrees of freedom. For example in 10 dimensions a Dirac fermion has 32 states, a Majorana or Weyl fermion has 16 states, and a Majorana-Weyl fermion has only 8 states. This possibility is crucial for the consistency of superstring theory.

9.5. Time Reversal

The last discrete symmetry we discuss is time reversal T. It is well-known that T must be an antiunitary transformation, meaning that it is antilinear, and furthermore

$$\langle T\Phi|T\Psi\rangle = \langle\Psi|\Phi\rangle\,.$$

With this in mind, we search for a transformation of the form

$$T^{-1}\psi(\mathbf{x},t)T = \mathcal{T}\psi(\mathbf{x},-t),$$

with $\mathcal{T}$ an appropriate matrix. From antiunitarity we have

$$\left\langle\Phi|T^{-1}\psi T\Psi\right\rangle = \langle\psi T\Psi|T\Phi\rangle = \left\langle T\Psi|\psi^\dagger T\Phi\right\rangle = \left\langle T^{-1}\psi^\dagger T\Phi|\Psi\right\rangle,$$

from which it follows that

$$T^{-1}\psi^\dagger(t)T = (T^{-1}\psi T)^\dagger = \psi^\dagger(-t)\mathcal{T}^\dagger.$$

Thus

$$T^{-1}HT = \int d^3x\psi^\dagger(\mathbf{x},-t)\mathcal{T}^\dagger\left(-\frac{1}{i}\boldsymbol{\alpha}^*\cdot\nabla+\beta m\right)\mathcal{T}\psi(\mathbf{x},-t).$$

If we choose $\mathcal{T}$ to be unitary, then invariance of H will be achieved (given conservation of energy $dH/dt=0$, which follows from the field equation) if and only if $\mathcal{T}$ commutes with β and α^2 and anticommutes with α^1 and α^3. Clearly the most general solution of these conditions is

$$\mathcal{T} = e^{i\tau}\gamma^1\gamma^3 = ie^{i\tau}\Sigma_2,$$

so that the transformation law becomes

$$T^{-1}\psi(\mathbf{x},t)T = ie^{i\tau}\Sigma_2\psi(\mathbf{x},-t). \tag{9.7}$$

This transformation law on ψ implies that for b,d. It is easiest to do this for the rest frame Σ_3 basis, because then the two-spinors ϕ_μ are real. Thus

$$\begin{aligned} u^*_\mu(-\mathbf{p}) =& \sqrt{\omega(\mathbf{p})+m}\begin{pmatrix}\phi_\mu\\ \frac{-\boldsymbol{\sigma}^*\cdot\mathbf{p}}{m+\omega(\mathbf{p})}\phi_\mu\end{pmatrix}\\ =&\Sigma_2\sqrt{\omega(\mathbf{p})+m}\begin{pmatrix}\sigma_2\phi_\mu\\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{m+\omega(\mathbf{p})}\sigma_2\phi_\mu\end{pmatrix}\\ =&i^{2\mu}\Sigma_2u_{-\mu}(\mathbf{p})\end{aligned} \tag{9.8}$$

Using $v=i\gamma^2u^*$, it is just a few steps to show that

$$v^*_\mu(-\mathbf{p}) = i^{2\mu}\Sigma_2v_{-\mu}(\mathbf{p}). \tag{9.9}$$

Because T is antilinear, the l.h.s. of (9.7) involves $u^*e^{-ix\cdot p}T^{-1}bT$ and $v^*e^{+ix\cdot p}T^{-1}d^\dagger T$, so (9.8) and (9.9) allow us to infer from (9.7) that

$$\begin{aligned} T^{-1}b_\mu(\mathbf{p})T =& ie^{i\tau}i^{-2\mu}b_{-\mu}(-\mathbf{p}) & T^{-1}b^\dagger_\mu(\mathbf{p})T =& -ie^{-i\tau}i^{+2\mu}b^\dagger_{-\mu}(-\mathbf{p})\\ T^{-1}d^\dagger_\mu(\mathbf{p})T =& ie^{i\tau}i^{-2\mu}d^\dagger_{-\mu}(-\mathbf{p}) & T^{-1}d_\mu(\mathbf{p})T =& -ie^{-i\tau}i^{+2\mu}d_{-\mu}(-\mathbf{p})\end{aligned}$$

The reversal of signs of momentum and spin label is intuitively correct since time reversing a motion reverses both momentum and angular momentum. The μ dependence of the phase is perhaps less intuitive, but

follows straightforwardly by using angular momentum raising and lowering operators together with the action of time reversal on angular momentum. If we had used the helicity basis, the helicity label would not be reversed by time reversal (remember it is $\mathbf{J}\cdot\mathbf{p}/|\mathbf{p}|$); unfortunately the phases that are induced are angle dependent and not very illuminating.

9.6. Violation of the Discrete Symmetries and the CPT Theorem

Having shown that C, P, and T are symmetries of the Dirac equation it is instructive to contemplate how things must be changed to violate these symmetries. For example, we have seen that both parity and charge conjugation are violated with Weyl fermions, but in such a way that CP remains a symmetry. More generally, one can consider adding noninvariant terms to the Hamiltonian. In exercises, it is shown how the bilinears $\bar\psi_A\Gamma\psi_B$ transform under these symmetries for $\Gamma = (I, i\gamma_5, \gamma^\mu, \gamma_5\gamma^\mu, \sigma^{\mu\nu})$. Under parity they transform with a factor of $(+,-,+,-,+)(-)^S$ times the bilinear evaluated with $\mathbf{x}\to-\mathbf{x}$, and where S is the number of *spatial* indices in the tensor component. So, examples of parity odd Lorentz invariants would be

$$\bar\psi i\gamma_5\psi \qquad\qquad \bar\psi\gamma^\mu\psi\bar\psi\gamma_5\gamma_\mu\psi. \tag{9.10}$$

Adding such terms to the energy density would appear to violate parity. One must be careful that the violation is not an illusion. For example the term $\psi^\dagger\boldsymbol{\alpha}\cdot\nabla\psi$ is invariant under the chiral symmetry

$$\psi\to e^{i\alpha\gamma_5}\psi$$

under which

$$\bar\psi\psi\to\cos 2\alpha\bar\psi\psi+\bar\psi i\gamma_5\psi\sin 2\alpha$$

so the added term $\bar\psi i\gamma_5\psi$ can be rotated into $\bar\psi\psi$ and parity violation disappears.

With regard to charge conjugation, $\bar\psi_A\Gamma\psi_B$ transforms to $(+,+,-,+,-)$ times $\bar\psi_B\Gamma\psi_A$. For example the first of (9.10) is invariant under charge conjugation (if $A=B$) but the second is odd. Under CP the result is $(+,-,-,-,-)(-)^S$. Thus it is the first of (9.10) that would violate CP. Since such a term can be rotated away by a chiral transformation we see that CP is a bit tricky to violate. For example in the Standard Model one needs at least three generations of quarks and leptons to frustrate the ability to transform away apparently CP violating couplings! Fortunately there is solid evidence for this number of generations.[⋆]

Finally we come to time reversal, under which the bilinears can be shown to transform into $(+,-,+,+,-)(-)^S$ times the bilinear with $t\to-t$. Note that $\bar\psi\gamma^\mu\psi$ transforms as expected for a current and $\bar\psi\sigma^{kl}\psi$ as expected for an angular momentum. It is only the first of (9.10) that violates time reversal: it is as tricky to violate as CP. In fact there is a deep connection between T and CP known as the CPT theorem.

⋆ The simplest version of QCD, the strong interaction sector of the standard model, violates the chiral symmetry used to rotate away $\bar\psi i\gamma^5\psi$. Then one could get CP violation with a smaller number of generations. To be compatible with the experimental size of CP violation the coefficient of such a term would have to be so tiny that a modified form of QCD which restores this symmetry (and predicts axions) is usually postulated. Then one is back to the three generation requirement.

Composing the three discrete symmetries we find

$$(CPT)^{-1}\psi(x)CPT = e^{-i\tau}(i\gamma^2\gamma^0\gamma^1\gamma^3)\psi(-x) = -e^{-i\tau}\gamma_5(\psi^\dagger)^T(-x).$$

Now applying this transformation to the bilinears, remembering the antilinear property of CPT, we find

$$\begin{aligned}(CPT)^{-1}\bar{\psi}_A(x)\Gamma\psi_B(x)CPT =& \psi_A^T(-x)\gamma_5\beta\Gamma^*\gamma_5(\psi_B^\dagger)^T(-x)\\ =& -\bar{\psi}_B(-x)\beta\gamma_5\Gamma^\dagger\beta\gamma_5\psi_A(-x)\\ =&(-)^{n_\Gamma}\bar{\psi}_B(-x)\beta\Gamma^\dagger\beta\psi_A(-x)\\ =&(-)^{n_\Gamma}(\bar{\psi}_A(-x)\Gamma\psi_B(-x))^\dagger\end{aligned} \tag{9.11}$$

where n_Γ is the number of Lorentz indices carried by Γ. The CPT theorem states the impossibility of violating this symmetry in quantum field theory. We shall not go through the rigorous proof here, but from the transformation law of the bilinears it is clear what is behind the theorem. Since each Lorentz index must be contracted with another in forming a Lorentz scalar polynomial of the bilinears, all of the $(-)^{n_\Gamma}$'s will cancel in the CPT transform of the polynomial. If we denote the Hamiltonian by some function $H(\bar{\psi}_A(x)\Gamma\psi_B(x))$ of the bilinears, we have

$$\begin{aligned}(CPT)^{-1}H(\bar{\psi}_A(x)\Gamma\psi_B(x))CPT =& H^*((CPT)^{-1}\bar{\psi}_A(x)\Gamma\psi_B(x)CPT)\\ =& H^*((\bar{\psi}_A(-x)\Gamma\psi_B(-x))^\dagger)\end{aligned}$$

where by H^* we mean that all of the complex numbers appearing in the formation of H as a function of the bilinears are complex conjugated. Apart from ordering of operators, the last line is just what we mean by the hermitian conjugate of H, if we set $t = 0$ (conservation of energy means H is constant) and integrate over $\mathbf{x}$. So up to operator ordering questions (which can be sorted out for local interactions), a hermitian Hamiltonian must be CPT invariant.

10. Representations of the Poincaré Group for General Spin

We have so far encountered via simple quantum field theories the relativistic quantum description of free particles of spin 0 or spin 1/2. It is useful at this point to realize that much of what we have found is not really tied to field theory but rather to simple requirements of Poincaré invariance. In quantum mechanics it is a general fact that a symmetry group must be realized by a unitary or antiunitary representation. The latter possibility only occurs for some discrete symmetries (time reversal being the physical example). Our goal in this chapter is to obtain the unitary realization of the Poincaré group for multiparticle states.

The Poincaré group consists of Lorentz transformations together with translations

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu + a^\mu.$$

where Λ preserves Minkowski scalar products

$$\eta_{\rho\sigma}\Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu = \eta_{\mu\nu}.$$

The Λ's can be divided into 4 disjoint sets according to the signs of $\det\Lambda$ and $\Lambda^0{}_0$. This is because it is easy to show from the above property that $(\det\Lambda)^2 = 1$ and $(\Lambda^0{}_0)^2 \geq 1$. Thus a continuous variation of Λ always stays within one of these sets. In the following we restrict ourselves to the *proper* Lorentz Group, *i.e.*with $\det\Lambda = +1$ and $\Lambda^0{}_0 \geq +1$. The complete Lorentz group is then obtained by adjoining parity and time reversal.

Lorentz transformations with $\Lambda^0{}_k = \Lambda^k{}_0 = 0$ are simply rotations and form a subgroup. We know from basic quantum mechanics all the unitary irreducible representations of the Rotation group, namely those labelled by angular momentum $j = 0, 1/2, 1, 3/2, \ldots$. The unitary representations of the Lorentz group must be extensions of these. Let us ask then how to construct this extension for a free massive particle of spin s. Such a particle must be described by a set of at least $2s+1$ momentum space wave functions $f_a(\mathbf{p})$. This much follows just from the Rotation group. We shall find a representation of the Poincaré group with this minimal number of components.

The basic idea, due to Wigner, is to exploit the fact that one can always bring a massive particle to rest by a Lorentz transformation. Define a "standard boost" $B_\mathbf{p}$ which boosts a particle at rest to one with momentum $\mathbf{p}$. Let us introduce momentum eigenstates via

$$|f\rangle \equiv \int d^3p \sum_a |\mathbf{p}, a\rangle\, f_a(\mathbf{p}).$$

Then a momentum eigenstate of the particle can be related to the state at rest by

$$|\mathbf{p}, a\rangle \equiv \sqrt{\frac{m}{\omega(\mathbf{p})}}\, U(B_\mathbf{p})\, |\mathbf{0}, a\rangle\,. \tag{10.1}$$

The multiplicative constant is necessary because we want U to be unitary. To understand this point, notice that the relation of the three momentum $\mathbf{p}'$ of a boosted particle to its initial momentum is nonlinear:

$$p'^k = \Lambda^k{}_l p^l + \Lambda^k{}_0 \omega(\mathbf{p}).$$

The Jacobian of this nonlinear transformation of variables is $\partial(\mathbf{p}')/\partial(\mathbf{p}) = \omega(\mathbf{p}')/\omega(\mathbf{p})$. The easiest way to see this is to observe that $\int d^4p\,\delta(p^2+m^2)$ is a Lorentz invariant; integrating over p^0 then shows that

$\int d^3p/\omega(\mathbf{p})$ is an invariant, which implies the above value for the Jacobian. A general Lorentz transformation on a particle state of momentum $\mathbf{p}$, which boosts it to $\mathbf{p}'$, can be expressed

$$\Lambda = B_{\mathbf{p}'}(B_{\mathbf{p}'}^{-1}\Lambda B_{\mathbf{p}})B_{\mathbf{p}}^{-1}.$$

The transformation in parentheses leaves a particle at rest at rest, and is therefore simply a rotation. Applying $U(\Lambda)$ to (10.1) we have

$$\begin{aligned} U(\Lambda)\,|\mathbf{p},a\rangle =& \sqrt{\frac{m}{\omega(\mathbf{p})}}U(B_{\mathbf{p}'})U(B_{\mathbf{p}'}^{-1}\Lambda B_{\mathbf{p}})\,|\mathbf{0},a\rangle \\ =& \sqrt{\frac{m}{\omega(\mathbf{p})}}U(B_{\mathbf{p}'})\,|\mathbf{0},b\rangle\, D^s_{ba}(B_{\mathbf{p}'}^{-1}\Lambda B_{\mathbf{p}}) \\ =& \sqrt{\frac{\omega(\mathbf{p}')}{\omega(\mathbf{p})}}\,|\mathbf{p}',b\rangle\, D^s_{ba}(B_{\mathbf{p}'}^{-1}\Lambda B_{\mathbf{p}}). \end{aligned}$$

Here $D^s_{ba}(R)$ is just the standard representation matrix of the rotation group with spin s. It is now easily checked that this defines a unitary representation of the Lorentz group on single particle states of spin s. Of course, on momentum eigenstates of a free particle, space-time translations are trivially realized by multiplication of the state by the phase $e^{-ia\cdot p}$ for the space-time translation by amount a^μ.

Let us return to the "standard boost" $B_{\mathbf{p}}$. It is clearly not uniquely determined since it can be preceded by an arbitrary rotation and followed by a rotation about the axis parallel to $\mathbf{p}$. There are two widely used choices for this boost. The simplest choice is the pure boost parallel to $\mathbf{p}$ which we will call $B^0_{\mathbf{p}}$. The second choice is dictated by choosing helicity states. It is described as follows. First agree that the spin states of the particle at rest be labelled by λ the eigenvalue of J_3. Then first boost the particle along the z axis to momentum $|\mathbf{p}|\hat{z}$. Then apply a rotation that carries the z axis to the direction of $\mathbf{p}$. This latter rotation can be taken to be

$$R_0(\mathbf{p}) \equiv e^{-i\phi J_3}e^{-i\theta J_2}e^{+i\phi J_3}$$

where (θ,ϕ) are the polar angles of $\mathbf{p}$. Then the helicity preserving standard boost is given by

$$B^h_{\mathbf{p}} = R_0(\mathbf{p})B^0(|\mathbf{p}|\hat{z}).$$

Then, clearly, $|\mathbf{p},\lambda\rangle \equiv \sqrt{m/\omega}B^h_{\mathbf{p}}\,|\mathbf{0},\lambda\rangle$ is an eigenstate of momentum $\mathbf{p}$ with helicity λ, since rotations do not change helicity. This is easy to see from the general transformation law. For a massive particle, helicity *is* changed by a general Lorentz transformation, *i.e.* $B_{\mathbf{p}'}^{-1}\Lambda B_{\mathbf{p}}$ can be any rotation. However, if one considers the massless limit of this rotation for any fixed Λ, he discovers that it always approaches a rotation about the z axis. Thus for *massless* particles helicity is actually a Lorentz invariant. Thus it can be consistent for massless particles with spin to exist in only one helicity state.

Having understood how a single particle state transforms under Lorentz transformations it is straightforward to find the transformation law for states with any number of free particles of varying mass and spin, which can be viewed as tensor products of single particle states. We can incorporate bose or fermi statistics by introducing a vacuum state $|0\rangle$ and creation and annihilation operators for each species of particle,*e.g.*:

$$|\mathbf{p},\lambda\rangle \equiv b^{i\dagger}_\lambda(\mathbf{p})\,|0\rangle\,,$$

where i labels the species. When we are dealing with only a limited number of species, we typically choose different letters for different types of particles, *e.g.* $a, a^\dagger$ for neutral scalar particles, $b, b^\dagger$ $(d, d^\dagger)$ for particles

(antiparticles), etc. Then the transformation law for a general multiparticle state is completely defined by

$$\begin{aligned}
U(\Lambda)\,|0\rangle =& |0\rangle \\
U(\Lambda)b^{i\dagger}_{\lambda}(\mathbf{p})U^{-1}(\Lambda) =& b^{i\dagger}_{\lambda'}(\Lambda\mathbf{p})D^{s}_{\lambda'\lambda}(B^{-1}_{\Lambda\mathbf{p}}\Lambda B_{\mathbf{p}}) \\
U(\Lambda)b^{i}_{\lambda}(\mathbf{p})U^{-1}(\Lambda) =& b^{i}_{\lambda'}(\Lambda\mathbf{p})D^{s*}_{\lambda'\lambda}(B^{-1}_{\Lambda\mathbf{p}}\Lambda B_{\mathbf{p}}).
\end{aligned}$$

Of course the last transformation law is just the Hermitian conjugate of the second one since U is supposed to be unitary. We complete the description by writing down the energy and momentum operators

$$\begin{aligned}
H =& \int d^3p \sum_i \sqrt{\mathbf{p}^2 + m_i^2} \sum_\lambda b^{i\dagger}_{\lambda}(\mathbf{p})b^{i}_{\lambda}(\mathbf{p}) \\
\mathbf{P} =& \int d^3p\mathbf{p} \sum_i \sum_\lambda b^{i\dagger}_{\lambda}(\mathbf{p})b^{i}_{\lambda}(\mathbf{p}).
\end{aligned}$$

The above discussion might mislead one into thinking that the problem of free quantum field theory for any spin is completely solved. Indeed, we have solved the problem of constructing a relativistic quantum description of any number of free particles with any spins. However, the equivalence of this description to local field theory is not yet transparent. We have explicitly seen how this works for spin 0 and 1/2. which are described by scalar and Dirac fields respectively. The scalar field is supposed to have the Lorentz transformation properties $\phi'(x') = \phi(x)$. We relate this to the general discussion by first identifying the momentum eigenstates with a creation operator applied to the vacuum $|\mathbf{p}\rangle = a^\dagger(\mathbf{p})\,|0\rangle$. Assuming the vacuum is Lorentz invariant, we conclude from the general discussion that

$$U(\Lambda)a^\dagger(\mathbf{p})U^{-1}(\Lambda) = \sqrt{\frac{\omega(\mathbf{p}')}{\omega(\mathbf{p})}}a^\dagger(\mathbf{p}').$$

With this result we can then evaluate how the scalar field transforms

$$\begin{aligned}
U^\dagger(\Lambda)\phi(x)U(\Lambda) =& \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}\sqrt{\frac{\omega(\mathbf{p}')}{\omega(\mathbf{p})}}(a(\mathbf{p}')e^{ix\cdot p} + a^\dagger(\mathbf{p}')e^{-ix\cdot p}) \\
=& \int \frac{d^3p'}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p}')}}(a(\mathbf{p}')e^{ix\cdot\Lambda p'} + a^\dagger(\mathbf{p}')e^{-ix\cdot\Lambda p'}) \\
=& \phi(\Lambda^{-1}x)
\end{aligned}$$

as desired.

Notice that Lorentz covariance alone is achieved by the positive frequency part of the field

$$\phi^+(x) = \int \frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}a(\mathbf{p})e^{ix\cdot p}.$$

But such a field would not commute with its adjoint at space-like separations as a local field must.

$$[\phi^{(+)}(x), \phi^{(-)}(y)] = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega}e^{i(x-y)\cdot p}.$$

Compare this to the result for the total field

$$[\phi(x), \phi(y)] = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega}(e^{i(x-y)\cdot p} - e^{-i(x-y)\cdot p}).$$

If $(x-y)^2 > 0$ there is a Lorentz frame for which $x^0 = y^0$. In that frame the r.h.s. is manifestly zero. It must be zero in all frames by Lorentz covariance. Notice also that the all important minus sign on the r.h.s

came because we chose commutation relations for $a, a^\dagger$. Had we tried to make the scalar particles fermions by imposing anticommutation relations, the two terms would add and locality would be lost. This is the famous spin-statistics connection for the scalar field. For the Dirac field fermi statistics was necessary for stability rather than locality, but the spin-statistics connection is nonetheless fixed.

The Dirac field shows us how we must generalize these considerations to develop a field theory for particles with spin. As shown in an exercise the Dirac field transforms under Lorentz transformations as

$$\psi'(x') = e^{-i\lambda_{\mu\nu}\sigma^{\mu\nu}/2}\psi(x),$$

where $(e^{-\lambda})^\mu_\nu = \Lambda^\mu_\nu$. The 4×4 matrices $\sigma^{\mu\nu}$ provide a finite dimensional representation of the Lorentz group which is necessarily *not* unitary. This nonunitarity is associated with the noncompactness of the Lorentz group. The nonunitarity of these matrices does not conflict with the unitarity of the action of the Lorentz group on the state space which is just that on multi-particle states we have just discussed. By expressing the field in terms of creation and annihilation operators we can see how the unitary representation on particle states induces the desired field transformation.

From the transformation properties of a spin 1/2 particle it follows that

$$\begin{aligned}
U^\dagger(\Lambda)d^\dagger_\lambda(\mathbf{p})U(\Lambda) =& \sqrt{\frac{\omega(\mathbf{p}')}{\omega(\mathbf{p})}}d^\dagger_{\lambda'}(\mathbf{p}')D^{1/2}_{\lambda'\lambda}(B^{-1}_{\mathbf{p}'}\Lambda^{-1}B_{\mathbf{p}})\\
U^\dagger(\Lambda)b_\lambda(\mathbf{p})U(\Lambda) =& \sqrt{\frac{\omega(\mathbf{p}')}{\omega(\mathbf{p})}}b_{\lambda'}(\mathbf{p}')D^{1/2*}_{\lambda'\lambda}(B^{-1}_{\mathbf{p}'}\Lambda^{-1}B_{\mathbf{p}}).
\end{aligned}$$

where $p' = \Lambda^{-1}p$. Focus on the way b enters the Dirac field:

$$\begin{aligned}
&U^\dagger(\Lambda)\int\frac{d^3p}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}\sum_\lambda b_\lambda(\mathbf{p})u_\lambda(\mathbf{p})e^{ix\cdot p}U(\Lambda)\\
&=\int\frac{d^3p'}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p}')}}b_{\lambda'}(\mathbf{p}')u_\lambda(\Lambda\mathbf{p}')e^{i\Lambda^{-1}x\cdot p'}D^{1/2*}_{\lambda'\lambda}(B^{-1}_{\mathbf{p}'}\Lambda^{-1}B_{\mathbf{p}})\\
&=\int\frac{d^3p'}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p}')}}b_{\lambda'}(\mathbf{p}')u_\lambda(\Lambda\mathbf{p}')e^{i\Lambda^{-1}x\cdot p'}D^{1/2}_{\lambda\lambda'}(B^{-1}_{\mathbf{p}}\Lambda B_{\mathbf{p}'})\\
&=e^{-i\lambda_{\mu\nu}\sigma^{\mu\nu}/2}\int\frac{d^3p'}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p}')}}\sum_{\lambda'}b_{\lambda'}(\mathbf{p}')u_{\lambda'}(\mathbf{p}')e^{i\Lambda^{-1}x\cdot p'},
\end{aligned}$$

which is exactly the desired field transformation. The term involving $d^\dagger$ works in an exactly similar way. In obtaining the last line we used the identity

$$e^{-i\lambda_{\mu\nu}\sigma^{\mu\nu}/2}u_\lambda(\mathbf{p}) = \sum_{\lambda'}u_{\lambda'}(\Lambda\mathbf{p})D^{1/2}_{\lambda'\lambda}(B^{-1}_{\Lambda\mathbf{p}}\Lambda B_{\mathbf{p}}),$$

which is a simple consequence of the way Lorentz covariance works in the first quantized interpretation of the Dirac wave function.

Clearly the first step in generalizing to higher spin fields is to classify all of the finite dimensional representations of the Lorentz group. We would like to find the possible representation matrices so that a

multi-component field $\psi_\alpha(x)$ will have the transformation law

$$\psi'_\alpha(x) = D_{\alpha\beta}(\Lambda)\psi_\beta(\Lambda^{-1}x).$$

Let us first cast the algebra of Lorentz generators

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\nu\rho}M_{\mu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} + \eta_{\nu\sigma}M_{\mu\rho})$$

in terms of the generators for rotations $J_k = \epsilon_{klm}M_{lm}/2$ and for boosts $K_k = M_{0k}$:

$$\begin{aligned}[J_k, J_l] =& i\epsilon_{klm}J_m \\ [J_k, K_l] =& i\epsilon_{klm}K_m \\ [K_k, K_l] =& -i\epsilon_{klm}J_m.\end{aligned}$$

Now notice that the linear combinations $\mathbf{J}_\pm \equiv (\mathbf{J} \pm i\mathbf{K})/2$ satisfy the algebra of two mutually commuting angular momentum algebras

$$\begin{aligned}[J_k^\pm, J_l^\pm] =& i\epsilon_{klm}J_m^\pm \\ [J_k^+, J_l^-] =& 0.\end{aligned}$$

We know from elementary quantum mechanics what all of the finite dimensional representations of the rotation group are: they are labelled by the eigenvalues $j(j+1)$ of the Casimir operator $\sum_k J^{k2}$ with $j = 0, 1/2, 1, 2, \ldots$ all nonnegative integers and half integers. The representation j has dimension $2j+1$. It follows that all the finite dimensional representations of the Lorentz group are labelled by the pair of eigenvalues $j_+(j_+ + 1), j_-(j_- + 1)$ of the pair of Casimir operators $\sum_k J_\pm^{k2}$ where $2j_+, 2j_-$ are any pair of nonnegative integers. All of these representations are equivalent to a unitary representation, *i.e.* $\mathbf{J}_\pm$ are both represented by hermitian matrices. This means that the rotation generators $\mathbf{J} = \mathbf{J}_+ + \mathbf{J}_-$ are represented by hermitian matrices, but the boost generators $\mathbf{K} = -i(\mathbf{J}_+ - \mathbf{J}_-)$ are represented by antihermitian matrices. Thus the finite dimensional representations of the Lorentz group are not equivalent to unitary ones. This is associated with the fact that the Lorentz group is noncompact. We already encountered this nonunitarity in the representation of the Lorentz group by gamma matrices.

We denote the representation matrices by $D(j_+, j_-)$. The simplest nontrivial representations are $D(1/2, 0)$ and $D(0, 1/2)$. In the first the generators J_- are represented by 0 and the generators J_+ by $\sigma/2$. This means that the angular momentum is represented by $\mathbf{J} = \sigma/2$ and the boost generators by $\mathbf{K} = -i\sigma/2$. The other two dimensionial representation has the same representative for $\mathbf{J}$ but the boost is represented by $\mathbf{K} = +i\sigma/2$. It is clear that these representations are not equivalent to each other, since any similarity transformation which could reverse the sign of $\mathbf{K}$ would do the same to $\mathbf{J}$. However these two inequivalent representations can be related by complex conjugation. In fact it is easy to see from the properties of the Pauli matrices that $D(1/2, 0)^* = \sigma_2 D(0, 1/2)\sigma_2$. In general, the representation $D(k, m)^*$ is equivalent to $D(m, k)$[⋆], so the only real irreducible representations have $j_+ = j_-$. Of course $D(k, m) \oplus D(m, k)$ is real but it is also reducible.

Notice that parity reverses the sign of the boost generators but not the sign of the angular momentum. Thus the representations $D(k, m)$ and $D(m, k)$ are also related by parity. We encountered this fact with the Dirac field which admits the parity symmetry. It exploits the reducible representation $D(1/2, 0) \oplus D(0, 1/2)$

⋆ In $D(k, m)^*$ the generators are $-\mathbf{J}^*, -\mathbf{K}^*$ so *e.g.* J_+ is represented by $-J^* - iK^* = -(J_-)^*$ and *vice versa.*

to achieve this. To make this quite explicit we note that the representatives of the Lorentz generators are the 4×4 matrices

$$\mathbf{J} = \frac{1}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}$$
$$\mathbf{K} = \frac{-i}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix}$$

which are just the components of $\sigma^{\mu\nu}/2$ constructed out of gamma matrices in the so-called natural representation

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$
$$\boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}.$$

Defining $\sigma^\mu = (I, \boldsymbol{\sigma})$, and $\bar{\sigma}^\mu = (I, -\boldsymbol{\sigma})$ we can unify these two in the single equation

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}$$

The potential spin content of a field in a given representation is typically richer than one might desire. For example, since $J = J_+ + J_-$ the representations of the rotation group contained in $D(k,m)$ include all spins that arrise from adding spin k to spin m: $|k-m|, |k-m|+1, \ldots, k+m$. Thus if our desire is to describe a given spin, depending on our choice of representation, we might bring in several other spins as well. The choice of field content is not unambiguous. Even for spin 1/2 we have noted various possibilities, *e.g.* Dirac, Majorana, and Weyl. Weyl fermions make use of $D(1/2,0)$, but since this is not a real representation, the hermitian conjugate field, which transforms under $D(0,1/2)$, must also be introduced and represents the anti-particle.

When we come to spin 1, two possiblities come to mind. $D(1,0) \oplus D(0,1)$ or $D(1/2,1/2)$. The latter contains potentially both spin 1 and spin 0 and is in fact the representation of a four-vector field. The former seems to contain spin 1 twice. It is easy to see that it is the antisymmetric tensor product of $D(1/2,1/2)$ with itself and thus represents an antisymmetric second rank tensor. The field strengths $F_{\mu\nu}$ of electromagnetism spring to mind, so we might decide that the first choice is best. However, we know that in quantum mechanics it is necessary to use the potentials A_μ which transform under the second choice. Then gauge invariance is essential to eliminate unwanted spin states.

This general discussion of higher spin serves to indicate some of the subtleties and complexities that must be confronted. In fact, consistent fully interacting quantum field theories have never been constructed for spins higher than 2 (the graviton). Furthermore, ultraviolet divergences have so far caused incurable difficulties for theories with spin higher than 1 including quantum gravity. Since gravity is very much present in the real world, it is clear that there is much to do before we can claim that quantum field theory can describe all of physics.

11. Quantum Field Equations with External Fields

11.1. Electromagnetic Fields

The coupling of the Dirac equation to an external elctromagnetic field is dictated by the principle of gauge invariance. In classical electrodynamics, it is possible to avoid potentials and formulate all equations of motion in terms of the electric and magnetic fields $F_{\mu\nu}$. However, the potential $A_\mu(x)$ is indispensable to an economical description of the coupling of a quantum particle to electomagnetism. Fundamentally, this is because the Hamiltonian and Lagrangian play a much more central role in quantum dynamics than in classical dynamics, and the potential appears explicitly in the Hamiltonian and Lagrangian. (Recall that the Schrödinger equation involves the Hamiltonian explicitly.) The field strengths are related to the potential via

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x),$$

but it is clear that the potential is not given uniquely in terms of the field strength: If the potential is changed by a *gauge transformation*

$$A_\mu \to A_\mu + \partial_\mu \Lambda, \tag{11.1}$$

the field strength $F_{\mu\nu}$ is unchanged. It is therefore important to introduce the potential into the Schrödinger equation in a way which preserves gauge invariance.

Let us first ask how gauge invariance is realized in classical particle electrodynamics. In order that the Euler-Lagrange equations reproduce the Lorentz force law, the scalar and vector potentials $(A^0, \mathbf{A})$ must enter the Lagrangian through the terms

$$-qA^0(x) + q\dot{\mathbf{x}} \cdot \mathbf{A}(x).$$

Because of the term linear in velocity, the momentum conjugate to $\mathbf{x}$ becomes

$$\mathbf{p} = \mathbf{p}_{\mathbf{A}=0} + q\mathbf{A},$$

where $\mathbf{p}_{\mathbf{A}=0}$ is the conjugate momentum with vanishing vector potential. Furthermore, when we form the Hamiltonian $H = \dot{\mathbf{x}} \cdot \mathbf{p} - L$, the term linear in velocity cancels, so

$$H = H_0(\mathbf{p}_{\mathbf{A}=0}, \mathbf{x}) + qA^0 = H_0(\mathbf{p} - q\mathbf{A}, \mathbf{x}) + qA^0.$$

If we subject A to a gauge transformation, the Lagrangian changes by the amount

$$q(\partial_0 \Lambda + \dot{\mathbf{x}} \cdot \nabla \Lambda) = q\frac{d\Lambda}{dt},$$

a total time derivative, so that the action $\int_{t_1}^{t_2} L$ changes by the amount $q\Lambda(\mathbf{x}(t_2), t_2) - q\Lambda(\mathbf{x}(t_1), t_1)$ and the added terms have no effect on the Euler-Lagrange equations. At first sight the Hamiltonian doesn't look invariant, but notice that the transformation

$$\mathbf{p}' = \mathbf{p} - q\nabla\Lambda \qquad\qquad \mathbf{x}' = \mathbf{x}$$

is a canonical transformation with generating function $W_2(\mathbf{x}, \mathbf{p}', t) = \mathbf{x} \cdot \mathbf{p}' + q\Lambda(\mathbf{x}, t)$. Furthermore if the generating function is time dependent the canonical transformation includes changing the Hamiltonian by $\partial_t W_2 = q\partial_t \Lambda$, so that after the canonical transformation the gauge transformed Hamiltonian is identical to the old one with the substitutions $\mathbf{x} \to \mathbf{x}'$, $\mathbf{p} \to \mathbf{p}'$.

The way A_μ enters the Schrodinger equation for a charged particle is now clear. The substitution $\mathbf{p} \to \mathbf{p} - q\mathbf{A}$ corresponds in the Schrödinger equation to

$$\frac{1}{i}\nabla \to \frac{1}{i}\nabla - q\mathbf{A}.$$

Furthermore the addition of $qA^0 = -qA_0$ to the Hamiltonian is prescribed by the substitution rule

$$\frac{1}{i}\frac{\partial}{\partial t} \to \frac{1}{i}\frac{\partial}{\partial t} - qA_0,$$

so both substitutions can be given the compact expression

$$\partial_\mu \to \partial_\mu - iqA_\mu \tag{11.2}$$

which is known as the minimal substitution rule. The gauge invariance of the Schrödinger equation is achieved by postulating in addition to (11.1) the change

$$\psi \to e^{iq\Lambda}\psi. \tag{11.3}$$

If we recall that in the semi-classical approximation $\psi \sim e^{iS}$, where S is the classical action, we recognize that this change of the wave function under gauge transformation is the quantum analogue of the classical change in the action.

The gauge transformation makes an arbitrary, local, redefinition of the phase of the wave function. In fact, one could take the attitude that invariance under such local phase changes is desirable from a physical point of view. (Since global phase changes are unobservable perhaps local ones should also be.) In that case one would be forced to introduce the electromagnetic field to realize the invariance! It is obvious that the Schrödinger equation is invariant under the combined changes (11.1) and (11.3). As a special case, the Dirac equation is invariant under the same gauge transformations. When we interpret the Dirac equation as a field equation, (11.3) is a transformation on fields as is (11.1), so the two are really on similar footing.

To sum up the above discussion, we display the Dirac equation in a potential A_μ:

$$i\gamma \cdot (\partial - iqA)\psi = m\psi.$$

Also the corresponding second quantized Hamiltonian is given by

$$H_A(t) = \int d^3x(\psi^\dagger(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi - j^\mu A_\mu).$$

Notice that a simple consequence of the Dirac equation is current conservation

$$\partial_\mu j^\mu = 0,$$

even when $A \neq 0$. This can be understood as a consequence of the gauge invariance of quantum evolution. To see this we have to consider the unitary evolution operator $U(t,t_0)$ ⋆ defined by

$$i\partial_t U(t,t_0) = H_S(t)U(t,t_0) \qquad\qquad U(t_0,t_0) = I, \tag{}$$

where H_S is the Schrödinger picture Hamiltonian. If we make a small change δA_μ in A_μ, **keeping the**

⋆ U gives the unitary transformation between Heisenberg and Schrödinger pictures. Its analogue in classical mechanics is the generator $S(q,P,t)$ of the canonical transformation mapping the initial phase space variables Q,P to those at time t, q,p. The analogue of the following equation for U is the Hamilton-Jacobi equation for S,

$$\frac{\partial S}{\partial t} = -H(q,\frac{\partial S}{\partial q},t).$$

Schrödinger picture dynamical variables fixed, U changes by the amount†

$$\begin{aligned}\delta U(t,t_0) =& i\int_{t_0}^{t} dt' d^3x U(t,t')\delta A_\mu(x') j^\mu(\mathbf{x}) U(t',t_0)\\ =& U(t,t_0) i\int_{t_0}^{t} dt' d^3x \delta A_\mu(x') j^\mu(\mathbf{x},t')\end{aligned} \tag{11.4}$$

with $j^\mu(\mathbf{x},t') = U^\dagger(t',t_0) j^\mu(\mathbf{x}) U(t',t_0)$ the Heisenberg picture current operator‡. To obtain this result simply differentiate (11.4) with respect to time and show that it satisfies () to first order in δA. Use is also made of the closure relation

$$U(t,t')U(t',t_0) = U(t,t_0)$$

which is a simple consequence of the differential equation satisfied by U. Under a gauge transformation $\delta A = \partial\Lambda$ so we see that

$$\delta U = Ui\int_{t_0}^{t} d^4x \partial_\mu \Lambda j^\mu(x).$$

The requirement that U be invariant under gauge transformations that vanish at t_0, t is then that $\partial_\mu j^\mu = 0$.

11.2. Nonabelian Gauge Fields

The local phase transformation (11.3) on charged fields can be generalized (Yang-Mills). Suppose that ψ carries an internal index k. Then in place of (11.3) we can consider

$$\psi_k(x) \to \sum_l \Omega_{kl}(x)\psi_l(x)$$

or, with suppressed indices

$$\psi(x) \to \Omega(x)\psi(x) \tag{11.5}$$

with $\Omega(x)$ in a unitary matrix representation of some continuous group G. In this language (11.3) corresponds to the choice $G = U(1)$, multiplication by a phase, an abelian group. If we require the dynamics to be

† In the classical theory a change in the parameters of the Hamiltonian would give rise to a change in S satisfying

$$\frac{\partial \delta S}{\partial t} = -\sum_k \frac{\partial \delta S}{\partial q_k}\frac{\partial H}{\partial p_k} - \delta H(q, \frac{\partial S}{\partial q}, t).$$

Now if we put $q = q(t)$ and $\partial S/\partial q = p(t)$ the Hamilton-Jacobi equation just says that $q(t), p(t)$ satisfy Hamilton's equations so $\partial H/\partial p = \dot{q}$ and we have

$$\frac{d\delta S}{dt} = -\delta H(q(t), p(t), t)$$

or $\delta S = -\int_{t_0}^{t} dt' \delta H(q(t'), p(t'), t')$, the classical analogue of the following equation.

‡ Equation (11.4) has a generalization to an arbitrary physical system: If one makes any small change δH_S in the Schrödinger picture hamiltonian, with the Scrödinger picture operators fixed, the corresponding change in the evolution operator $U(t,t_0)$ is given by $\delta U = -iU\int_{t_0}^{t} dt' \delta H(t')$, where $\delta H(t) \equiv U^\dagger \delta H_S U$ is the change in the Schrödinger hamiltonian transformed to the Heisenberg picture, which is the same as the change in the Heisenberg hamiltonnian $H(t)$, keeping the Heisenberg dynamical variables fixed. The proof of this is exactly the same as that of Eq.(11.4).

invariant under the nonabelian local gauge transformations, we must introduce a nonabelian gauge field to absorb the noncovariance of

$$\partial_\mu \psi \to \Omega(\partial_\mu + \Omega^{-1}\partial_\mu \Omega)\psi.$$

In analogy with the electromagnetic case, we need to introduce a matrix valued potential $A_\mu(x)$ via the substitution

$$\partial_\mu \to \partial_\mu - igA_\mu(x) \equiv D_\mu. \tag{11.6}$$

Then the requirement $D_\mu \psi \to \Omega D_\mu \psi$ translates to

$$A_\mu \to \Omega A_\mu \Omega^{-1} - \frac{i}{g}\partial_\mu \Omega \Omega^{-1}. \qquad (11.7).$$

Clearly A takes values in the Lie algebra of the group G. If ψ is a Dirac field, a gauge invariant dynamics is given by the field equation

$$i\gamma \cdot D\psi = m\psi,$$

or the corresponding second quantized Hamiltonian

$$H_A(t) = \int d^3x(\psi^\dagger(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi - g\bar{\psi}A_\mu\gamma^\mu\psi).$$

Just as in the electromagnetic case we can consider how the quantum evolution operator changes under a small change of A, and identical steps lead to:

$$\delta U(t,t_0) = U(t,t_0)i\int_{t_0}^{t} dt' d^3x \bar{\psi} g \delta A_\mu \gamma^\mu \psi.$$

An infinitesimal gauge transformation $\Omega = I + ig\epsilon G$ corresponds to

$$\delta A_\mu = \epsilon(\partial_\mu G - ig[A_\mu, G]).$$

For G which vanish initially and finally the corresponding change in U is

$$\begin{aligned}\delta U(t,t_0) =& U(t,t_0)ig\epsilon\int_{t_0}^{t} dt' d^3x \\ & G_{ab}(-\partial_\mu(\bar{\psi}_a\gamma^\mu\psi_b) - ig(A_\mu)_{ca}\bar{\psi}_c\gamma^\mu\psi_b + ig(A_\mu)_{bc}\bar{\psi}_a\gamma^\mu\psi_c)\end{aligned}$$

so gauge invariance implies the following generalization of current conservation ("covariant conservation")

$$D_\mu j^\mu \equiv \partial_\mu j^\mu - ig[A_\mu, j^\mu] = 0,$$

where the current is a matrix operator

$$j^\mu(x)_{ba} \equiv \bar{\psi}_a(x)\gamma^\mu\psi_b(x).$$

11.3. External Gravitational Fields

According to the Principle of Equivalence, an external gravitational field is described by introducing a space-time dependent metric $\eta_{\mu\nu} \to g_{\mu\nu}(x)$ which then must enter the field equations in a generally covariant way. This prescription suffices for bosonic fields but new concepts must be brought in for fermionic fields. At this stage we shall confine our discussion to a real scalar field. The minimal generally covariant classical action is given by

$$S = -\frac{1}{2}\int d^4x\sqrt{-g}(g^{\mu\nu}(x)\partial_\mu\phi\partial_\nu\phi + m^2\phi^2).$$

To construct the Hamiltonian, we first define the conjugate momentum

$$\pi(x) \equiv \frac{\delta S}{\delta\dot{\phi}(x)} = -\sqrt{-g}g^{0\nu}(x)\partial_\nu\phi.$$

Then the Hamiltonian is

$$\begin{aligned} H(t) &= \int d^3x(\pi(x)\dot{\phi}(x) + \frac{1}{2}\sqrt{-g}(g^{\mu\nu}(x)\partial_\mu\phi\partial_\nu\phi + m^2\phi^2)) \\ &= \int d^3x\left(-\frac{1}{2}\frac{\pi^2}{g^{00}\sqrt{-g}} + \frac{g^{0k}}{g^{00}}\pi\partial_k\phi + \frac{1}{2}\sqrt{-g}\left(g^{kl} - \frac{g^{0k}g^{0l}}{g^{00}}\right)\partial_k\phi\partial_l\phi + \frac{m^2}{2}\sqrt{-g}\phi^2\right). \end{aligned}$$

Just as with gauge fields we can ask how the evolution operator changes under a small change $\delta g_{\mu\nu}$ in the metric, $\delta U = -iU\int_{t_0}^t dt'\delta H(t')$, where δH is computed holding ϕ, π fixed. The easiest way to do this is to evaluate the change in the Lagrangian L at fixed $\phi, \dot{\phi}$. Since L is related to H by a Legendre transform, $\delta H = -\delta L$.

$$\delta L = -\frac{1}{2}\int d^3x\sqrt{-g}\delta g^{\mu\nu}(x)T_{\mu\nu}(x)$$

with $T_{\mu\nu}$ the energy momentum tensor

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}(g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi + m^2\phi^2).$$

Thus we have finally*

$$\delta U = iU\frac{1}{2}\int_{t_0}^{t} d^4x'\sqrt{-g}\delta g_{\mu\nu}(x')T^{\mu\nu}(x').$$

Under the infinitesimal general coordinate transformation $x^\mu = x'^\mu + \xi^\mu(x')$, the metric changes according to

$$\begin{aligned} g_{\mu\nu}(x) =& g'_{\mu\nu}(x') - \frac{\partial\xi^\rho}{\partial x^\mu}g_{\rho\nu} - \frac{\partial\xi^\rho}{\partial x^\nu}g_{\rho\mu} \\ =& g'_{\mu\nu}(x') - D_\mu\xi_\nu - D_\nu\xi_\mu + \Gamma^\rho_{\mu\sigma}\xi^\sigma g_{\rho\nu} + \Gamma^\rho_{\nu\sigma}\xi^\sigma g_{\rho\mu} \\ =& g'_{\mu\nu}(x') - D_\mu\xi_\nu - D_\nu\xi_\mu + \xi^\rho\partial_\rho g_{\mu\nu} \\ =& g'_{\mu\nu}(x) - D_\mu\xi_\nu - D_\nu\xi_\mu \end{aligned}$$

An infinitesimal change of integration variables is just a surface term:

$$\int d^4x'\mathcal{L}'(x') = \int d^4x(1-\partial_\rho\xi^\rho)(1-\xi^\rho\partial_\rho)\mathcal{L}'(x) = \int d^4x\mathcal{L}'(x) - \int d^4x\partial_\rho(\xi^\rho\mathcal{L}'(x))$$

so choosing $\delta g_{\mu\nu} = -D_\mu\xi_\nu - D_\nu\xi_\mu$ should give $\delta U = 0$ for all ξ vanishing sufficiently rapidly at infinity, if the quantum field dynamics is invariant under general coordinate transformations. Thus general coordinate

* Note that since $g^{\mu\nu}$ is the inverse matrix to $g_{\mu\nu}$, $\delta g^{\mu\nu} = -g^{\mu\rho}\delta g_{\rho\sigma}g^{\sigma\nu}$.

invariance implies that the energy momentum tensor is covariantly conserved:

$$D_\mu T^{\mu\nu} = 0.$$

In the limit of flat space (no gravity) this condition reduces to ordinary energy-momentum conservation.

11.4. Asymptotic States and Matrix Elements

In discussing time dependent processes, it is convenient to introduce asymptotic states which are eigenstates of $H(\pm\infty)$. We denote by $|in\rangle$ the ground state of $H(-\infty)$ and by $|out\rangle$ the ground state of $H(+\infty)$. The normal situation will be one in which all external fields vanish at sufficiently early and late times. Thus $|in\rangle$ and $|out\rangle$ will typically be ground states of $H_0(-\infty)$ and $H_0(+\infty)$ respectively. Although these operators are not the same (because their time evolution is governed by H not H_0), the spectra of the two Hamiltonians *are* identical: $H_0(t) = U^{-1}(t,-\infty)H_{0S}U(t,-\infty)$. By convention we are identifying the Schrödinger and Heisenberg pictures at $t = -\infty$. Thus, if $|in\rangle$ is the ground state of $H_0(-\infty) = H_{0S}$, the state $\langle in|\, U(\infty,-\infty)$ is an eigenstate of $H_0(+\infty)$ with the same eigenvalue and hence the ground state. Thus we can and shall fix phases by defining

$$\langle out| \equiv \langle in|\, U(\infty,-\infty).$$

We stress that this is the true "out" state only when $H_S(\infty) = H_S(-\infty) \equiv H_0$.

If the time dependence of H_S is *adiabatic*, *i.e.* very slow on the time scale set by the level spacings, the Adiabatic Theorem assures us that an eigenstate of $H_S(-\infty)$ evolves to an eigenstate of $H_S(t)$ for all t for which adiabatic conditions apply, even after a long enough time to change H_S by a finite amount. For example, the state $|in\rangle$ will be an eigenstate of $H(t)$ for all t for which adiabatic time variation applies. In particular, the ground state eigenvalue $E_G(t)$ must not get close to the next higher eigenvalue as t varies. If this situation holds for all time, it follows that the state $|in\rangle$ is a phase times the state $|out\rangle$, or $\langle out|$ is this same phase times $\langle in|$. This phase is easily evaluated in terms of the time dependent ground state energy $E_G(t)$ of $H_S(t)$ by applying the Schrödinger equation to $\langle in|\, U(t,-\infty)\, |in\rangle$ and using the adiabatic theorem $H_S(t)U(t,-\infty)\, |in\rangle = E_G(t)U(t,-\infty)\, |in\rangle$:

$$\langle out|in\rangle = \exp\left\{-i\int_{-\infty}^{\infty} dt E_G(t)\right\} \qquad \text{Adiabatic Conditions.}$$

Note carefully that adiabatic conditions would *not* apply if the ground state energy got close to an excited level as time evolved. In particular, it would not apply in processes with pair production when $|\,\langle out|in\rangle\,| < 1$.

It is only for very particular external fields that one can solve the Dirac equation or any field equation exactly. Important examples include the static Coulomb potential for which one can find all the energy eigenvalues and eigenstates, arbitrary constant field strengths, and plane waves. The Coulomb potential is of special importance since the exact solution in that case is the starting point for the relativistic theory of atomic energy levels. Still the complete dynamics is never exactly given by these special cases and perturbation theory is the important tool for evaluating corrections to the exactly soluble (idealized) case, which can be zero external fields or one of the above cases. When we quantize the electromagnetic field perturbation theory is essentially our *only* tool for computing radiative corrections due to the quantum nature of the electromagnetic field.

12. External Field Perturbations

As an important example, let us apply perturbation theory to the Dirac field in an external electromagnetic field. Since we shall work in interaction picture throughout, we shall not append the I subscript to indicate interaction picture: that will be understood throughout this section. Then the persistence amplitude in the presence of A is

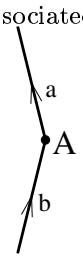
$$\begin{aligned}\langle out|in\rangle_A &= \langle 0|\, T e^{i\int d^4x j^\mu(x) A_\mu(x)}\, |0\rangle \\ &= \sum_{n=0}^{\infty} \frac{i^n}{n!}(\frac{q}{2})^n \int d^4x_1 \cdots d^4x_n A_{\mu_1} \cdots A_{\mu_n} \\ &\qquad \langle 0|\, T[[\bar\psi(x_1), \gamma^{\mu_1}\psi(x_1)] \cdots [\bar\psi(x_n), \gamma^{\mu_n}\psi(x_n)]]\, |0\rangle\,.\end{aligned}$$

It is now a matter of applying Wick's theorem to evaluate the nth term of the series.

We shall organize the calculation with Feynman diagrams. The Wick expansion expresses the time ordered product in terms of propagators completely characterized by two points. Each propagator is represented by a line connecting the two points, directed from the argument of $\bar\psi$ to that of ψ:

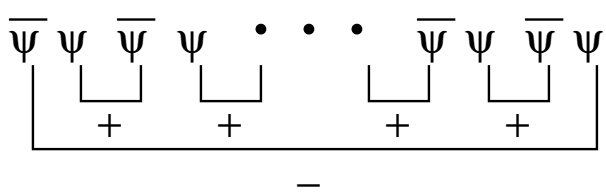
$$S_F(x-y)_{ab}.$$

The lines terminate on vertices associated with the field A:



$$iq\int d^4x A_\mu(x)\gamma^\mu_{ab}.$$

For each closed loop there is a trace over Dirac indices and a multiplicative factor of -1. This last factor is due to the anticommuting property of the Dirac field and arises because the order of fields in the product of currents that is contracted to form the loop differs from that of the contributing propagators by an odd number of interchanges.



Finally there is a combinatoric factor arising from a sometimes partial cancellation of the $1/n!$ multiplying the nth order term.

<u>Connected Diagrams</u> Terms in the perturbation series described by disconnected diagrams factorize into a product of the values of each connected subdiagram. Thus all the useful information is contained in the subset of connected diagrams, and it is useful to know how the final answer is expressed in terms of connected diagrams only. Roughly speaking, the sum of all diagrams is simply the exponential of the sum

of all connected diagrams. This statement applies to the expansion of

$$\langle 0| Te^{-i\int dt H_I'(t)} |0\rangle \tag{12.1}$$

for a completely general Hamiltonian. The reason is purely combinatoric. Call the value of the sum of connected graphs at order k, $G_k^c/k!$. Consider the terms at order $n = \sum_{k=1}^{\infty} kr_k$ in the perturbation series described by r_k appearances of the connected graphs of order k. (Note that all but a finite number of the r_k are 0.) The contractions leading to these terms can occur in $n!/\prod_k[(k!)^{r_k} r_k!]$ ways. This is because there are $n!$ ways to assign the n hamiltonians to the n vertices of all the connected subgraphs, but this overcounts the number of contractions by a factor of $k!$ for each of the connected subgraphs of order k, because different orders of the assignment to each subgraph do not give distinct contractions, and overcounts by a factor of $r_k!$ for each group of identical subgraphs for the same reason. Thus the value of the order n terms is

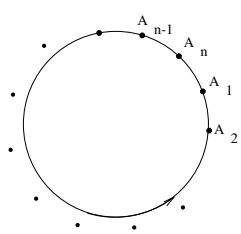
$$\frac{1}{n!}\sum_{\sum kr_k=n} \frac{n!}{\prod_k[(k!)^{r_k} r_k!]} \prod_k (G_k^c)^{r_k} = \sum_{\sum kr_k=n} \prod_k \left[\frac{1}{r_k!}(\frac{G_k^c}{k!})^{r_k}\right].$$

Note that the factors in square brackets are 1 if all $r_k = 0$. Summing over all n simply relaxes the constraint on the summation over the r_k, so we have finally

$$\langle 0| Te^{-i\int dt H_I'(t)} |0\rangle = \prod_{k=1}^{\infty}\left[\sum_{r_k=0}^{\infty} \frac{1}{r_k!}(\frac{G_k^c}{k!})^{r_k}\right] = \prod_{k=1}^{\infty} e^{G_k^c/k!} = e^{\sum_k G_k^c/k!},$$

which is the desired result.

We now turn to the connected diagrams for an external electromagnetic field. At order n in the fields, the diagrams contributing to $\ln\langle out|in\rangle$ are



$$= -(iq)^n \int d^4x_1 \cdots d^4x_n$$
$$\mathrm{Tr}[\gamma\cdot A(x_1)S_F(x_1-x_2)\cdots\gamma\cdot A(x_n)S_F(x_n-x_1)]. \tag{12.2}$$

Since there are $(n-1)!$ distinct diagrams with the same value (after the coordinate integrations) the net combinatoric factor is $(n-1)!/n! = 1/n$. This factor can be interpreted as $1/S_\Gamma$ where S_Γ is the symmetry number of the graph γ. It is also worth noting that the same formula applies to the case of nonabelian gauge field $A_{\mu ab}$ with the understanding that the trace includes the trace over the internal indices as well as the spinor indices.

In fact, it is instructive to regard the coordinates as (continuous) indices, so $S_{F x_1 a, x_2 b}$ is a matrix and $(\gamma\cdot A)_{x_1 x_2} \equiv \delta(x_1-x_2)\gamma\cdot A$ is a matrix, so the term as a whole can be regarded as a grand trace

$$-\frac{1}{n}\mathbf{Tr}[iq\gamma\cdot AS_F]^n$$

and the sum over all n is then recognized as the Taylor expansion for a logarithm:

$$\ln\langle out|in\rangle_A = \mathbf{Tr}[\ln(I - iq\gamma\cdot AS_F)].$$

Making use of the identities $\det A = \exp\{Tr\ln A\}$ and $\det AB = \det A \det B$, which are fundamental prop-

erties of the determinant, and noting that in this matrix notation $S_F = -i(m + (1/i)\gamma\cdot\partial)^{-1}$, we see that

$$\begin{aligned}\langle out|in\rangle_A =& \det(I - q\gamma\cdot A(m + \frac{1}{i}\gamma\cdot\partial)^{-1}) \\ =& \frac{\det(m + \frac{1}{i}\gamma\cdot\partial - q\gamma\cdot A)}{\det(m + \frac{1}{i}\gamma\cdot\partial)} \\ \equiv& \frac{\det(m - i\gamma\cdot D)}{\det(m - i\gamma\cdot\partial)}\end{aligned} \tag{12.3}$$

where we have defined the covariant derivative operator as $D = \partial - iqA$. The denominator in (12.3) serves to normalize $\langle out|in\rangle$ to 1 at $A = 0$, which amounts to choosing the zero of energy to be the ground state energy of H_0. In practice these "explicit formulae" for $\langle out|in\rangle$ can not be evaluated exactly for general A, although for special external potentials such as those corresponding to constant field strength it is possible. Nonetheless they give the most efficient derivation of the perturbation series in powers of the external fields, and as we shall see give some general insight into the meaning of the amplitudes we are calculating.

Furry's Theorem. There appears to be a connected diagram (12.2) for every n. But for the electromagnetic case, only those with even n are nonvanishing. To see why this is true use the trace property $\mathrm{Tr}\, A^T = \mathrm{Tr}\, A$ to show that

$$\begin{aligned}&\mathrm{Tr}[\gamma\cdot A(x_1)S_F(x_1 - x_2)\cdots\gamma\cdot A(x_n)S_F(x_n - x_1)] = \\ &\qquad \mathrm{Tr}[S_F^T(x_n - x_1)\gamma^T\cdot A(x_n)\cdots S_F^T(x_1 - x_2)\gamma^T\cdot A(x_n)].\end{aligned}$$

But $\gamma^{\mu T} = -(\gamma^0\gamma^2)^{-1}\gamma^\mu\gamma^0\gamma^2$. From this it follows that

$$S_F^T(x) = -i(\gamma^0\gamma^2)^{-1}\int\frac{d^4p}{(2\pi)^4}\frac{m + \gamma\cdot p}{m^2 + p^2}e^{ip\cdot x}\gamma^0\gamma^2 = (\gamma^0\gamma^2)^{-1}S_F(-x)\gamma^0\gamma^2$$

So we have

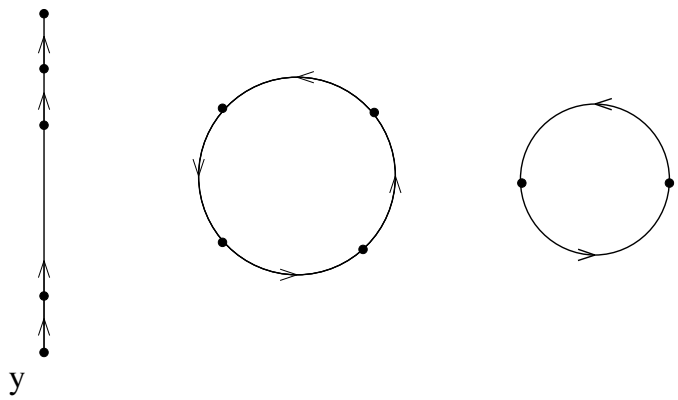
$$\begin{aligned}&\mathrm{Tr}[\gamma\cdot A(x_1)S_F(x_1 - x_2)\cdots\gamma\cdot A(x_n)S_F(x_n - x_1)] = \\ &\qquad (-)^n\,\mathrm{Tr}[S_F(x_1 - x_n)\gamma\cdot A(x_n)\cdots S_F(x_2 - x_1)\gamma\cdot A(x_1)].\end{aligned}$$

After integrating over the x's the only difference between the left and right sides is the labelling of dummy integration variables and the factor $(-)^n$ on the right. Thus for odd n both sides must vanish, *i.e.* all connected diagrams for $\langle out|in\rangle$ with an odd number of A's vanish. This is Furry's theorem. A more basic way to understand the result is to note that the substitution $A \to -A$ can be undone by the charge conjugation transformation under which $j \to -j$. Thus $\langle out|in\rangle$ must be an even function of A.

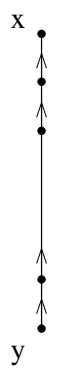
The perturbation series can be similarly "summed" for the propagator in the presence of external gauge fields

$$\begin{aligned}\langle out|\, T\psi(x)\bar\psi(y)\,|in\rangle_A =& \sum_{n=0}^{\infty}\frac{i^n}{n!}\int d^4x_1\cdots d^4x_n \\ &\langle 0|\, T[A\cdot j(x_1)\cdots A\cdot j(x_n)\psi(x)\bar\psi(y)]\,|0\rangle\,.\end{aligned}$$

The connected subdiagrams are of two types: (1) In a given term there are any number of closed loop diagrams of the sort contributing to $\langle out|in\rangle_A$; (2) In each term there is exactly one subdiagram with a continuous line running from y to x



Summing up all of the first type of diagram gives simply an overall factor of $\langle out|in\rangle_A$ multiplying the value of each diagram of the second type. Apart from the net combinatoric factor the latter diagram has at order n the value



$$= -(iq)^n \int d^4x_1 \cdots d^4x_n$$
$$S_F(x-x_1)\gamma\cdot A(x_1)S_F(x_1-x_2)\cdots\gamma\cdot A(x_n)S_F(x_n-y). \tag{12.4}$$

In fact the combinatoric factor is just 1 since there are $n!$ distinct contractions leading to this diagram (giving identical values since they only differ in the labelling of dummy integration variables), and that precisely cancels the $1/n!$ coming from the Dyson formula. In the matrix notation introduced above, we recognize the sum of all diagrams, including the disconnected closed loops, as

$$\begin{aligned}\langle out|in\rangle_A \sum_{n=0}^{\infty} S_F(iq\gamma\cdot AS_F)^n &= \langle out|in\rangle_A\, S_F \frac{1}{I - iq\gamma\cdot AS_F}\\ &= \langle out|in\rangle_A \frac{-i}{m - i\epsilon - i\gamma\cdot D}.\end{aligned}$$

In other words the right hand side is just proportional to a Green function for the differential operator $m - i\gamma\cdot D$:

$$\langle out|\, T\psi(x)\bar{\psi}(y)\,|in\rangle_A = \langle out|in\rangle_A\, S(x,y;A)$$

where

$$(m - i\gamma\cdot D)S(x,y;A) = -i\delta^4(x-y),$$

and the boundary condition is specified by giving an infinitesimal negative imaginary part to m. Since $A=0$ in the distant past and future, this prescription corresponds to the requirement of only positive frequencies as $t\to+\infty$ and negative frequencies as $t\to-\infty$, which are manifest properties of $\langle out|\, T\psi(x)\bar{\psi}(y)\,|in\rangle_A$. The fact that this latter quantity is this Green function is also immediately seen by applying the differential operator and remembering the contribution from differentiating the step functions implicit in the time ordering symbol.

By making use of our result for $\langle out| T\psi(x)\bar{\psi}(y)\,|in\rangle_A$, we can give a much quicker derivation of the determinant formula for $\langle out|in\rangle_A$. Looking back to Eq(11.4) we see that if we make a change δA in A, the change in $\langle out|in\rangle$ is

$$\delta\,\langle out|in\rangle_A = \langle in|\,\delta U(\infty,-\infty)\,|in\rangle = i\int d^4x\delta A_\mu(x)\,\langle out|\,j^\mu(x)\,|in\rangle_A\,.$$

On the other hand the current matrix element can be related to the Green function of the previous paragraph

$$\begin{aligned}\langle out|\,j^\mu(x)\,|in\rangle_A &= \langle out|\,\frac{q}{2}[\bar{\psi}(x),\gamma^\mu\psi(x)]\,|in\rangle_A\\ &= -\frac{q}{2}\,\mathrm{Tr}\,\gamma^\mu\,\langle out|\,[\psi(x),\bar{\psi}(x)]\,|in\rangle_A\\ &= -\,q\,\mathrm{Tr}\,\gamma^\mu\,\langle out|\,T[\psi(x)\bar{\psi}(x)]\,|in\rangle_A\\ &= -\,q\,\mathrm{Tr}\,\gamma^\mu S(x,x;A)\,\langle out|in\rangle_A\,.\end{aligned}$$

where we are interpreting $\theta(0)=1/2$. Using $\delta\,\langle out|in\rangle\,/\,\langle out|in\rangle = \delta\ln\langle out|in\rangle$ we obtain

$$\begin{aligned}\delta\ln\langle out|in\rangle_A =& -iq\int d^4x\,\mathrm{Tr}[\gamma\cdot\delta A S(x,x;A)]\\ =&\mathbf{Tr}[-q\gamma\cdot\delta A(m-i\gamma\cdot D)^{-1}]\\ =&\delta\mathbf{Tr}\ln[m-i\gamma\cdot D]\end{aligned}$$

which implies our previous formula up to a multiplicative A independent constant which is fixed by requiring $\langle out|in\rangle_{A=0}=1$. Actually this initial condition is somewhat artificial; it would be more natural to simply take $\langle out|in\rangle_A = \det(m-i\gamma\cdot D)$. This corresponds to not removing the sea contribution to the energy in the case $A=0$.

It is instructive to see how the formal expression for the sea energy comes out of this evaluation. When $A=0$ the matrix element (in momentum basis) $(p'|\ln(m-\gamma\cdot\partial)|p) = \delta(p'-p)\ln(m+\gamma\cdot p)$, so when we take the trace over the continuous momentum indices, we set $p'=p$ and get an overall factor of $\delta^4(0)$. To interpret this singular factor, think of the integral representation for $\delta(p)=\int e^{ip\cdot x}d^4x/(2\pi)^4$. As $p\to 0$ this is just $VT/(2\pi)^4$ where V is the volume of space and T the duration of time. With this interpretation, we have

$$\begin{aligned}\mathbf{Tr}\ln(m-i\gamma\cdot\partial) =& \frac{VT}{(2\pi)^4}\int d^4p\,\mathrm{Tr}\ln(m+\gamma\cdot p)\\ =&\frac{VT}{(2\pi)^4}\int d^4p\;2\ln(m^2+p^2)\end{aligned}$$

where we have used the fact that the eigenvalues of $\gamma\cdot p$ are $+\sqrt{-p^2}$ twice and $-\sqrt{-p^2}$ twice. To compare this to the sea energy we obtained earlier, we need to interpret the integral over p^0. This is of course a divergent integral, but if we differentiate once with respect to m^2 to get the mass dependence, we get a convergent integral which can be evaluated by closing the contour in the upper half complex p^0 plane

$$\int\frac{dp^0}{2\pi}\frac{1}{\omega(\mathbf{p})^2-p^{02}-i\epsilon} = \frac{i}{2\omega} = i\frac{d}{dm^2}\omega(\mathbf{p}).$$

Thus

$$\ln\langle out|in\rangle = 2iT\frac{V}{(2\pi)^3}\int d^3p[\omega(\mathbf{p})+C]$$

where the constant is at least independent of the mass. Since $\langle out|in\rangle\sim e^{-iE_{sea}T}$ we see that we recover our previous result for E_{sea} at least as far as the mass dependence is concerned. The constant C is itself infinite and complex: the Wick rotation to imaginary p^0 gives a factor of i, but the contours at infinty are nonvanishing and complex.

The occurrence of complex energies in the Minkowski definition of $\langle out|in\rangle$ motivates the idea that these vacuum amplitudes should be defined fundamentally in Euclidean space and then continued back to Minkowski space as the application demands. In Euclidean space-time we put $T = -iT_E$ and the vacuum amplitude would have the behavior $\langle out|in\rangle \sim \exp(-E_{sea}T_E)$. If we repeat the calculation of $\mathbf{Tr}\ln(m - i\gamma\cdot\partial)$ in Euclidean space we get $VT_E \int d^4p_E 2\ln(m^2+p^2)/(2\pi)^4$ which is manifestly real albeit divergent, and we get the formula

$$E_{sea} = -2\frac{V}{(2\pi)^4}\int d^4p_E \ln(m^2+p^2).$$

This still has the mass dependence of our earlier result and the mass indepentent discrepancy is at least real. In effect, working in Euclidean space-time from the beginning discards undesirable complex contours from Wick rotations that fail to vanish only because of the poor high momentum (ultra-violet) behavior of the integrands of vacuum diagrams. Consequently, among modern field theorists Euclidean space-time is widely accepted as the best way to define vacuum amplitudes.

Finally, we consider the calculation of time ordered products of an arbitrary number of Dirac fields. Clearly the nonvanishing ones have an equal number of ψ's and $\bar{\psi}$'s:

$$\langle out|\, T[\psi(x_1)\cdots\psi(x_n)\bar{\psi}(y_n)\cdots\bar{\psi}(y_1)]\,|in\rangle\,.$$

To each term there will correspond any number of connected closed loop diagrams and n connected diagrams of exactly the type contributing to the $n = 2$ case. Each of these latter subdiagrams consists of a line from one of the y_k to one of the x_{P_k}. The closed loops sum to an overall factor of $\langle out|in\rangle$. For each distinct pairing of the x's with the y's the other subdiagrams sum to a product of n factors $S(x_{P_k}, y_k; A)$. In other words, there is a Wick expansion for time ordered products of Dirac fields in the presence of an external field exactly as in the $A = 0$ case. All one does is substitute $S_F(x-y) \to S(x,y;A)$ and multiply by an overall factor of $\langle out|in\rangle_A$.

13. Scattering in External Fields

Our time dependent formalism is ideally suited for defining transition amplitudes. Since external fields are turned off in the far future and distant past, particle states can be defined in exactly the same manner as the *in* and *out* vacuua. We can expand the Heisenberg field operators at early(late) times in terms of annihilation and creation operators $b^{in}_\lambda(\mathbf{p}), d^{in}_\lambda(\mathbf{p})$ $(b^{out}_\lambda(\mathbf{p}), d^{out}_\lambda(\mathbf{p}))$, since they are free fields there. These are of course eigenoperators for $H_0(-\infty)(H_0(+\infty))$ respectively. They are of course determined up to a phase by the labels $\lambda, \mathbf{p}$. We can fix the phases of the *out* operators in terms of the *in* operators by defining $\Omega^{out} \equiv U_I^{-1}(\infty,-\infty)\Omega^{in}U_I(\infty,-\infty)^\star$. Then incoming particle states are obtained by applying $b^{in\dagger}, d^{in\dagger}$ to $|in\rangle$ and outgoing particle states are obtained by applying b^{out}, d^{out} to $\langle out|$.

The transition amplitudes between multiparticle states at early times and multiparticle states at late times can be immediately transcribed to interaction picture:

$$\langle out|\, b_1^{out}\cdots d_N^{out} d_M^{in\dagger}\cdots b_1^{in\dagger}\,|in\rangle = \langle 0,I|\, b_{I1}\cdots d_{IN}U_I(\infty,-\infty)d^\dagger_{IM}\cdots b^\dagger_{I1}\,|0,I\rangle\,.$$

The rules for expanding these amplitudes in perturbation theory are very similar to those for the outin matrix elements of time ordered products of Dirac field operators. The disconnected closed loops sum up to an overall factor of $\langle out|in\rangle_A$. The creation and annihilation operators can either contract against each other or against one of the Dirac fields in U_I:

$$\langle 0,I|\, b_\lambda(\mathbf{p})\bar\psi(x)\,|0,I\rangle = \frac{1}{(2\pi)^{3/2}\sqrt{2\omega}}\bar u_\lambda(\mathbf{p})e^{-ip\cdot x}$$

$$\langle 0,I|\, d_\lambda(\mathbf{p})\psi(x)\,|0,I\rangle = \frac{1}{(2\pi)^{3/2}\sqrt{2\omega}}v_\lambda(\mathbf{p})e^{-ip\cdot x}$$

$$\langle 0,I|\, \psi(x)b^\dagger_\lambda(\mathbf{p})\,|0,I\rangle = \frac{1}{(2\pi)^{3/2}\sqrt{2\omega}}u_\lambda(\mathbf{p})e^{ip\cdot x}$$

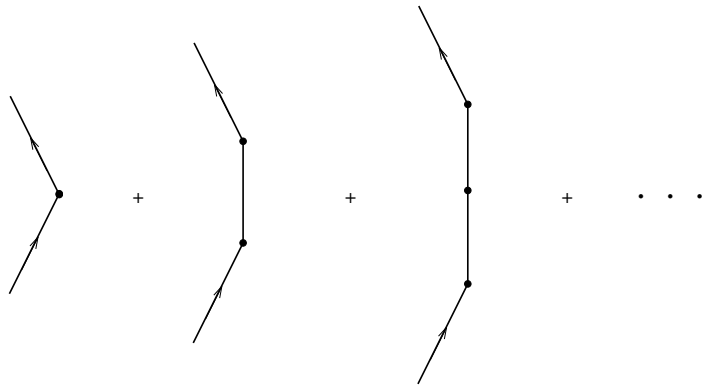
$$\langle 0,I|\, \bar\psi(x)d^\dagger_\lambda(\mathbf{p})\,|0,I\rangle = \frac{1}{(2\pi)^{3/2}\sqrt{2\omega}}\bar v_\lambda(\mathbf{p})e^{ip\cdot x}$$

To illustrate how to use these rules consider the process in which a single particle at early time is scattered by the external potential to a single particle at late times.

$$\langle out|\, b^{out}_\lambda(\mathbf{q})b^{in\dagger}_\lambda(\mathbf{p})\,|in\rangle = \langle out|in\rangle_A\,[\delta_{\lambda'\lambda}\delta^3(\mathbf{q}-\mathbf{p}) + \langle 0,I|\, b_\lambda(\mathbf{q})U_I b^\dagger_\lambda(\mathbf{p})\,|0,I\rangle^c]$$

where the superscript c denotes the restriction to connected diagrams containing at least one vertex.

$\star$ Note that since $U = U_0U_I$ and $U_0^{-1}\Omega^{in}U_0$ is just a numerical phase times Ω^{in} this definition does ensure that $\Omega^{out}$is an eigenoperator of $H_0(+\infty)$.



The first diagram has the value

$$\frac{1}{(2\pi)^3\sqrt{4\omega(\mathbf{q})\omega(\mathbf{p})}}\int d^4x e^{i(p-q)\cdot x}\bar{u}_{\lambda'}(\mathbf{q})iq\gamma\cdot A(x)u_\lambda(\mathbf{p}).$$

The second diagram has two vertices and a factor of $S_F(y-x)$ between them. But notice that the second plus all the higher diagrams just amounts to replacing this S_F by $S_F(y,x;A)$, so the sum of all the diagrams but the first has the value

$$\frac{1}{(2\pi)^3\sqrt{4\omega(\mathbf{q})\omega(\mathbf{p})}}\int d^4x d^4y e^{i(p\cdot x-q\cdot y)}\bar{u}_{\lambda'}(\mathbf{q})iq\gamma\cdot A(y)S_F(y,x;A)iq\gamma\cdot A(x)u_\lambda(\mathbf{p}).$$

We put all this together in the form

$$\langle out|\, b_\lambda^{out}(\mathbf{q})b_\lambda^{in\dagger}(\mathbf{p})\,|in\rangle = \langle out|in\rangle_A\,[\delta_{\lambda'\lambda}\delta^3(\mathbf{q}-\mathbf{p})+\mathcal{M}(\mathbf{q},\mathbf{p};A)],$$

where

$$\begin{aligned}\mathcal{M}(\mathbf{q},\mathbf{p};A) =& \frac{1}{(2\pi)^3\sqrt{4\omega(\mathbf{q})\omega(\mathbf{p})}}\int d^4x d^4y e^{i(p\cdot x-q\cdot y)}\\ &\bar{u}_{\lambda'}(\mathbf{q})[iq\gamma\cdot A(x)\delta(x-y)+iq\gamma\cdot A(y)S_F(y,x;A)iq\gamma\cdot A(x)]u_\lambda(\mathbf{p}).\end{aligned} \tag{13.1}$$

We can put this last formula in a more suggestive form by defining the free Dirac plane wave functions

$$\psi^0_{\mathbf{p}\lambda}(x) \equiv \frac{1}{(2\pi)^{3/2}\sqrt{2\omega}}u_\lambda(\mathbf{p})e^{ip\cdot x}$$

and noting that

$$\psi_{\mathbf{p}\lambda}(x) \equiv \psi^0_{\mathbf{p}\lambda}(x) + \int d^4y S_F(x,y;A)iq\gamma\cdot A(y)\psi^0_{\mathbf{p}\lambda}(y)$$

is a solution of the Dirac equation in the presence of A:

$$(m+\frac{1}{i}\gamma\cdot D)\psi_{\mathbf{p}\lambda}=0,$$

with the boundary condition that at early times the only positive frequency components are contained in the term $\psi^0_{\mathbf{p}\lambda}(x)$. Thus we can write

$$\mathcal{M}(\mathbf{q},\mathbf{p};A) = \int d^4x\bar{\psi}^0_{\mathbf{q}\lambda'}(x)iq\gamma\cdot A(x)\psi_{\mathbf{p}\lambda}(x),$$

which is reminiscent of the corresponding formula in nonrelativistic quantum mechanics for the scattering of a particle from an external potential.

In the case of a static potential, the time dependence of both $\psi^0_{\mathbf{p}\lambda}(x)$ and $\psi_{\mathbf{p}\lambda}(x)$ is given by a multiplicative phase $e^{-i\omega(\mathbf{p})t}$ and the time integral then provides an energy conserving delta function

$$\mathcal{M}(\mathbf{q},\mathbf{p};A(\mathbf{x})) = 2\pi\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))\int d^3x\bar{\psi}^0_{\mathbf{q}\lambda'}(\mathbf{x})iq\gamma\cdot A(\mathbf{x})\psi_{\mathbf{p}\lambda}(\mathbf{x}).$$

Recall from basic scattering theory that if the scattering matrix for a particle from a static potential is written

$$\langle\mathbf{q},out|\mathbf{p},in\rangle = \delta_{\lambda'\lambda}\delta(\mathbf{q}-\mathbf{p}) - 2\pi i\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})$$

then the differential scattering cross section is given by

$$\begin{aligned}\frac{d\sigma}{d\Omega} =& \frac{d^3q}{d\Omega}\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))\frac{(2\pi)^4}{v}|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\\ =& q^2dq\delta(\omega(\mathbf{q})-\omega(\mathbf{p}))\frac{(2\pi)^4}{v}|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\\ =& p\omega(\mathbf{p})\frac{(2\pi)^4}{v}|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\\ =& \omega(\mathbf{p})^2(2\pi)^4|T_{\lambda'\lambda}(\mathbf{q},\mathbf{p})|^2,\end{aligned}$$

where v is the speed of the incident particle. Comparing with our expression we find that for electron scattering from a static potential

$$T_{\lambda'\lambda}(\mathbf{q},\mathbf{p}) = i\int d^3x\bar{\psi}^0_{\mathbf{q}\lambda'}(\mathbf{x})iq\gamma\cdot A(\mathbf{x})\psi_{\mathbf{p}\lambda}(\mathbf{x}).$$

Note that in the absence of pair production, the factor $\langle out|in\rangle$, the vacuum persistence amplitude, is a pure phase and doesn't contribute in the absolute square of T. In lowest order in A (the Born approximation) one simply replaces ψ by ψ^0.

Let us note some tricks that are useful in calculating cross sections for processes with unobserved final spins and unpolarized beams. When we calculate the absolute square of a spinor matrix element we can make use of the identities

$$\begin{aligned}\sum_\lambda u_\lambda\bar{u}_\lambda &= m-\gamma\cdot p\\ \sum_\lambda v_\lambda\bar{v}_\lambda &= -m-\gamma\cdot p.\end{aligned}$$

For example these allow us to write

$$\begin{aligned}\sum_{\lambda\lambda'}|\bar{u}_{\lambda'}\gamma\cdot\tilde{A}(\mathbf{q}-\mathbf{p})u_\lambda(\mathbf{p})|^2 =& \mathrm{Tr}[\gamma\cdot\tilde{A}(m-\gamma\cdot p)\gamma\cdot\tilde{A}^*(m-\gamma\cdot q)]\\ =& 4(-m^2-p\cdot q)\tilde{A}\cdot\tilde{A}^* + 4(p\cdot\tilde{A}q\cdot\tilde{A}^* + q\cdot\tilde{A}p\cdot\tilde{A}^*)\\ =& 2(p-q)^2\tilde{A}\cdot\tilde{A}^* + 4(p\cdot\tilde{A}q\cdot\tilde{A}^* + q\cdot\tilde{A}p\cdot\tilde{A}^*)\end{aligned}$$

Physical quantities should be gauge invariant and cross sections are no exception. We can reveal the gauge invariance of this last formula by introducing $\tilde{F}_{\mu\nu}(q)\equiv i(q_\mu\tilde{A}_\nu - q_\nu\tilde{A}_\mu)$ which is the Fourier transform of the gauge invariant field strength. Then

$$\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu *} = 2(q^2\tilde{A}\cdot\tilde{A}^* - q\cdot\tilde{A}q\cdot\tilde{A}^*).$$

Recalling that the argument of $\tilde{A}$ is $q-p$ we then find that the squared spinor matrix element can be

rewritten

$$\sum_{\lambda\lambda'} |\bar{u}_{\lambda'}\gamma\cdot\tilde{A}(\mathbf{q}-\mathbf{p})u_\lambda(\mathbf{p})|^2 = \tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu *} + 2(p+q)\cdot\tilde{A}\ (p+q)\cdot\tilde{A}^*.$$

The first term is manifestly gauge invariant and the second is gauge invariant by virtue of the identity

$$(q+p)\cdot(q-p) = q^2 - p^2 = -m^2 + m^2 = 0.$$

The differential cross section for unpolarized electron scattering in a general static potential with final spin unobserved is then

$$\begin{aligned}\sum_{\lambda'} \frac{d\sigma}{d\Omega}^{Born}_{Unpolarized} =& \frac{1}{16\pi^2}\frac{1}{2}\sum_{\lambda'\lambda} |\bar{u}_{\lambda'}q\gamma\cdot\tilde{A}(\mathbf{q}-\mathbf{p})u_\lambda(\mathbf{p})|^2 \\ =& \frac{e^2}{16\pi^2}(\frac{1}{2}\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu *} + (p+q)\cdot\tilde{A}\ (p+q)\cdot\tilde{A}^*) \\ =& \frac{\alpha}{4\pi}(\frac{1}{2}\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu *} + |(p+q)\cdot\tilde{A}|^2).\end{aligned}$$

One can easily confirm that this agrees with our previous results for the special case of the Coulomb potential.

We expect that the transition amplitudes for antiparticles should be obtained from (13.1) by the substitution $A \to -A$. On the other hand the rules seem to give a different prescription:

$$\langle out|\, d^{out}_\lambda(\mathbf{q}) d^{in\dagger}_\lambda(\mathbf{p})\, |in\rangle = \langle out|in\rangle_A\, [\delta_{\lambda'\lambda}\delta^3(\mathbf{q}-\mathbf{p}) + \bar{\mathcal{M}}(\mathbf{q},\mathbf{p};A)],$$

with

$$\begin{aligned}\bar{\mathcal{M}}(\mathbf{q},\mathbf{p};A) = -\frac{1}{(2\pi)^3\sqrt{4\omega(\mathbf{q})\omega(\mathbf{p})}}\int d^4x d^4y e^{i(p\cdot x - q\cdot y)} \\ \bar{v}_\lambda(\mathbf{p})[iq\gamma\cdot A(x)\delta(x-y) + iq\gamma\cdot A(x)S_F(x,y;A)iq\gamma\cdot A(y)]v_{\lambda'}(\mathbf{q}).\end{aligned} \tag{13.2}$$

In fact it is not hard to show that $\bar{\mathcal{M}}(\mathbf{q},\mathbf{p};A) = \mathcal{M}(\mathbf{q},\mathbf{p};-A)$ by inserting $v = i\gamma^2 u^*$ into (13.2) and transposing the matrix element, using $i\gamma^2\gamma^0\gamma^{\mu T}\gamma^0 i\gamma^2 = -\gamma^\mu$ and $i\gamma^2\gamma^0 S^T(x,y;A)\gamma^0 i\gamma^2 = -\gamma^\mu = S_F(y,x;-A)$. The latter fact can be seen either term by term in the expansion in A or by examining the defining Green function equation for $S_F(x,y;A)$.

In addition to the scattering of electrons or positrons, an external field can induce electron positron pair production and also pair annihilation. Taking the first case, for example, we have

$$\langle out|\, b^{out}_\lambda(\mathbf{q}_1) d^{out}_{\lambda'}(\mathbf{q}_2)\, |in\rangle = \langle out|in\rangle_A\, [\mathcal{M}^{PairCreate}(\mathbf{q}_1,\mathbf{q}_2;A)],$$

with

$$\begin{aligned}\mathcal{M}^{PairCreate}(\mathbf{q}_1,\mathbf{q}_2;A) = -\frac{1}{(2\pi)^3\sqrt{4\omega(\mathbf{q}_1)\omega(\mathbf{q}_2)}}\int d^4x d^4y e^{-i(q_1\cdot x + q_2\cdot y)} \\ \bar{u}_\lambda(\mathbf{q}_1)[iq\gamma\cdot A(x)\delta(x-y) + iq\gamma\cdot A(x)S_F(x,y;A)iq\gamma\cdot A(y)]v_{\lambda'}(\mathbf{q}_2).\end{aligned} \tag{13.3}$$

In this case, of course, there is no delta function term representing only vacuum persistence. Similarly for pair annihilation, we have

$$\langle out|\, b^{in\dagger}_\lambda(\mathbf{q}_1) d^{in\dagger}_{\lambda'}(\mathbf{q}_2)\, |in\rangle = \langle out|in\rangle_A\, [\mathcal{M}^{PairAnnih}(\mathbf{q}_1,\mathbf{q}_2;A)],$$

with

$$\begin{aligned}\mathcal{M}^{PairAnnih}(\mathbf{q}_1,\mathbf{q}_2;A) =& \frac{1}{(2\pi)^3\sqrt{4\omega(\mathbf{q}_1)\omega(\mathbf{q}_2)}}\int d^4x d^4y e^{-i(q_2\cdot x + q_1\cdot y)} \\ & \bar{v}_\lambda(\mathbf{q}_2)[iq\gamma\cdot A(x)\delta(x-y) + iq\gamma\cdot A(x)S_F(x,y;A)iq\gamma\cdot A(y)]u_{\lambda'}(\mathbf{q}_1).\end{aligned} \tag{13.3}$$

Notice the prominent appearance of the Green function $S_F(x,y;A)$ for the Dirac Equation with an

external field in all of the four basic processes, electron scattering, positron scattering, pair production, and pair annihilation. Moreover, we have also seen how to express the vacuum persistence amplitude in terms of this same Green function. Thus we see that the solutions of the first quantized Dirac equation are of direct utility in finding the physical properties of the second quantized theory.

As a final note we show how S_F even contains information about the bound states of an electron in a static external field. Actually it is better to deal not with an exactly static field, but with a field that is adiabatically switched on at some very early time $-T$, static for a very long time interval $\approx 2T$ and then adiabatically switched off at a very late time T. Then according to the adiabatic theorem, $|in\rangle$ is proportional to the ground state of the Hamiltonian $H(t)$ at all times, and $|out\rangle = |in\rangle \langle out|in\rangle$. In particular, $|in\rangle$ is the ground state $|G\rangle$ of H_A the hamiltonian for the static potential under study, since $H(t) = H_A$ for $-T << t << T$. Now fix the times $T >> x^0 > y^0 >> -T$. Then we have

$$\begin{aligned}
S_F(x,y)\langle out|in\rangle &= \langle out|\,\psi(x)\bar{\psi}(y)\,|in\rangle \\
&= \langle out|in\rangle\,\langle G|\,\psi(x)\bar{\psi}(y)\,|G\rangle \\
&= \langle out|in\rangle \sum_n \langle G|\,\psi(x)\,|n\rangle\,\langle n|\,\bar{\psi}(y)\,|G\rangle \\
&= \langle out|in\rangle \sum_n e^{-i(E_n-E_G)(x^0-y^0)}\,\langle G|\,\psi(\mathbf{x},0)\,|n\rangle\,\langle n|\,\bar{\psi}(\mathbf{y},0)\,|G\rangle
\end{aligned}$$

From this last formula, we see that a harmonic analysis of S_F in the time variable $t = x^0 - y^0$ yields the possible energy eigenvalues $E_n - E_G$ of the energy eigenstates $|n\rangle$. These states have the quantum numbers of a one electron state because they are created from the vacuum by $\bar{\psi}$. Note that the harmonic components for $x^0 < y^0$ have the interpretation as $-(E_{\bar{n}} - E_G)$ where $E_{\bar{n}}$ are the energy levels of one positron states. Because S_F satisfies the homogeneous Dirac equation for $t \neq 0$, the possible energy eigenvalues are solutions of the time independent Dirac equation

$$(m + \frac{1}{i}\boldsymbol{\gamma}\cdot\mathbf{D} + qA^0\gamma^0 - (E_n - E_G))\psi = 0, \qquad t > 0$$

or

$$(m + \frac{1}{i}\boldsymbol{\gamma}\cdot\mathbf{D} + qA^0\gamma^0 + (E_{\bar{n}} - E_G))\psi = 0, \qquad t < 0.$$

Thus, again, the solution of the energy eigenvalue problem for the first quantized theory is directly applicable to that for the second quantized theory.

<u>Relation to Time Ordered Products.</u> (This subsection can be skipped in a first reading.) It is useful to establish how the scattering amplitudes are related to matrix elements of time ordered products. First let us define

$$\begin{aligned}
b^\dagger_\lambda(\mathbf{p},t) &= \int d^3x \frac{1}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} e^{ip\cdot x}\bar{\psi}(x)\gamma^0 u_\lambda(\mathbf{p}) \\
d^\dagger_\lambda(\mathbf{p},t) &= \int d^3x \frac{1}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} e^{ip\cdot x}\bar{v}_\lambda(\mathbf{p})\gamma^0\psi(x) \\
b_\lambda(\mathbf{p},t) &= \int d^3x \frac{1}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} e^{-ip\cdot x}\bar{u}_\lambda(\mathbf{p})\gamma^0\psi(x) \\
d_\lambda(\mathbf{p},t) &= \int d^3x \frac{1}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} e^{-ip\cdot x}\bar{\psi}(x)\gamma^0 v_\lambda(\mathbf{p}).
\end{aligned} \tag{13.4}$$

For free fields all these operators are constant in time and are just the creation and annihilation operators of the Dirac field. For a Dirac field in the presence of external fields which vanish at early and late times they are *not* constant but approach the "out" creation and annihilation operators at $t = \infty$ and the "in" operators at $t = -\infty$.

Now we shall make use of the following "reduction" trick:

$$F(t=+\infty)T[\phi_1(x_1)\cdots\phi_N(x_N)] \mp T[\phi_1(x_1)\cdots\phi_N(x_N)]F(t=-\infty) = \int_{-\infty}^{\infty} dt \frac{\partial}{\partial t} T[F(t)\phi_1(x_1)\cdots\phi_N(x_N)]$$

where $F(t)$ is any function of Heisenberg operators at time t, and the ϕ's are generic field operators. The sign choice is $+$ for $F(t)$ *and* $T[\cdots]$ fermionic operators, but $-$ in all other cases. In particular, F can be any of the expressions (13.4) where the time appearing on the r.h.s. is set to t. This leads to a series of formulae for the commutation of creation and annihilation operators with time ordered products. For instance, take F to be the r.h.s. of the expression for $b_\lambda(\mathbf{p})$.

$$\begin{aligned}
&b_\lambda^{out}(\mathbf{p})T[\phi_1(x_1)\cdots\phi_N(x_N)] \mp T[\phi_1(x_1)\cdots\phi_N(x_N)]b_\lambda^{in}(\mathbf{p}) = \\
&\quad \frac{\bar{u}_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \int d^4x \gamma^0 \frac{\partial}{\partial t}(e^{-ip\cdot x}T[\psi(x)\phi_1(x_1)\cdots\phi_N(x_N)]) \\
&= \frac{\bar{u}_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \int d^4x e^{-ip\cdot x}\left(\gamma^0\frac{\partial}{\partial t} + i\gamma^0\omega(\mathbf{p})\right)T[\psi(x)\phi_1(x_1)\cdots\phi_N(x_N)] \\
&= \frac{\bar{u}_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \int d^4x e^{-ip\cdot x} i\,(m - i\gamma\cdot\partial)\,T[\psi(x)\phi_1(x_1)\cdots\phi_N(x_N)]
\end{aligned}$$

where in the last line we used

$$\bar{u}\gamma^0\omega e^{-ip\cdot x} = \bar{u}(m + i\boldsymbol{\gamma}\cdot\nabla)e^{-ip\cdot x}$$

and then integrated by parts. Clearly this derivation can be repeated for each of the expressions in (13.4) leading to the reduction formulae:

$$\begin{aligned}
&b_\lambda^{out}(\mathbf{p})T[\phi_1(x_1)\cdots\phi_N(x_N)] \mp T[\phi_1(x_1)\cdots\phi_N(x_N)]b_\lambda^{in}(\mathbf{p}) = \\
&\quad \frac{\bar{u}_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \int d^4x e^{-ip\cdot x} i\,(m - i\gamma\cdot\partial)\,T[\psi(x)\phi_1(x_1)\cdots\phi_N(x_N)] \\
&b_\lambda^{out\dagger}(\mathbf{p})T[\phi_1(x_1)\cdots\phi_N(x_N)] \mp T[\phi_1(x_1)\cdots\phi_N(x_N)]b_\lambda^{in\dagger}(\mathbf{p}) = \\
&\quad \int d^4x T[\bar{\psi}(x)\phi_1(x_1)\cdots\phi_N(x_N)] i\left(-m - i\gamma\cdot\overleftarrow{\partial}\right)e^{ip\cdot x}\frac{u_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \\
&d_\lambda^{out}(\mathbf{p})T[\phi_1(x_1)\cdots\phi_N(x_N)] \mp T[\phi_1(x_1)\cdots\phi_N(x_N)]d_\lambda^{in}(\mathbf{p}) = \\
&\quad \int d^4x T[\bar{\psi}(x)\phi_1(x_1)\cdots\phi_N(x_N)] i\left(-m - i\gamma\cdot\overleftarrow{\partial}\right)e^{-ip\cdot x}\frac{v_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \\
&d_\lambda^{out\dagger}(\mathbf{p})T[\phi_1(x_1)\cdots\phi_N(x_N)] \mp T[\phi_1(x_1)\cdots\phi_N(x_N)]d_\lambda^{in\dagger}(\mathbf{p}) = \\
&\quad \frac{\bar{v}_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \int d^4x e^{ip\cdot x} i\,(m - i\gamma\cdot\partial)\,T[\psi(x)\phi_1(x_1)\cdots\phi_N(x_N)]
\end{aligned} \tag{13.5}$$

The reduction formulae can be used to systematically reduce scattering amplitudes to expressions directly involving time ordered products.

We shall illustrate the procedure for the case of a particle scattering in an external field. The application

to antiparticle scattering and pair production and annihilation will be left to the reader:

$$\begin{aligned}
\langle out|\, b^{out}_{\lambda'}(\mathbf{q}) b^{in\dagger}_{\lambda}(\mathbf{p})\, |in\rangle &= \langle out|\, b^{in}_{\lambda'}(\mathbf{q}) b^{in\dagger}_{\lambda}(\mathbf{p})\, |in\rangle \\
&+ \frac{\bar{u}_{\lambda'}(\mathbf{q})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{q})}} \int d^4x e^{-iq\cdot x} i\,(m - i\gamma\cdot\partial)\,\langle out|\, \psi(x) b^{in\dagger}_{\lambda}(\mathbf{p})\, |in\rangle \\
&= \langle out|in\rangle\, \delta_{\lambda\lambda'}\delta(\mathbf{q}-\mathbf{p}) + \frac{\bar{u}_{\lambda'}(\mathbf{q})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{q})}} \int d^4x d^4y \\
e^{-iq\cdot x} i\,(m - i\gamma\cdot\partial)\,\langle out|\, T[\bar{\psi}(y)\psi(x)]\, |in\rangle\, i(-m - i\gamma\cdot\overleftarrow{\partial}_y) e^{ip\cdot y} &\frac{u_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}} \\
&= \langle out|in\rangle \left[\delta_{\lambda\lambda'}\delta(\mathbf{q}-\mathbf{p}) + \frac{\bar{u}_{\lambda'}(\mathbf{q})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{q})}} \int d^4x d^4y \right. \\
&\left. e^{-iq\cdot x} i\,(m - i\gamma\cdot\partial) S_F(x,y;A) i(m + i\gamma\cdot\overleftarrow{\partial}_y) e^{ip\cdot y} \frac{u_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}\right]
\end{aligned}$$

This is identical to our original expression as can be seen by writing $m - i\gamma\cdot\partial = m - i\gamma\cdot D + q\gamma\cdot A$ so that

$$i(m - i\gamma\cdot\partial) S_F(x,y;A) i(m + i\gamma\cdot\overleftarrow{\partial}_y) = \delta(x-y) i(m + i\gamma\cdot\overleftarrow{\partial}_y) + iq\gamma\cdot A S_F(x,y;A) i(m + i\gamma\cdot\overleftarrow{\partial}_y).$$

The first term on the r.h.s contributes a term $\bar{u}(q)(m + \gamma\cdot q)u(p) = 0$ by the Dirac equation. The second term can be simplified using the fact that S is also a Green function in its second argument:

$$S_F(m + i\gamma\cdot\overleftarrow{\partial}_y - q\gamma\cdot A) = -i\delta(x-y).$$

Thus we have

$$i(m - i\gamma\cdot\partial) S_F(x,y;A) i(m + i\gamma\cdot\overleftarrow{\partial}_y) = iq\gamma\cdot A\delta(x-y) + iq\gamma\cdot A S_F(x,y;A) iq\gamma\cdot A$$

as desired.

Now let us consider a little more closely the meaning of the reduction formula

$$\begin{aligned}
\langle out|\, b^{out}_{\lambda'}(\mathbf{q}) b^{in\dagger}_{\lambda}(\mathbf{p})\, |in\rangle = \langle out|in\rangle &\left[\delta_{\lambda\lambda'}\delta(\mathbf{q}-\mathbf{p}) + \frac{\bar{u}_{\lambda'}(\mathbf{q})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{q})}} \int d^4x d^4y \right. \\
&\left. e^{-iq\cdot x} i\,(m - i\gamma\cdot\partial) S_F(x,y;A) i(m + i\gamma\cdot\overleftarrow{\partial}_y) e^{ip\cdot y} \frac{u_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}\right].
\end{aligned} \tag{13.6}$$

The factors $(m - i\gamma\cdot\partial)$ and $i(m + i\gamma\cdot\overleftarrow{\partial}_y)$ look as though they should give zero because after integrating by parts, they become $(m + \gamma\cdot q)$ and $(m + \gamma\cdot p)$ respectively and these factors give zero next to the Dirac spinors. The error in this reasoning is of course that the surface terms at $t = \pm\infty$ are not zero (being in fact the scattering amplitudes themselves!). To get a clearer idea of what is happening, suppose we continue q^0, p^0 a little away from their "on-shell" values of $\omega(\mathbf{q}), \omega(\mathbf{p})$. Then the surface terms oscillate at infinity and are effectively zero (*e.g.*in any smooth wave packet). Then we can integrate by parts and the second term becomes

$$\begin{aligned}
&\frac{\bar{u}_{\lambda'}(\mathbf{q})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{q})}} i\,(m + \gamma\cdot q) \int d^4x d^4y e^{-iq\cdot x + ip\cdot y} S_F(x,y;A) i(m + \gamma\cdot p) \frac{u_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}. \\
&\equiv \frac{\bar{u}_{\lambda'}(\mathbf{q})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{q})}} i\,(m + \gamma\cdot q)\, T(q,p) i(m + \gamma\cdot p) \frac{u_\lambda(\mathbf{p})}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}.
\end{aligned} \tag{13.7}$$

where the r.h.s. defines $T(q,p)$, the Fourier transform of the Green function. The only way for this to be nonzero as q^0, p^0 approach there on-shell values is for T to acquire poles in this limit. The residues of these poles are then related to the scattering amplitudes. This is the content of the reduction formula, but there is a rather direct way to see how these poles come about.

Consider first of all the region of integration $x^0 > T$ where ψ is a free field, with $b^{out}, d^{out\dagger}$. Then this region contributes to $T(q,p)$ the bit

$$\int_T^\infty dt e^{i(q^0-\omega(\mathbf{q}))t} \frac{(2\pi)^{3/2}}{\sqrt{2\omega(\mathbf{q})}} \langle out| \sum_{\lambda'} b^{out}(\mathbf{q}) u_{\lambda'}(\mathbf{q}) \bar{\psi}(y) |in\rangle$$

For $q^0 = \omega$ the integrand is time independent so the integral over an infinite range gives a divergence. To study it, give q^0 a small positive imaginary part and do the integral:

$$\frac{ie^{i(q^0-\omega(\mathbf{q}))T}}{q^0-\omega(\mathbf{q})+i\epsilon} \frac{(2\pi)^{3/2}}{\sqrt{2\omega(\mathbf{q})}} \langle out| \sum_{\lambda'} b^{out}(\mathbf{q}) u_{\lambda'}(\mathbf{q}) \bar{\psi}(y) |in\rangle$$

For the range of integration with $x^0 < -T$ ψ contains $b^{in}, d^{in\dagger}$ and stands on the right so only the second operator contributes. The integral over this range is not singular for $q^0 \to \omega$.

Similar considerations apply to the integral over y^0. This time the region $y^0 < -T$ contributes a pole at $p^0 = \omega(\mathbf{p})$ involving $b^{in\dagger}$. There is also a pole from the region $y^0 > T$ involving $b^{out\dagger}$. This contribution doesn't vanish because of the presence of b^{out} from the first reduction; it just gives delta functions. Going through all these steps leads to

$$T(q,p) \underset{q^0,p^0 \to \omega(\mathbf{q}),\omega(\mathbf{p})}{\sim} \frac{i}{q^0-\omega(\mathbf{q})} \frac{i}{p^0-\omega(\mathbf{p})} \left[\frac{(2\pi)^3}{\sqrt{4\omega(\mathbf{q})\omega(\mathbf{p})}}\right] \sum_{\lambda',\lambda} u_{\lambda'}(\mathbf{q}) \bar{u}_\lambda(\mathbf{p})$$
$$\left[\langle out| b^{out}_{\lambda'}(\mathbf{q}) b^{in\dagger}_\lambda(\mathbf{p}) |in\rangle - \langle out|in\rangle \delta_{\lambda'\lambda} \delta(\mathbf{q}-\mathbf{p})\right]$$

which is of course exactly the behavior required to satisfy the reduction formula.

14. Vacuum Polarization

Before leaving external field problems, it is interesting to consider the effect an external field has on the vacuum (negative energy sea in the Dirac case). We shall calculate the response of the vacuum to the application of a weak electromagnetic field. For definiteness consider the Dirac field described by

$$H_A = H_0 - \int d^3x j^\mu(\mathbf{x},t) A_\mu(\mathbf{x},t).$$

Assume that $A \to 0$ as $t \to -\infty$ and that the system starts out in the ground state of $H_A(-\infty) = H_0(-\infty)$.

We should expect the field to induce charge and current densities in the vacuum. A simple measure of these induced currents is the expectation value of the Heisenberg picture current operator in the system state $\langle in| j^\mu(x) |in\rangle_A$. We may express this matrix element in interaction picture and then develop it in an expansion in powers of A. In the limit of very weak fields we can neglect all terms beyond those linear in A:

$$\begin{aligned}\langle in| j^\mu(x) |in\rangle &= \langle in| U_I^{-1}(t,-\infty) j_I^\mu(x) U_I(t,-\infty) |in\rangle_A \\ &= \langle 0,I| \left(Te^{i\int_{-\infty}^t j_I\cdot A}\right)^\dagger j_I^\mu(x) Te^{i\int_{-\infty}^t j_I\cdot A} |0,I\rangle \\ &\approx i\int d^4y\theta(t-t_y) \langle 0,I| [j_I^\mu(\mathbf{x},t), j_I^\nu(y)] |0,I\rangle A_\nu(y).\end{aligned}$$

There is no term independent of A because the vacuum expectation value of the current vanishes in the absence of applied fields.

This is the linear response to an applied field and is characterized by the response function

$$R^{\mu\nu}(x) \equiv i\theta(t) \langle 0| [j^\mu(x), j^\nu(0)] |0\rangle$$

where here and in the following we drop the subscripts I and it is understood that the currents are those of free fields. An important physical property of the response function is Einstein causality: $R^{\mu\nu}$ vanishes for spacelike argument $x^2 > 0$ as follows from the fact that local operators commute at space-like separations. Thus application of an external field at the origin at $t = 0$ cannot evoke a response at $\mathbf{x}$ until enough time has elapsed for light to travel from the origin to $\mathbf{x}$. This property is *not* shared by the time ordered product.

<u>Retarded Commutators from Time Ordered Products</u> The Wick expansion we have developed works best for time ordered products, so it is helpful that we can work out a relationship between the response function and the expectation value of the time-ordered product. This relationship is a general one that depends only on the time variable, so we suppress spatial and internal labels and consider two hermitian operators $O_1(t), O_2(t)$. We shall actually relate the Fourier transforms of the two quantities:

$$\begin{aligned}R(\omega) &\equiv i\int dte^{i\omega t}\theta(t) \langle G| [O_1(t), O_2(0)] |G\rangle \\ T(\omega) &\equiv i\int dte^{i\omega t} \langle G| T[O_1(t)O_2(0)] |G\rangle\end{aligned}$$

where $|G\rangle$ is the Ground State of the system, *assumed* to be nondegenerate. Now using $\theta(t) = 1 - \theta(-t)$ we

have

$$\theta(t)[O_1(t), O_2(0)] = T[O_1(t)O_2(0)] - O_2(0)O_1(t)$$

so the difference between R and T involves

$$\begin{aligned}
-i \int dt e^{i\omega t} \langle G| O_2(0)O_1(t) |G\rangle \\
&= -i \int dt e^{i\omega t} \sum_n \langle G| O_2(0) |n\rangle \langle n| O_1(0) |G\rangle e^{-i(E_G - E_n)t} \\
&= -2\pi i \delta(\omega) \langle G| O_2(0) |G\rangle \langle G| O_1(0) |G\rangle \\
&\quad -i \int dt e^{i\omega t} \sum_{n \neq G} \langle G| O_2(0) |n\rangle \langle n| O_1(0) |G\rangle e^{-i(E_G - E_n)t}
\end{aligned}$$

The important feature of this result is that by virtue of the fact that E_G is the lowest energy eigenvalue, the r.h.s. *vanishes* for positive frequency $\omega > 0$. Thus in this case $R(\omega) = T(\omega)$. Next we find a relation for negative frequency. For this case we relate R to the anti-time-ordered product:

$$\theta(t)[O_1(t), O_2(0)] = O_1(t)O_2(0) - \bar{T}[O_1(t)O_2(0)].$$

Now inserting a complete set of states allows us to conclude that the Fourier transform of the first term vanishes for negative frequency $\omega < 0$. Thus in this case we have $R(\omega) = -\bar{T}(\omega)$ where

$$\bar{T}(\omega) \equiv i \int dt e^{i\omega t} \langle G| \bar{T}[O_1(t)O_2(0)] |G\rangle = -T^*(-\omega).$$

where we used the assumption that O_1, O_2 are hermitian.

In summary we have found that

$$R(\omega) = \begin{cases} T(\omega) & \omega > 0 \\ T^*(-\omega) & \omega < 0. \end{cases}$$

For ω near zero we observe that the ground state contributes to T but not to R. The contribution to T is

$$T_G(\omega) = 2\pi i \delta(\omega) \langle G| O_2(0) |G\rangle \langle G| O_1(0) |G\rangle .$$

If there is a gap separating E_G from the rest of the spectrum this is the only zero frequency discrepancy between R and T.

Finally when we consider this relationship for field operators, it is natural to quote it for the spatial and temporal Fourier transform:

$$R(p^0, \mathbf{p}) = \begin{cases} T(p^0, \mathbf{p}) & p^0 > 0 \\ T^*(-p^0, -\mathbf{p}) & p^0 < 0. \end{cases}$$

Again the vacuum contributes to T (and not R) the amount

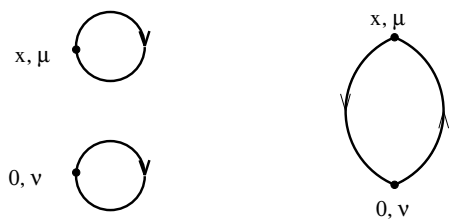
$$T_G(p^0) = (2\pi)^4 i \delta^4(p) \langle 0| O_1 |0\rangle \langle 0| O_2 |0\rangle .$$

Calculation of Vacuum Polarization If we Fourier transform the current induced by an external field, we obtain

$$\begin{aligned}\tilde{J}^\mu(k) &= \int d^4x e^{-ik\cdot x} \langle in| j^\mu(x) |in\rangle \\ &= \int d^4x e^{-ik\cdot x} \int d^4y R^{\mu\nu}(x-y) A_\nu(y) + O(A^2) \\ &= R^{\mu\nu}(k)\tilde{A}_\nu(k) + O(A^2)\end{aligned}$$

and we have just obtained the relation of $R^{\mu\nu}$ to $T^{\mu\nu}$.

To calculate $T^{\mu\nu}$ we first apply the Wick expansion to the time ordered product of four Dirac fields contained in the two current amplitude. There are two distinct contractions corresponding to the diagrams



The disconnected diagrams vanish because $\langle 0| j^\mu |0\rangle = 0$ in the absence of external fields. (This is a simple consequence of charge conjugation invariance.) The unique connected diagram has the value

$$-Q^2 \operatorname{Tr}[\gamma^\mu S_F(x)\gamma^\nu S_F(-x)]$$

where the minus sign comes from the single closed fermi loop. Inserting the known Fourier representation for S_F and carrying out the integration over x in the evaluation of $T^{\mu\nu}$ leads to

$$T^{\mu\nu}(k) = iQ^2 \int \frac{d^4p}{(2\pi)^4} \operatorname{Tr}\left(\gamma^\mu \frac{m - p\cdot\gamma}{m^2+p^2-i\epsilon}\gamma^\nu \frac{m-(p-k)\cdot\gamma}{m^2+(p-k)^2-i\epsilon}\right).$$

We immediately see from this expression that the integration over momentum is quadratically divergent at high momentum. The origin of this divergence is that $S_F(x)$ behaves like $1/x^3$ at small x which means that the two current amplitude behaves like $1/x^6$ which means that its Fourier transform is ill-defined.

Before dealing with this divergence, let us simplify the integrand by first evaluating the trace

$$\begin{aligned}N^{\mu\nu}(p,k) &\equiv \operatorname{Tr}(\gamma^\mu(m-p\cdot\gamma)\gamma^\nu(m-(p-k)\cdot\gamma)) \\ &= 8p^\mu p^\nu - 4(p^\mu k^\nu + p^\nu k^\mu) - 4\eta^{\mu\nu}(m^2 + p\cdot(p-k))\end{aligned}$$

and secondly combining denominators using the Feynman trick

$$\frac{1}{AB} = \int_0^1 dx \frac{1}{[Ax+B(1-x)]^2}$$

which is trivial to derive. Then

$$T^{\mu\nu}(k) = iQ^2 \int \frac{d^4p}{(2\pi)^4} \int_0^1 dx \frac{N^{\mu\nu}(p,k)}{[m^2+(p-kx)^2-i\epsilon+x(1-x)k^2]^2}$$

where we have completed the square in the denominator.

Next we do a step which is not quite legitimate in view of the quadratic divergence, which is to change integration variables $p \to p + kx$. After this the denominator depends only on p^2, so all terms in the numerator linear in p^μ integrate to zero and can be dropped. Furthermore terms of the form $p^\mu p^\nu$ can be replaced by $p^2\eta^{\mu\nu}/4$, since $\int d^4p f(p^2)p^\mu p^\nu$ must be proportional to $\eta^{\mu\nu}$ and the proportionality constant is then determined to be $\int d^4p f(p^2)p^2/4$ by comparing the trace of both sides. Thus we have the replacements

$$\begin{aligned} N^{\mu\nu}(p+xk,k) \to & 8(\frac{p^2}{4}\eta^{\mu\nu} + x^2k^\mu k^\nu) - 8xk^\mu k^\nu - 4\eta^{\mu\nu}(m^2 + p^2 - x(1-x)k^2) \\ \to & -2\eta^{\mu\nu}p^2 + 4x(1-x)(k^2\eta^{\mu\nu} - 2k^\mu k^\nu) - 4m^2\eta^{\mu\nu} \end{aligned}$$

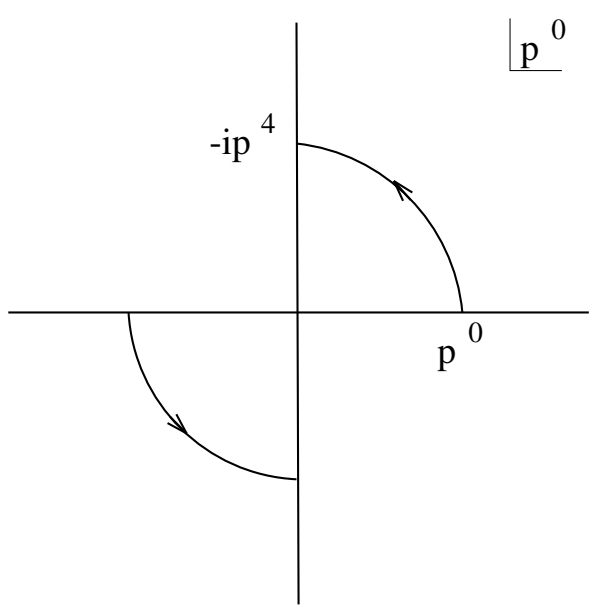
After all these steps so far we have reduced the integrals to

$$T^{\mu\nu}(k) = iQ^2 \int \frac{d^4p}{(2\pi)^4} \int_0^1 dx \frac{-2\eta^{\mu\nu}p^2 + 4x(1-x)(k^2\eta^{\mu\nu} - 2k^\mu k^\nu) - 4m^2\eta^{\mu\nu}}{[m^2 + p^2 - i\epsilon + x(1-x)k^2]^2}.$$

Next we wish to evaluate the p integral. It is easiest to think about this evaluation after the Wick rotation to Euclidean momenta, so the integral is over 4 dimensional Euclidean space and the integrand is $O(4)$ invariant. Then the angular integrals can be done and the integral reduced to a one dimensional one. The rotation of the p^0 contour to the imaginary axis must avoid the singularities due to the vanishing of the denominator which occurs at

$$p^0 = \pm\sqrt{m^2 + \mathbf{p}^2 + x(1-x)k^2 - i\epsilon}$$

These poles remain in the fourth and second quadrant of the complex p^0 plane for all values of k^2. However they get infinitesimally close to the imaginary axis for $x(1-x)k^2 \leq -m^2 - \mathbf{p}^2$ which we shall see is responsible for singular behavior in the result as a function of k^2. As long as we stick to $k^2 > -4m^2$, though, the poles stay well within their respective quadrants, and a counterclockwise contour rotation by 90 degrees encounters no singularities.



After the Wick rotation we change variables to $p^0 = ip^4$ so that $d^4p = id^4p_E$ and $p^2 = \mathbf{p}^2 + (p^4)^2$. Going to polar coordinates, $d^4p_E = p^3 dp d\Omega$, we wish to evaluate the angular integrals $d\Omega$. A useful trick to do this in any number of dimensions is to integrate a Gaussian $e^{-\vec{p}^2}$ in both Cartesian and polar coordinates. In

Cartesian coordinates in D dimensions we get $\pi^{D/2}$. In polar coordinates it is

$$\Omega_D \int_0^\infty p^{D-1} dp e^{-p^2} = \Omega_D \frac{1}{2}\Gamma(D/2)$$

Comparing we arrive at

$$\Omega_D = \frac{2\pi^{D/2}}{\Gamma(D/2)}.$$

Notice that for $D = 1, 2, 3$ this gives the well known results $2, 2\pi, 4\pi$. For our case $D = 4$ and the result is $2\pi^2$. We also record here the useful formula

$$\int d^D p \frac{(p^2)^m}{(p^2 + A^2)^n} = \frac{2\pi^{D/2}}{\Gamma(D/2)} \int_0^\infty dp \frac{p^{D-1+2m}}{(p^2 + A^2)^n} = \frac{A^{D+2m-2n}\pi^{D/2}\Gamma(m + D/2)\Gamma(n - m - D/2)}{\Gamma(D/2)\Gamma(n)}.$$

The fact that the r.h.s. is a perfectly defined analytic function of D, m, n allows for dimensional regularization as we shall see later.

Putting all this together and cutting off the p integral at Λ we have so far

$$T^{\mu\nu}_\Lambda(k) = -\frac{Q^2}{8\pi^2} \int_0^\Lambda p^3 dp \int_0^1 dx \frac{-2\eta^{\mu\nu}(p^2 + 2m^2) + 4x(1-x)(k^2\eta^{\mu\nu} - 2k^\mu k^\nu)}{[m^2 + p^2 - i\epsilon + x(1-x)k^2]^2}.$$

The p integrals are now elementary:

$$\begin{aligned}
\int_0^\Lambda \frac{p^3 dp}{[p^2 + C]^2} &= \frac{1}{2}\left[\ln\frac{\Lambda^2 + C}{C} - 1\right] + \frac{1}{2}\frac{C}{\Lambda^2 + C} \\
&= \frac{1}{2}\left[\ln\frac{\Lambda^2}{C} - 1\right] + O(\Lambda^{-2}) \\
\int_0^\Lambda \frac{p^5 dp}{[p^2 + C]^2} &= \Lambda^2 - C\ln\frac{\Lambda^2 + C}{C} - \frac{1}{2}\frac{\Lambda^4}{\Lambda^2 + C} \\
&= \frac{1}{2}\Lambda^2 + \frac{1}{2}C - C\ln\frac{\Lambda^2}{C} + O(\Lambda^{-2}),
\end{aligned}$$

where for us $C = m^2 + x(1-x)k^2 - i\epsilon$. Putting these results into the expression for T gives

$$\begin{aligned}
T^{\mu\nu}_\Lambda(k) = &-\frac{Q^2}{8\pi^2}\int_0^1 dx \Big[-\eta^{\mu\nu}(\Lambda^2 - C - 2C(\ln\frac{\Lambda^2}{C} - 1)) \\
&- (2\eta^{\mu\nu}m^2 - 2x(1-x)(k^2\eta^{\mu\nu} - 2k^\mu k^\nu)\left[\ln\frac{\Lambda^2}{C} - 1\right]\Big] \\
= &-\frac{Q^2}{2\pi^2}\int_0^1 dx x(1-x)(k^2\eta^{\mu\nu} - k^\mu k^\nu)(\ln\frac{\Lambda^2}{m^2 + x(1-x)k^2 - i\epsilon} - 1) \\
&+ \frac{Q^2}{8\pi^2}\eta^{\mu\nu}(\Lambda^2 - m^2 - \frac{k^2}{6}).
\end{aligned} \tag{14.1}$$

We could do the last integral over x, but it is actually easier to see the properties of T directly from the integral representation (14.1).

We have been casual about the way we cutoff the momentum integral, and now we can see a bad consequence of this: a violation of gauge invariance. In Fourier components a gauge transformation on the external field has the form $\tilde{A}_\mu(k) \to \tilde{A}_\mu + k_\mu \tilde{\Lambda}(k)$. The induced current was given by $R^{\mu\nu}(k)\tilde{A}_\nu(k)$, so gauge invariance would imply $R^{\mu\nu}(k)k_\nu = 0$, and in view of the relation between R and T, $T^{\mu\nu}(k)k_\nu = 0$. Clearly, the last line of (14.1) fails to satisfy this condition. The reason for this error can be traced to insufficient care with the manner in which we regularized the divergent integral. It is fortunate that the momentum dependence of the erroneous terms is a simple polynomial. This is in fact a characteristic of all such errors induced by ultraviolet divergences: Differentiating the integral a finite number of times with respect to the external momenta renders it convergent, so a finite number of derivatives must kill the mistake. In this case it would require three derivatives to kill the mistake. We shall later discuss gauge invariant regularization procedures that prevent such mistakes from occurring, provided of course that the theory can be consistently quantized. But for now we shall be satisfied with simply adjusting the polynomial dependence of our results to be consistent with gauge invariance. The nonpolynomial part of T can of course not be removed by such an adjustment. Making this adjustment, and at the same time absorbing the -1 in a rescaling of the cutoff, we then obtain

$$\begin{aligned} T^{\mu\nu}_{GI}(k) =& (k^\mu k^\nu - k^2\eta^{\mu\nu})T(k^2) \\ T(k^2) =& \frac{Q^2}{2\pi^2}\int_0^1 dx x(1-x)\ln\frac{\Lambda^2 e^{-\gamma}}{m^2 + x(1-x)k^2 - i\epsilon}. \end{aligned}$$

It is important to appreciate that gauge invariant regularization does not cure the problem of ultraviolet divergences, although it does reduce its severity. Our initial expression for T was quadratically divergent, but we have seen that gauge invariance effectively reduces the divergence to a logarithmic one. We shall see that this last divergence, although present in the quantities we are calculating, disappears after expressing the answer in terms of physically measureable parameters. Note also that our polynomial adjustment of $T^{\mu\nu}$ to make it gauge invariant allows an undetermined constant γ. However this ambiguity is linked to the cutoff dependence, and will disappear along with the latter in physical quantities.

<u>The Physics of Vacuum Polarization</u> Our result for $T^{\mu\nu}$ can now be used to give us the response function

$$R^{\mu\nu}_{GI}(k) = (k^\mu k^\nu - k^2\eta^{\mu\nu})R(k^2)$$

where

$$R(k^2) = \begin{cases} T(k^2) & k^0 > 0 \\ T^*(k^2) & k^0 < 0 \end{cases}.$$

Note that since T is real for $k^2 > -4m^2$, the two cases merge for that range of momentum. Since we have incorporated gauge invariance in our answer, we are free to fix a convenient gauge for discussing the physical interpretation of our result. Let us choose Lorentz gauge, $k^\mu \tilde{A}_\mu = 0$. Then the Fourier transform of the current induced by the external field is simply

$$\langle j_\mu(k)\rangle = -k^2 R(k^2)\tilde{A}^e_\mu(k).$$

We must now recognize that the induced currents will produce induced fields via Maxwell's equations. As long as A^e the external field is sufficiently weak the induced currents and the induced field A^{IND} will also be weak, and it will be consistent to assert that the total current is given by the response function times the

total field $A^e + A^{IND}$.

$$\langle j_\mu(k)\rangle_{TOTAL} = -k^2 R(k^2)(\tilde{A}^e_\mu(k) + \tilde{A}^{IND}_\mu(k)).$$

It is this total current that we should use in Maxwell's equations to calculate A^{IND}

$$k^2 \tilde{A}^{IND}_\mu(k) = \langle j_\mu(k)\rangle_{TOTAL} = -k^2 R(k^2)(\tilde{A}^e_\mu(k) + \tilde{A}^{IND}_\mu(k)).$$

This gives a self-consistent equation for A^{IND} in terms of the external field.

$$\tilde{A}^{IND}_\mu(k) = -\frac{R(k^2)}{1+R(k^2)}\tilde{A}^e_\mu(k).$$

Finally, if we add the induced field to the external field, we obtain the total field

$$\tilde{A}^{TOT}_\mu(k) = \frac{1}{1+R(k^2)}\tilde{A}^e_\mu(k).$$

In summary we have calculated the total electromagnetic field that arises in the "medium" of the Dirac sea in the presence of an externally applied field. The externally applied fields are what are traditionally called the $\vec{D}$ and $\vec{H}$ fields. (Recall that the sources of these fields are the external charge and current density respectively.) On the other hand the total fields are traditionally given the name $\vec{E}$ and $\vec{B}$. The dielectric "constant" of the medium is defined by $\vec{D} = \epsilon\vec{E}$ and the magnetic permeability by $\vec{B} = \mu\vec{H}$. Thus we can interpret our calculation by attributing a k dependent dielectricity and magnetic permeability to the Dirac sea

$$\epsilon(k^2) = 1/\mu(k^2) = 1 + R(k^2).$$

Since R is positive for static fields ($k^0 = 0$), the vacuum is a polarizable diamagnetic medium. The fact that $\epsilon = 1/\mu$ means that the velocity of light is unaltered by the medium (*i.e.* the medium preserves Poincarè invariance). The effectiveness of the medium in screening external fields is reduced at shorter wavelength (R decreases as k^2 increases).

Charge Renormalization We now come to the resolution of the logarithmic divergence that remains in our expression for $R(k^2)$. We begin by asking how we measure charge. We seem to have particles and antiparticles of charge $\pm Q$ in the theory. However this is not the measured charge, the charge we could define by $e^2_{ph} \equiv \lim_{R\to\infty} 4\pi R^2 F(R)$ where R is the spatial separation between two such charged particles and F is the force exerted by one on the other. In other words, Q represents the external or "bare" charge, which acts as source to the $\vec{D}$ field. The Fourier component of the latter field is just $-iQ\mathbf{k}/\mathbf{k}^2$. The measured force is given by $Q\vec{E}$ the Fourier component of which is $-iQ^2\mathbf{k}/[\mathbf{k}^2(1+R(k^2))]$. The long distance part of the force is controlled by the Fourier components with $k \approx 0$ which are clearly those of a Coulomb force with effective charge squared of

$$e^2 = \frac{Q^2}{1+R(0)}. \tag{14.2}$$

Putting $k = 0$ in our expression for R gives us

$$e^2 = \frac{Q^2}{1+(Q^2/12\pi^2)\ln(\Lambda^2/m^2)}.$$

It is e and not the parameter Q that we measure in experiments, all of which are performed within the "medium" represented by the vacuum. The fine structure constant is $\alpha = e^2/4\pi \approx 1/137$.

If e is the measured charge, then the measured electric field should be related to measured force by

$$\begin{aligned}\vec{F} =& e\vec{E}_{meas} = Q\vec{E} = \frac{Q}{1+R(k^2)}\vec{D} \\ =& \frac{e(1+R(0))}{1+R(0)+[R(k^2)-R(0)]}\frac{\vec{D}}{\sqrt{1+R(0)}} \\ =& \frac{e}{1+[R(k^2)-R(0)]/(1+R(0))}\frac{\vec{D}}{\sqrt{1+R(0)}}\end{aligned}$$

Now since $\vec{D}$ is simply proportional to Q, $\vec{D}_{meas} \equiv \vec{D}/\sqrt{1+R(0)}$ is what we can call the measured $\vec{D}$ field since it has Q replaced by e. Thus we have the following relationship between measured fields

$$\vec{E}_{meas} = \frac{1}{1+[R(k^2)-R(0)]/(1+R(0))}\vec{D}_{meas}.$$

The measured dielectric constant is accordingly

$$\epsilon(k^2) = 1 + \frac{R(k^2)-R(0)}{1+R(0)} \approx 1 + \frac{e^2}{2\pi^2}\int_0^1 dx x(1-x)\ln\frac{m^2}{m^2+x(1-x)k^2}$$

and we see that the cutoff dependence has disappeared when we express measured quantities in terms of measured parameters. This is what is known as Renormalizability, and is a feature of a wide class of quantum field theories. Such theories encounter infinities in intermediate stages of a given calculation, but the measureable quantities always come out finite.

After renormalization, the dielectric constant is fixed to be 1 at $k = 0$, *i.e.*the medium is effectively absent then. Before, we argued that the effects of the medium should be reduced at large k. By going to large k we should begin to see more and more of the bare charge. We can phrase this by defining a k dependent coupling by $\alpha(k^2) = e^2/4\pi\epsilon(k^2)$. Then as k^2 increases from zero, corresponding to shorter distances, α increases until it blows up at some finite value of k^2. It's clear that this will happen because the ln starts out at zero goes negative and behaves monotonically without bound. The pole occurs at a value of $k^2 \sim m^2\exp(12\pi^2/e^2)$. We have seen how poles in amplitudes are associated with particles. Unfortunately this pole occurs at space-like momentum, *i.e.*imaginary mass, a tachyon. This is the physically unacceptable Landau Ghost. If it were really present it would signify an inconsistency of electrodynamics. Fortunately, the whole issue is completely open since we have made approximations in our calculation that amount to weak coupling perturbation theory, and as we have seen the effective coupling gets strong at values of k^2 much less than the ghost mass, thus invalidating perturbation theory.

One way to understand why this approximation has led to this problem is to return to the bare expressions

$$e^2 = \frac{1}{1/Q^2 + (1/12\pi^2)\ln(\Lambda^2/m^2)}.$$

Now in the renormalization procedure, we attempt to take $\Lambda \to \infty$ holding e fixed. But this is only possible if Q is imaginary, which would mean we started with a Hamiltonian which was not hermitian. In order to escape this conclusion, the relation between bare and measured coupling would have to be fundamentally altered by higher order corrections. This is a logical possibility, but many field theorists including Landau doubted that the problem would go away. At this point it is appropriate to mention that for some quantum

field theories the sign in front of the logarithmic divergence is opposite to that in QED:

$$g^2 = \frac{1}{1/g_0^2 - b\ln(\Lambda^2/m^2)}, \qquad b > 0.$$

In this case one can take $\Lambda \to \infty$ with g fixed and g_0 real. This phenomenon is known as asymptotic freedom since the bare coupling goes to 0 through real values as $\Lambda \to \infty$. These quantum field theories can be renormalized consistently in the weak coupling limit as long as one restricts one's attention to very large momenta. The other side of the coin is that they become strong coupling theories at low momenta, and so weak coupling approximations are useless for studying their particle spectrum.

Next we turn to the interpretation of the singularity in T or R as $k^2 \to -4m^2$. At this point the logarithm has a branch point, leading to a discontinuity depending on whether one goes past the branch point in the upper or lower half plane:

$$\begin{aligned}
T_\pm(k^2) =& \frac{Q^2}{2\pi^2}\int_0^1 dx x(1-x)\ln\frac{\Lambda^2}{|m^2+x(1-x)k^2|} \\
&\mp i\frac{Q^2}{2\pi}\int_0^1 dx x(1-x)\theta(-x(1-x)k^2-4m^2) \\
=& \frac{Q^2}{2\pi^2}\int_0^1 dx x(1-x)\ln\frac{\Lambda^2}{|m^2+x(1-x)k^2|} \\
&\mp i\frac{Q^2}{2\pi}\int_0^1 dx x(1-x)\theta(-x(1-x)k^2-4m^2).
\end{aligned}$$

The integral in the second term contributes only if $k^2 < -4m^2$ when the range of x contributing is

$$\frac{1}{2} - \sqrt{\frac{1}{4}+\frac{m^2}{k^2}} < x < \frac{1}{2} + \sqrt{\frac{1}{4}+\frac{m^2}{k^2}}.$$

The integral over that range is elementary and yields

$$\begin{aligned}
T_\pm(k^2) =& \frac{Q^2}{2\pi^2}\int_0^1 dx x(1-x)\ln\frac{\Lambda^2}{|m^2+x(1-x)k^2|} \\
&\mp i\frac{Q^2}{2\pi}\frac{1}{3}\theta(-k^2-4m^2)\sqrt{\frac{1}{4}+\frac{m^2}{k^2}}\left(1-\frac{2m^2}{k^2}\right).
\end{aligned}$$

The $-i\epsilon$ prescription tells us to choose the lower half plane continuation, *i.e.* the lower $(+)$ sign is to be taken.

$$\begin{aligned}
T(k^2) =& \frac{Q^2}{2\pi^2}\int_0^1 dx x(1-x)\ln\frac{\Lambda^2}{|m^2+x(1-x)k^2|} \\
&+ i\frac{\alpha}{3}\theta(-k^2-4m^2)\sqrt{1+\frac{4m^2}{k^2}}\left(1-\frac{2m^2}{k^2}\right). \\
R(k^2) =& \frac{Q^2}{2\pi^2}\int_0^1 dx x(1-x)\ln\frac{\Lambda^2}{|m^2+x(1-x)k^2|} \\
&+ i\frac{\alpha_0}{3}\epsilon(k^0)\theta(-k^2-4m^2)\sqrt{1+\frac{4m^2}{k^2}}\left(1-\frac{2m^2}{k^2}\right),
\end{aligned}$$

where we have used the definition of the bare fine structure constant $\alpha_0 = Q^2/4\pi$ and the relation between

R and T.

We can associate the appearance of an imaginary part of T with pair production which becomes energetically possible for energies larger than $2m$. To see the connection recall the variational equation satisfied by the vacuum persistence amplitude

$$\begin{aligned}\delta \ln \langle out|in\rangle =& i\int d^4x \frac{\langle out|\, j^\mu(x)\,|in\rangle}{\langle out|in\rangle}\delta A_\mu(x)\\ \approx& -\int d^4x d^4y\, \langle 0,I|\, T[j_I^\mu(x) j_I^\nu(y)]\,|0,I\rangle\, A_\nu(y)\delta A_\mu(x)\end{aligned}$$

where we used first order perturbation theory in the external field to approximate the current matrix element. The results of our calculation gave

$$i\,\langle 0,I|\, T[j_I^\mu(x) j_I^\nu(y)]\,|0,I\rangle = \int \frac{d^4k}{(2\pi)^4} e^{ik(x-y)}(k^\mu k^\nu - k^2\eta^{\mu\nu})T(k^2),$$

so we find, in weak field approximation,

$$\langle out|in\rangle \approx \exp\left\{\frac{i}{2}\int \frac{d^4k}{(2\pi)^4} T(k^2)\tilde{A}_\mu(-k)(k^\mu k^\nu - k^2\eta^{\mu\nu})\tilde{A}_\nu(k)\right\}.$$

It is illuminating to express this in terms of the Fourier components of the field strengths

$$\begin{aligned}\tilde{F}_{\mu\nu}(k) =& i(k_\mu \tilde{A}_\nu(k) - k_\nu \tilde{A}_\mu(k))\\ \tilde{F}_{\mu\nu}(-k)\tilde{F}^{\mu\nu}(k) =& -2(k^\mu k^\nu - k^2\eta^{\mu\nu})\tilde{A}_\mu(-k)\tilde{A}_\nu(k)\\ \langle out|in\rangle \approx& \exp\left\{-\frac{i}{4}\int \frac{d^4k}{(2\pi)^4} T(k^2)\tilde{F}_{\mu\nu}(-k)\tilde{F}^{\mu\nu}(k)\right\}.\end{aligned}$$

As long as T is real, this is a pure phase and $|\langle out|in\rangle| = 1$. But when T acquires an imaginary part as it does for $k^2 < -4m^2$ this is no longer true. Specifically,

$$|\langle out|in\rangle|^2 \approx \exp\left\{\frac{1}{2}\int \frac{d^4k}{(2\pi)^4} \mathrm{Im}T(k^2)\tilde{F}_{\mu\nu}(-k)\tilde{F}^{\mu\nu}(k)\right\}$$

which is the probability that there is *no* pair creation. The probability of pair creation is $1 - |\langle out|in\rangle|^2$ and is approximately given by

$$\begin{aligned}\mathrm{P}_{pair} \approx& -\frac{1}{2}\int \frac{d^4k}{(2\pi)^4} \mathrm{Im}T(k^2)\tilde{F}_{\mu\nu}(-k)\tilde{F}^{\mu\nu}(k)\\ \approx& -\int \frac{d^4k}{(2\pi)^4}\frac{\alpha}{6}\theta(-k^2-4m^2)\sqrt{1+\frac{4m^2}{k^2}}\left(1-\frac{2m^2}{k^2}\right)\tilde{F}_{\mu\nu}(-k)\tilde{F}^{\mu\nu}(k),\end{aligned}$$

where we have replaced the bare fine structure constant $\alpha_0 = \alpha(1+R(0)) \to \alpha$ which is correct to the order we are calculating*.

$\star$ Strictly speaking, this calculation of $\langle out|in\rangle$ is for a fixed external field for which QA is fixed and finite. Renormalization applies only when the induced fields are included. In that case the $\tilde{F}$ appearing in these formulas are the total fields, and for them $Q\tilde{F}$ is indeed held fixed as the cutoff is removed. That is, $\sqrt{1+R(0)}\tilde{F}$ is the measured field strength.

Note that one could also find this result by calculating $|\langle out| b^{out} d^{out} |in\rangle|^2$ directly in perturbation theory and summing over all final states. It is of course important that this probability be positive (and $|\langle out|in\rangle| < 1$). Indeed, $\tilde{F}_{\mu\nu}(-k)\tilde{F}^{\mu\nu}(k)$ is negative for k timelike, because then there is a Lorentz frame where $\mathbf{k} = 0$ which implies $\mathbf{B} = 0$ so $\tilde{F}_{\mu\nu}(-k)\tilde{F}^{\mu\nu}(k) = -2|\tilde{\mathbf{E}}|^2$. Since it is an invariant, it must be negative in all frames.

When the fields have support only where T is real, the amplitude $\langle out|in\rangle$, even though it is a pure phase, gives information about the energy of the system in the presence of external fields. To get this connection consider a static field with adiabatic switching off at early and late times:

$$F_{\mu\nu}(x) = F_{\mu\nu}(\mathbf{x})\epsilon(t).$$

We take ϵ to be a symmetric function of t with central value 1 and gradual fall off to zero at times roughly $\pm T$. The exact shape is unimportant: we only need that $\tilde{\epsilon}(k^0)$ peaked at 0 with a width of order $1/T$. Clearly

$$\int dk^0 \tilde{\epsilon}(k^0)^2 = 2\pi \int dt \epsilon(t)\epsilon(-t) \approx 4\pi T.$$

Thus in the limit $T \to \infty$, $\tilde{\epsilon}^2$ can be approximated by $4\pi T\delta(k^0)$. Thus in this limit our approximate formula for $\langle out|in\rangle$ reads

$$\langle out|in\rangle \approx \exp\left\{-2iT \int \frac{d^3k}{(2\pi)^3} \frac{1}{4} \tilde{F}_{\mu\nu}(-\mathbf{k})\tilde{F}^{\mu\nu}(\mathbf{k}) T(\mathbf{k}^2)\right\}.$$

The coefficient of $-2iT$ is just the energy of the Dirac system in the presence of static external fields. Since T is positive for spacelike k, and

$$\frac{1}{4}\tilde{F}_{\mu\nu}(-\mathbf{k})\tilde{F}^{\mu\nu}(\mathbf{k}) = \frac{1}{2}[|\tilde{\mathbf{B}}|^2 - |\tilde{\mathbf{E}}|^2]$$

we see that the energy increases under the addition of a magnetic field and decreases with an electric field in accord with our conclusion that the vacuum is a diamagnetic dielectric medium.

<u>Superconductivity and the Higgs Mechanism</u> There is one physical phenomenon which can occur in quantum gauge field theories, although not for QED in four dimensions. It can happen in some theories that $R(k^2)$ possesses a pole at zero:

$$R(k^2) \underset{k^2 \to 0}{\sim} \frac{K}{k^2}.$$

Note that this infrared singularity does not occur in our previous calculation even for $m = 0$. But if it does occur, then

$$\tilde{A}(k)_{TOT} = \frac{1}{1 + R(k^2)}\tilde{A}_e \underset{k^2 \to 0}{\sim} \frac{k^2}{K} A_e \sim 0.$$

Thus A_{TOT} is screened at long wavelengths by the induced currents. In particular, for a static Coulomb potential $\tilde{A}^0 \sim \delta(k^0)/k^2$, the singularity at vanishing k disappears so in coordinate space the potential falls off faster than any power. Because of Lorentz covariance this screening is effective for *both* electric and magnetic fields. The screening of the magnetic fields means there is a Meissner effect, *i.e.* the vacuum in this situation is a relativistic superconductor. The vanishing of the total field at $k^2 = 0$ means that the vacuum cannot support massless photons. On the other hand there most likely is a negative value of k^2 call it $-M^2$ for which $1 + R(-M^2) = 0$. For such values of k, A_{TOT} can be nonzero even for vanishing external field. These waves correspond to particles of mass M. Thus this phenomenon, sometimes called the Higgs

mechanism, gives the photon a mass without violating gauge invariance. A nonrelativistic version of this effect was long known for superconductors. It was first discussed in the context of relativistic quantum field theory by Schwinger for QED in two space-time dimensions. Four dimensional versions were first discussed by Higgs, Englert and Brout, and Guralnik, Hagen and Kibble. In spite of the long list of discoverers, it seems that Higgs' name has stuck. The Higgs mechanism is at the heart of the electroweak unified gauge theory, because it is responsible for the masses of the W and Z bosons.

15. Operator Quantization of the Electromagnetic Field

Treated classically the e.m. field $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ satisfies Maxwell's equations

$$\partial_\nu F^{\mu\nu} = J^\mu,$$

which imply current conservation $\partial_\mu J^\mu = 0$ for consistency. So far we have developed the theory of Dirac and scalar quantum fields interacting with fixed external e.m. fields. For such systems the current is of course an operator, so Maxwell's equations imply that the e.m. field must also be a quantum operator, which inserted into the Dirac equation produces a nonlinear quantum field equation.

15.1. Quantized Electromagnetic Field Interacting with a Conserved Current

The first step in developing the quantum theory for A_μ is to understand the canonical structure of Maxwell's equations. One first notes that the equations follow from stationarity of the action

$$S = -\frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu} + \int d^4x A_\mu J^\mu,$$

which is gauge invariant provided $\partial_\mu J^\mu = 0$. It is straightforward to find the momentum conjugate to A_μ

$$\Pi^\mu = F_0{}^\mu = \partial_0 A^\mu - \partial^\mu A_0.$$

The spatial components of Π are just those of minus the electric field strength, but $\Pi^0 = 0$. This last fact poses a difficulty for quantization since it is inconsistent with nonvanishing canonical commutation relations. We know how to assign operator properties to $\vec{A}$ and $\vec{\Pi}$, but not to A_0. Before facing this difficulty, we construct the canonical Hamiltonian

$$\begin{aligned} H_{CAN} &= \int d^3x[\vec{\Pi}\cdot\vec{A} - \mathcal{L}] \\ &= \int d^3x\left[\frac{1}{2}\vec{\Pi}^2 + \frac{1}{2}(\nabla\times\vec{A})^2 - \vec{A}\cdot\vec{J} + \nabla A_0\cdot\vec{\Pi} - A_0 J^0\right]. \end{aligned}$$

Notice that the troublesome variable A_0 appears only linearly and in the last two terms. After an integration by parts the coefficient is just $-(\nabla\cdot\vec{\Pi} + J^0)$, which would vanish if we could use Gauss' Law.

Classically, we could certainly enforce Gauss' Law and then the canonically uncertain A_0 would disappear from the dynamics. Attempting to enforce Gauss' Law as a quantum operator equation would contradict the canonical commutation relations for $\vec{A}, \vec{\Pi}$

$$[A_k(\vec{x},t), \Pi^m(\vec{y},t)] = i\delta_k^m\delta(\vec{x}-\vec{y}),$$

so we postpone discussion of this point. If we leave the operator character of A_0 unspecified, it is clearly important to know how quantum evolution will be affected if we make a change in A_0. Let the evolution operator for a given A_0 be $U_{A_0}(t,-\infty)$. Then by a familiar argument

$$\delta U_{A_0}(t,-\infty) = iU_{A_0}(t,-\infty)\int d^4x\delta A_0(\nabla\cdot\vec{\Pi} + J^0)$$

where the operators multiplying U_{A_0} are in Heisenberg picture. The Heisenberg equations for $\vec{A}, \vec{\Pi}$ imply

$$\begin{aligned} \frac{\partial}{\partial t}(\nabla\cdot\vec{\Pi} + J^0) =& \frac{\partial J^0}{\partial t} + \nabla\cdot\vec{J}i\int d^3x[\nabla\cdot\vec{\Pi}, A_0](\nabla\cdot\vec{\Pi} + J^0) \\ =& i\int d^3x[\nabla\cdot\vec{\Pi}, A_0](\nabla\cdot\vec{\Pi} + J^0) \end{aligned} \tag{15.1}$$

Although we are not free to impose Gauss' Law as an operator equation, we can restrict our incoming

states to satisfy

$$(\nabla \cdot \vec{\Pi} + J^0)|_{t=-\infty} \, |in\rangle = 0.$$

It then follows from (15.1) that $(\nabla \cdot \vec{\Pi} + J^0) \, |in\rangle = 0$ for all time. By making this restriction we therefore arrange that

$$\delta U_{A_0} \, |in\rangle = 0$$

for arbitrary operator changes in A_0, that is, the quantum evolution is independent of how we treat A_0. This is the quantum analog of what is true in the classical treatment, and suffices to resolve the difficulties. By imposing Gauss' Law on initial states, we are free to make a choice for A_0 which simplifies the dynamical problem.

For example, suppose we want to "solve" the constraint by "imposing Coulomb gauge." We do this by writing

$$\begin{aligned} \vec{\Pi} =& \vec{\Pi}_T + \frac{1}{\nabla^2} \vec{\nabla} \nabla \cdot \vec{\Pi} \\ =& \vec{\Pi}_T - \frac{1}{\nabla^2} \vec{\nabla} J^0 + \frac{1}{\nabla^2} \vec{\nabla} (\nabla \cdot \vec{\Pi} + J^0) \end{aligned} \tag{15.2}$$

so that $\nabla \cdot \vec{\Pi}_T = 0$. Inserting (15.2) into the term in the Hamiltonian containing Π and judiciously integrating by parts, we obtain

$$\begin{aligned} \int d^3x \, \vec{\Pi}^2 =& \int d^3x \, \vec{\Pi}_T^2 + \int J^0 \left(-\frac{1}{\nabla^2} \right) J^0 \\ &+ \int \left(2 \frac{1}{\nabla^2} J^0 - \frac{1}{\nabla^2} (\nabla \cdot \vec{\Pi} + J^0) \right) (\nabla \cdot \vec{\Pi} + J^0) + \int [\frac{1}{\nabla^2} J^0, \nabla \cdot \vec{\Pi} + J^0]. \end{aligned}$$

The last term, which arises because we reordered operators so that $\nabla \cdot \vec{\Pi} + J^0$ stands on the right wherever it appears, formally vanishes because the fields entering J are canonically independent of Π and also J^0 commutes with itself at equal times. The next to last term has a factor of $\nabla \cdot \vec{\Pi} + J^0$ on the right. By choosing

$$A_0 = \frac{1}{\nabla^2} J^0 - \frac{1}{2\nabla^2} (\nabla \cdot \vec{\Pi} + J^0),$$

this term is cancelled by the term linear in A_0 and the Hamiltonian then simplifies to

$$H_{COUL} = \int d^3x \, \vec{\Pi}_T^2 + \int J^0 \left(-\frac{1}{\nabla^2} \right) J^0 + \int (\frac{1}{2} (\nabla \times \vec{A})^2 - \vec{A} \cdot \vec{J}).$$

Since only Π_T now enters the Hamiltonian, it is appropriate to make a similar decomposition of $\vec{A} = \vec{A}_T + \vec{\nabla}(1/\nabla^2) \nabla \cdot \vec{A}$. The gradient term does not contribute to the curl of A and an integration by parts allows that term in H_{COUL} to be simplified as

$$\begin{aligned} \int (\nabla \times \vec{A})^2 =& \int (\nabla \times \vec{A}_T)^2 = \int \vec{A}_T \cdot (\nabla \times (\nabla \times \vec{A}_T)) \\ =& \int \vec{A}_T \cdot (-\nabla^2) \vec{A}_T = \int \nabla_k \vec{A}_T \cdot \nabla_k \vec{A}_T. \end{aligned}$$

The longitudinal component of $\vec{A}$ appears in H_{COUL} through the coupling to $\vec{J}$. To interpret this term we have to consider a little more closely the inner product structure on the space of states. The states which satisfy the Gauss' Law constraint have the character of momentum eigenstates with a fixed eigenvalue. Such

states would of course have infinite norm. If we restrict our *in* states to satisfy Gauss' Law, and we want our *out* states to have a finite inner product with these *in* states, it follows that the *out* states should not be taken to satisfy the constraint. In fact, the natural dual space to momentum eigenstates are position eigenstates, with $\langle q|p\rangle \propto e^{iqp}$. Thus we should choose our *out* states to be eigenstates of $\nabla\cdot\vec{A}$, which is the conjugate variable to $\nabla\cdot\vec{\Pi}$. Since $\nabla\cdot A$ commutes with the Hamiltonian, as a Heisenberg picture operator it will be independent of time. Thus, if we choose our *out* states to be eigenstates of $\nabla\cdot\vec{A}(\vec{x},+\infty)$, with eigenvalue $A_L(\vec{x})$ they will also be eigenstates of $\nabla\cdot\vec{A}(\vec{x},-\infty)$ with the same eigenvalue.

We can always choose phases, as in the standard Schrödinger representation, so that $i\,\langle A', out|\,\vec{\Pi} = (\delta/\delta A')\,\langle A', out|$. Then the dependence of persistence amplitudes on the choice of A_L is determined by

$$\begin{aligned} i\delta\,\langle A_L, out|in\rangle &= \int d^3x\ \delta A_L(\vec{x})\,\langle A_L, out|\left(-\frac{1}{\nabla^2}\right)\nabla\cdot\Pi(\vec{x},+\infty)\,|in\rangle \\ &= \int d^3x\ \delta A_L(\vec{x})\,\langle A_L, out|\left(+\frac{1}{\nabla^2}\right)J^0(\vec{x},+\infty)\,|in\rangle\,. \end{aligned}$$

Since $J^0(\vec{x},t)$ is just the infinitesimal generator of local phase changes on the charged Heisenberg fields at time t, we see that an infinitesimal change in A_L can be compensated by a local gauge transformation on the charged fields at $t=+\infty$, which determine the definition of the *out* states. The principle of gauge invariance includes the statement that state vectors differing by such gauge changes describe the same physical state, since they will give identical predictions for all gauge invariant observations. Thus we are free to fix A_L to be any convenient function. Coulomb or radiation gauge corresponds to the choice $A_L=0$.

We have generally defined $\langle out|$ by the $t\to\infty$ limit of $\langle t| = \langle in|\,U(t,-\infty)$. Our choice of $\langle A_L, out|$ is implemented by the replacement $\langle in| \to \langle A_L, in|$ where the latter is the eigenstate of $\nabla\cdot A(\vec{x},-\infty)$ with eigenvalue A_L. Then the corresponding $\langle A_L, t|$ is the corresponding eigenstate of $\nabla\cdot A(\vec{x},t)$. Since $\langle A_L, t|$ satisfies $i(\partial/\partial t)\,\langle A_L, t| = \langle A_L, t|\,H_{COUL}(t)$, we see that replacing $\nabla\cdot A$ in H_{COUL} by A_L will yield the same $\langle A_L, out|$. In particular the radiation gauge choice $A_L=0$ leads to the effective Hamiltonian

$$H_{eff} = \int d^3x\ \vec{\Pi}_T^2 + \int J^0\left(-\frac{1}{\nabla^2}\right)J^0 + \int(\frac{1}{2}\partial_k\vec{A}_T\cdot\partial_k\vec{A}_T - \vec{A}_T\cdot\vec{J}). \tag{15.3}$$

One may use H_{eff} to compute any physical quantity. It only contains the transverse components of $\vec{A}$ and $\vec{\Pi}$. From the canonical commutation relations for these two quantities one can easily evaluate those for the transverse components

$$[A_{Tk}(\vec{x}), \Pi_{Tm}(\vec{y})] = i\left(\delta_{km} - \frac{\nabla_k\nabla_m}{\nabla^2}\right)\delta(\vec{x}-\vec{y}).$$

The operator acting on the delta function simply reflects the fact that the l.h.s. has vanishing divergence because of the transversality of the operators.

We have obtained the effective Coulomb gauge Hamiltonian (15.3) by reduction from a gauge independent quantization procedure. A much quicker route to the same answer is to fix the gauge before quantization by setting $\nabla\cdot A=0$ from the beginning. Then the Gauss' law constraint can be "solved" by setting

$$A^0(\vec{x},t) = -\frac{1}{\nabla^2}J^0(\vec{x},t) = \int d^3y\frac{J^0(\vec{y},t)}{4\pi|\vec{x}-\vec{y}|}.$$

The longitudinal component of $\vec{\Pi}$ is also eliminated because $\nabla\cdot\vec{\Pi} = \nabla^2 A^0$. One passes to quantum mechanics by promoting only $\vec{A}$ and Π_T to operators. (If the currents are operators, A^0 is an operator by virtue of the constraint, but it is not independent.) Then the transverse projector must appear on the r.h.s. of the canonical commutation relations.

Polarization and Helicity of Photons. An explicit realization of the commutation relations can be given in terms of creation and annihilation operators as follows:

$$A_{Tk}(\vec{x},0) = \int \frac{d^3k}{\sqrt{(2\pi)^3 2|\mathbf{k}|}}[a_k(\vec{k})e^{i\vec{k}\cdot\vec{x}} + a_k^\dagger(\vec{k})e^{-i\vec{k}\cdot\vec{x}}]$$
$$\Pi_{Tk}(\vec{x},0) = -\,i\int \frac{d^3k\sqrt{|\mathbf{k}|}}{\sqrt{(2\pi)^3 2}}[a_k(\vec{k})e^{i\vec{k}\cdot\vec{x}} - a_k^\dagger(\vec{k})e^{-i\vec{k}\cdot\vec{x}}]$$

with

$$[a_k(\vec{k}), a_m^\dagger(\vec{q})] = (\delta_{km} - \frac{k_k k_m}{\mathbf{k}^2})\delta(\vec{k}-\vec{q}).$$

Inserting these into H_{eff} gives

$$H_{eff} = \int d^3k|\vec{k}|\vec{a}^\dagger(\vec{k})\cdot\vec{a}(\vec{k}) - \int d^3x\vec{A}_T\cdot\vec{J} + E_0,$$

where E_0 is the usual (infinite) zero point energy of the oscillators which will be dropped from now on. This formula shows us immediately that for $\vec{J}=0$, the quantum e.m. field is interpretable as a system of massless bosons (photons). The vacuum $|0\rangle$ is defined by $a_k(\vec{k})\,|0\rangle = 0$ and the n photon state is represented by

$$a_{m_1}^\dagger(\vec{q}_1)a_{m_2}^\dagger(\vec{q}_2)\cdots a_{m_n}^\dagger(\vec{q}_n)\,|0\rangle\,.$$

Because of transversality there are two photon states for each momentum. These two polarization states will next be shown to correspond to the two helicities ± 1 of the photon.

First for fixed $\vec{k}$ let us introduce two (in general complex) basis vectors $\vec{\epsilon}_a$, $a=1,2$ for the plane perpendicular to $\vec{k}$, satisfying $\vec{k}\cdot\vec{\epsilon}_a = 0$ and the orthonormality and completeness relations

$$\vec{\epsilon}_a\cdot\vec{\epsilon}_b^{\,*} = \delta_{ab}$$
$$\sum_a \vec{\epsilon}_a^{\,m}\cdot\vec{\epsilon}_a^{\,n*} = \delta_{mn} - \frac{k^m k^n}{\mathbf{k}^2}.$$

We can then introduce two independent sets of creation and annihilation operators via

$$\vec{a}(\vec{k}) = \sum_a \vec{\epsilon}_a a_a(\vec{k}).$$

We shall relate the multiplicity associated with the index a to the spin of the photon. First recall the classical expression for the angular momentum carried by the e.m. field,

$$\vec{J} = \int d^3x\vec{x}\times(\vec{E}\times\vec{B})$$
$$= \int d^3x\sum_k E_k(\vec{x}\times\vec{\nabla})A_k - \int d^3x\vec{x}\times(\vec{E}\cdot\vec{\nabla})\vec{A}.$$

We can recognize the first term in the last line as the "orbital" angular momentum, which will not contribute to the helicity of a one photon state. This is because acting on a one photon state the ∇ is replaced by $\vec{k}$

and because of the cross product the term will be perpendicular to $\vec{k}$. The second term, after an integration by parts becomes

$$\begin{aligned}\vec{S} &= \int d^3x \vec{E} \times \vec{A} \\ &= -i \int d^3k \vec{a}^\dagger(\vec{k}) \times \vec{a}(\vec{k}).\end{aligned}$$

Applying $\vec{S}$ to a one photon state $a_a^\dagger(\vec{k})\,|0\rangle$, yields

$$\vec{S} a_a^\dagger(\vec{k})\,|0\rangle = i \sum_b (\vec{\epsilon}_a \times \vec{\epsilon}_b) a_b^\dagger(\vec{k})\,|0\rangle\,.$$

Thus we see that the 2×2 matrix $\vec{S}_{ab} = \vec{\epsilon}_a \times \vec{\epsilon}_b$ acts as a spin matrix on the index of the creation operator. To get the helicity interpretation, consider the case of $\vec{k} = k\hat{z}$. Then the helicity matrix is

$$S^3_{ab} = i(\epsilon^1_a \epsilon^2_b - \epsilon^2_a \epsilon^1_b).$$

This matrix is $diag\{1, -1\}$ with the choices

$$\vec{\epsilon}_1 = (1, i, 0)/\sqrt{2} \qquad \vec{\epsilon}_2 = (1, -i, 0)/\sqrt{2},$$

so with this choice of polarization vectors, $a_1^\dagger$ creates a photon with helicity $+1$ and $a_2^\dagger$ creates a photon with helicity -1. This establishes that the photon is a spin one particle. There is no zero helicity state for the photon: this is consistent with Poincaré invariance because the photon is massless.

The polarization vector enters scattering amplitudes multilinearly, with a factor of ϵ for each incoming photon and a factor ϵ^* for each outgoing photon. Its four-vector index forms a Minkowski scalar product with that of the vertex coupling the gauge potential to the charged fields. According to gauge invariance, this vertex satisfies current conservation: its scalar product with the momentum entering it gives zero. Thus changing each polarization vector by an amount proportional to its four-momentum leaves the scattering amplitude unaltered. In Coulomb gauge the polarization vector is of the form $\epsilon = (\boldsymbol{\epsilon}, 0)$ with $\mathbf{k} \cdot \boldsymbol{\epsilon} = 0$, so $k_\mu \epsilon^\mu = 0$. But $k_\mu k^\mu = 0$ since the photon is massless. Thus we can characterize the polarization vector completely by the covariant condition $k_\mu \epsilon^\mu = 0$. Any further specification, *e.g.* $\epsilon^0 = 0$, is merely a gauge choice which can be made at will and exploited to simplify detailed calculations. This is particularly advantageous in the calculation of Compton scattering for polarized photons.

15.2. Charged Fields Interacting with the Quantized Electromagnetic Field

We have seen that when the e.m. field interacts with a conserved current, the time component of the current enters the Hamiltonian only through the coefficient in the term linear in A_0 so that the coefficient of A_0 is just the Gauss' Law constraint with nonzero charge density. This feature is quite general, even when the e.m. field couples to dynamical charged fields. This is obvious in the case of the Dirac field because the Dirac field Hamiltonian is linear in A_μ. The A dependence of the scalar field Hamiltonian includes quadratic pieces in the *spatial* components of the potential, but nonetheless A_0 still enters only linearly after the Hamiltonian is expressed solely in terms of coordinates and momenta. This means that the elimination of A_0 in the passage to Coulomb gauge proceeds exactly as in the previous section.

Thus, the Hamiltonian for e.m. field plus charged fields in Coulomb gauge will quite generally be of the form

$$\begin{aligned} H_{eff} = &\int d^3x \left(\frac{1}{2}\mathbf{\Pi_T}^2 + \frac{1}{2}\partial_k \mathbf{A}_T \cdot \partial_k \mathbf{A}_T - \mathbf{A}_T \cdot \mathbf{J}_e \right) \\ &+ \int (J_e^0 + j^0) \left(-\frac{1}{2\nabla^2} \right) (J_e^0 + j^0) + H_{fields}|_{A_0=0}, \end{aligned} \tag{15.4}$$

where J_e^μ is an optional external current, j^μ is the current operator for the fields, and the subscript *fields* refers to the Hamiltonian operator for any dynamical fields in the system. For the Dirac field

$$\begin{aligned} j^0 =& Q[\bar{\psi}, \gamma^0 \psi] \\ H_{fields} =& \int d^3x \bar{\psi}(\frac{1}{i}\gamma \cdot \partial + m - Q\boldsymbol{\gamma} \cdot \mathbf{A})\psi, \end{aligned}$$

and for the scalar field,

$$\begin{aligned} j^0 =& -iQ(\pi\phi - \phi^\dagger \pi^\dagger) \\ H_{fields} =& \int d^3x (\pi\pi^\dagger + m^2 \phi^\dagger \phi + (\nabla + iQ\mathbf{A})\phi^\dagger \cdot (\nabla - iQ\mathbf{A})\phi. \end{aligned}$$

One obstacle to formulating an efficient perturbation theory for systems with interacting quantum fields is that one can't "turn off" the interactions at early and late times as is possible with externally applied fields. Thus *out* and *in* states are eigenstates of complicated interacting Hamiltonians. One approach to this difficulty is to artificially make the coupling constants time dependent and force them to vanish at early and late times. Another approach, which we shall favor, is to relax the requirement that the initial and final states be eigenstates of the Hamiltonian with vanishing external fields. Then one calculates in first instance a quantity that is not of immediate interest, but which can be simply related to such quantities.

A quantity of more or less direct physical interest is the vacuum expectation value of the time ordered product of several quantum fields. More generally the *outin* matrix element of such a time ordered product is relevant if time varying external fields are present. So let us consider how to obtain this quantity in perturbation theory by first calculating with general initial and final states. Using the evolution operator and assuming $t_1 > t_2 > \cdots > t_n$, we therefore consider

$$\begin{aligned} &\langle f| U(\infty, -\infty) T[\Omega_1(t_1) \cdots \Omega_n(t_n)] |i\rangle = \\ &\qquad \langle f| U(\infty, t_1)\Omega_{S1} U(t_1, t_2) \cdots U(t_{n-1}, t_n)\Omega_{Sn} U(t_n, -\infty) |i\rangle . \end{aligned}$$

Choose the time T so that all external fields vanish for times earlier than $-T$ and later than T. Then

$$\begin{aligned} U(t_n, -\infty)|i\rangle =& U(t_n, -T) e^{-i(\infty - T)H_S} |i\rangle \\ =& U(t_n, -T) e^{-i(\infty - T)E_G} \sum_r e^{-i(\infty - T)(E_r - E_G)} |r\rangle \langle r|i\rangle . \end{aligned}$$

We would now like to argue that the infinite oscillations wash out all contributions but the (assumed nondegenerate*) ground state. In a field theory this is quite plausible since the excited states correspond to particles so the sum over r is really an integral over a range of continuous energies. But even without

$\star$ There are interesting cases of degenerate vacua, when there is "spontaneous symmetry breakdown." In such cases the choice of initial and final states determines which of the degenerate vacua is picked out.

this smearing, we can make the washing out rigorous by calculating with imaginary time: $it = \beta > 0$. Then $i\infty$ is really $+\infty$ and all excited states are damped exponentially. Massless particle states could introduce a subtlety here, but the part of phase space that is not exponentially damped is infinitesimal: this has the effect of changing exponential damping to a power law damping. If we buy this argument, then we can assert quite generally that $U(t_n, -\infty)\,|i\rangle = U(t_n, -\infty)\,|0\rangle\,\langle 0|i\rangle$ and similarly $\langle f|\,U(\infty, t_1) = \langle f|0\rangle\,\langle 0|\,U(\infty, t_1)$.

Since we take (as usual) the Heisenberg and Schrödinger pictures to coincide at $t = -\infty$, then $|in\rangle = |0\rangle$ and $\langle out| = \langle 0|\,U(\infty, -\infty)$. Thus we have obtained the relation

$$\begin{aligned}\langle f|\,U(\infty,-\infty)T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|i\rangle = \\ \langle f|0\rangle\,\langle 0|i\rangle\,\langle out|\,T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|in\rangle\,.\end{aligned}$$

In other words calculating with any initial and final states that have finite overlap† with the true ground state gives us a constant times the desired matrix element. We can easily evaluate the multiplicative constant by considering by the same reasoning

$$\begin{aligned}\langle f|\,U(\infty,-\infty)\,|i\rangle &= \langle f|0\rangle\,\langle 0|i\rangle\,\langle out|in\rangle \\ &\to e^{-2i\infty E_G}\,\langle f|0\rangle\,\langle 0|i\rangle \qquad \text{External Fields} = 0.\end{aligned}$$

Putting this into our relation we obtain

$$\begin{aligned}\langle out|\,T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|in\rangle = \\ e^{-2i\infty E_G}\,\frac{\langle f|\,U(\infty,-\infty)T[\Omega_1(t_1)\cdots\Omega_n(t_n)]\,|i\rangle}{\langle f|\,U(\infty,-\infty)_{Ext=0}\,|i\rangle},\end{aligned} \tag{4.13}$$

where the subscript on U in the denominator denotes vanishing external fields. In field theory applications E_G is the energy of the vacuum, which is zero if we measure all energies relative to that of the vacuum. In the absence of gravity all physical quantities depend only on energy differences, so we lose nothing by doing this. Gravity couples directly to the energy density and therefore is sensitive to the energy as opposed to energy differences, but then E_G only appears in the combination $\Lambda \equiv E_G + \Lambda_0$, with Λ_0 the "bare" cosmological constant cosmological constant. Replacing Λ_0 by Λ in effect sets $E_G = 0$.

The formula (4.13) is a convenient starting point for developing perturbation theory. Any breakup

$$H_S(t) = H_0(t) + H'(t)$$

determines an interaction picture defined by

$$\Omega_I(t) = U_0^{-1}(t,-\infty)\Omega_S U_0(t,-\infty) = U_I^{-1}(t,-\infty)\Omega(t)U_I(t,-\infty),$$

where

$$\begin{aligned}i\dot{U} &= H_S(t)U = UH(t) \\ i\dot{U}_0 &= U_0 H_{0I}(t) \\ i\dot{U}_I &= H_I(t)U_I\end{aligned}$$

† The infinite number of degrees of freedom in quantum field theory requires care here: the overlap between different states in a theory with n degrees of freedom is typically f^n with $f < 1$. Since $n = \infty$, we should expect $\langle f|0\rangle\,\langle 0|i\rangle \sim e^{-\infty}$. In field theory $n = \infty$ because the volume of space is infinite and because space is continuous. Thus strict application of the above relation should be done in the presence of both an infrared and ultraviolet cutoff, which can then be removed after extracting the desired amplitude.

and all U's are the identity at $t = -\infty$. Then the evolution operator satisfies

$$U(t_1, t_2) = U_0(t_1, -\infty) U_I(t_1, t_2) U_0(t_2, -\infty).$$

Plugging these relations into (4.13) then gives

$$\begin{aligned} &\langle out| T[\Omega_1(t_1) \cdots \Omega_n(t_n)] |in\rangle = \\ &\qquad e^{-2i\infty E_G} \frac{\langle f| U_0(\infty, -\infty) T[U_I(\infty, -\infty) \Omega_{1I}(t_1) \cdots \Omega_{nI}(t_n)] |i\rangle}{\langle f| U_0(\infty, -\infty)_{Ext=0} U_I(\infty, -\infty)_{Ext=0} |i\rangle}. \end{aligned} \tag{4.14}$$

This formula is completely general: we have even allowed H_0 to contain time varying external fields, which is hardly ever done in practice. Since all operators in this formula are in interaction picture, it is most convenient to choose $|i\rangle, |f\rangle$ to have simple properties with respect to $H_{0I}(-\infty)$. Let us call the ground state of this operator $|in, 0\rangle$. Then $\langle in, 0| U_0(\infty, -\infty)$ is the ground state of $H_{0I}(+\infty)$ and therefore deserves the name $\langle out, 0|$. When all external fields vanish, H_{0I} is time indepenent and we call its ground state $|0, I\rangle \equiv |in, 0\rangle$ and its ground state energy E_0. Then $\langle in, 0| U_0(\infty, -\infty)_{Ext=0} = e^{-2i\infty E_0} \langle 0, I|$. Thus choosing $|i\rangle = |f\rangle = |0, I\rangle = |in, 0\rangle$ we obtain the useful formula

$$\begin{aligned} &\langle out| T[\Omega_1(t_1) \cdots \Omega_n(t_n)] |in\rangle = \\ &\qquad e^{-2i\infty (E_G - E_0)} \frac{\langle out, 0| T[U_I(\infty, -\infty) \Omega_{1I}(t_1) \cdots \Omega_{nI}(t_n)] |in, 0\rangle}{\langle 0, I| U_I(\infty, -\infty)_{Ext=0} |0, I\rangle}. \end{aligned} \tag{4.15}$$

In the usual case where we do not include external fields in H_0, the formula simplifies further

$$\begin{aligned} &\langle out| T[\Omega_1(t_1) \cdots \Omega_n(t_n)] |in\rangle = \\ &\qquad e^{-2i\infty E_G} \frac{\langle 0, I| T[U_I(\infty, -\infty) \Omega_{1I}(t_1) \cdots \Omega_{nI}(t_n)] |0, I\rangle}{\langle 0, I| U_I(\infty, -\infty)_{Ext=0} |0, I\rangle}. \end{aligned} \tag{4.16}$$

Using the Wick expansion one can describe the perturbation series for the numerators and denominators of these formulas using Feynman diagrams. The diagrams contributing to the denominator are all those completely disconnected from either external fields or from the points assigned to the operators in the numerator. The numerator contains this same sum of diagrams as a multiplicative factor. Thus the division by the denominator is achieved by simply deleting all such disconnected "vacuum" diagrams from the expansion of the numerator.

16. Ward Identities, Ultraviolet Divergences and Gauge Invariance Anomalies

16.1. Ward Identities

One might think that gauge-fixing would exhaust the implications of gauge invariance on correlation functions. This is not really true because physics must not depend on **which** gauge is chosen, and this gauge independence implies many relationships among different correlation functions known as Ward identities.

Recalling the gauge fixing procedure, it is clear that a necessary condition for gauge independence is that the part of the path integral describing the coupling of charged fields to the gauge fields, $\langle out|in\rangle_A$ must be invariant under a gauge change of A:

$$D_\mu \frac{\delta}{\delta A_\mu(z)} \langle out|in\rangle_A = 0.$$

Here we have written the condition for a general gauge theory; for QED $D = \partial$. If this is inserted for the $\langle out|in\rangle_A$ factor in the path integral for a correlation function of gauge invariant operators, it becomes the statement that the correlator of the divergence of the gauge current with any set of gauge invariant operators is conserved (abelian case) or related to another correlator (nonabelian case).

If gauge noninvariant operators also appear in the correlator, more contributions appear. As an important example of this, take the case of additional charged fields in QED. Then the gauge transform induces a change in the phase of each charged field $\psi \to e^{iQ_0\Lambda}\psi$. For example, if only $\psi, \bar\psi$ appear we find

$$\partial_\mu \left\langle j^\mu(z)\psi(x)\bar\psi(y)\right\rangle = -Q_0[\delta(x-z) - \delta(y-z)] \left\langle \psi(x)\bar\psi(y)\right\rangle .$$

The VEV (vacuum expectation) on the left hand side of this Ward identity is simply related to the vertex to which a photon propagator couples. In fact, all diagrams in which the current couples to a closed charged field line contribute nothing to the l.h.s.* so we can apply the identity to the one photon line irreducible vertex. To get an identity for the completely one particle irreducible vertex Γ, we write

$$\begin{aligned} i\left\langle j^\mu(z)\psi(x)\bar\psi(y)\right\rangle_{1\gamma IR} &= \int d^4x' d^4y' \left\langle \psi(x)\bar\psi(x')\right\rangle i\left\langle j^\mu(z)\psi(x')\bar\psi(y')\right\rangle_{1PIR} \left\langle \psi(y')\bar\psi(y)\right\rangle \\ &\equiv \int d^4x' d^4y' \left\langle \psi(x)\bar\psi(x')\right\rangle iQ_0\Gamma(z,x',y') \left\langle \psi(y')\bar\psi(y)\right\rangle . \end{aligned}$$

We state the relevant identity in momentum space, for which we define

$$\left\langle \psi(x)\bar\psi(y)\right\rangle = \int \frac{d^4p}{(2\pi)^4} e^{ip\cdot(x-y)} \frac{-i}{m_0 + \gamma\cdot p + \Sigma(p)},$$

where Σ is the electron's proper self energy part. In momentum space, the Ward identity then reads:

$$(p'-p)_\mu \Gamma^\mu(p',p) = \gamma\cdot(p'-p) + \Sigma(p') - \Sigma(p).$$

This is actually a generalized form of the original Ward Identity which follows from the generalized one plus an assumption on the absence of certain infrared singularities.

* This follows simply from the gauge invariance of $\langle out|in\rangle_A$.

To get it, take the derivative of the above equation with respect to p', and set $p' = p$. Then, assuming that $(p'-p)_\mu \partial/\partial p'_\nu \Gamma^\mu \to 0$ in this limit, we obtain

$$\Gamma^\mu(p,p) = \gamma^\mu + \frac{\partial \Sigma(p)}{\partial p_\mu}.$$

An example of the sort of singularity that would invalidate this derivation would be a contribution to Γ^μ of the form

$$\frac{(p'-p)^\mu \gamma \cdot (p'-p)}{(p'-p)^2}.$$

Such a singularity would imply a zero mass bound state coupled to the current, and can be shown to be absent in QED to any finite order in perturbation theory.

As a simple application of the Ward identity, we use it to show that the only contribution to charge renormalization comes from corrections to the photon propagator. We define charge by the photon coupling at zero photon momentum. Define Z_1 by

$$\Gamma^\mu(p,p) = \frac{1}{Z_1}\gamma^\mu + O(m + \gamma \cdot p).$$

By the definition of Z_2,

$$m_0 + \gamma \cdot p + \Sigma(p) = \frac{1}{Z_2}(m + \gamma \cdot p) + O(m + \gamma \cdot p)^2.$$

Then the Ward identity implies $Z_1 = Z_2$. Clearly the renormalized (physical) charge is related to the bare one by

$$Q = \frac{Z_1}{Z_2}\sqrt{Z_3}Q_0 = \sqrt{Z_3}Q_0$$

by the Ward identity.

Recall also that the Ward identity for diagrams with only photon external lines required that the vacuum polarization tensor was transverse:

$$\Pi_{\mu\nu}(q) = -(q^2\eta_{\mu\nu} - q_\mu q_\nu)\Pi(q^2).$$

Π is defined so that the sum of all 1PIR 2 photon diagrams is given by $iQ_0^2\Pi_{\mu\nu}(q)$. Thus the photon two point function is given by

$$\langle A_\mu(x)A_\nu(0)\rangle = \int \frac{d^4q}{(2\pi)^4} e^{iq\cdot x}\left[\frac{-i\alpha q_\mu q_\nu}{q^4} + \frac{-i}{q^2}\left(\eta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2}\right)\frac{1}{1+Q_0^2\Pi(q^2)}\right].$$

We note two consequences. (1) The longitudinal part of the two point function is not corrected. (2) The photon mass remains zero unless $\Pi(q^2)$ has a pole at $q^2 = 0$. We can read off the value of $Z_3 = 1/(1+Q_0^2\Pi(0))$.

In our calculation of vacuum polarization, we found that ultraviolet divergences spoiled the gauge invariance of the calculation. In this chapter we look into this question more closely. Our discussion generalizes that of K. Johnson, Brandeis Lectures, 1964.

The focus will be on the vacuum persistence amplitude of a charged field in the presence of a fixed external gauge field. This quantity is a basic ingredient of the path history quantization of gauge fields in interaction with charged fields, and as such must be gauge invariant for consistency. As long as the gauge field path integral is amenable to dimensional regularization, difficulties with gauge invariance must, in fact reside in the charged field sector which is the focus of this chapter.

Let us first sketch the important consequences of gauge invariance. Of course, gauge invariance demands that, in addition to the persistence amplitude itself, the outin matrix element of every gauge invariant observable should likewise be invariant under gauge transformations. For QED (the abelian case), the current operator $j^\mu(x)$ is one such observable. In the nonabelian case the current operator $j^\mu_a(x)$ is covariant rather than invariant under gauge transformations since gauge transformations act on the index a.

If we return to the variational equation for $\langle out|in\rangle$

$$\delta\langle out|in\rangle = i\int d^4x\delta A_\mu(x)\langle out|\, j^\mu(x)\,|in\rangle$$

we see that gauge invariance of $\langle out|in\rangle$ under infinitesimal gauge transformations that vanish sufficiently rapidly at infinity is equivalent to current conservation $\partial_\mu j^\mu = 0$ in the abelian case. (In the nonabelian case the requirement is $D_\mu j^\mu = 0$.) More is in fact true: since the conservation laws hold for arbitrary A, all terms with the same power of A must cancel among themselves.

We first try to understand how ultraviolet divergences can spoil current conservation. Consider the representation of the current matrix element in terms of the Green function

$$\frac{\langle out|\, j^\mu(x)\,|in\rangle}{\langle out|in\rangle} = -q\,\mathrm{Tr}[\gamma^\mu S_F(x,x;A)].$$

The two arguments of the Green function are coincident, but it is apparent from the free Dirac propagator that $S_F(x-y)\sim 1/(x-y)^3$ at short distances. The fact that $A\neq 0$ is not going to alter this fact. To exert a bit more caution, we should consider the Green function $S_F(x,y;A)$ at slightly separated points and study what happens as $y\to x$. The first thing to notice is that S_F is not gauge invariant because the Dirac field $\psi\to e^{iq\Lambda(x)}\psi$ under gauge transformations*, so

$$S_F(x,y;A)\to e^{iq[\Lambda(x)-\Lambda(y)]}S_F(x,y;A).$$

Superficially, one would think that the above expression for the current matrix element would be gauge invariant because the prefactor formally goes to 1 as $y\to x$. But the approach to unity is only linear so the

* In the nonabelian case the Dirac field transforms as $\psi\to\Omega(x)\psi$, where $\Omega(x)$ is an element of the nonabelian gauge group. Then the following equation is replaced by

$$S_F(x,y;A)\to\Omega(x)S_F(x,y;A)\Omega^\dagger(y).$$

As $y\to x$ this approaches a similarity transformation, but only linearly, so divergences spoil this property.

fact that S_F has a cubic short distance singularity means that gauge noninvariance can occur with quadratic or lower singularities. Thus point splitting alone will not yield a gauge invariant definition of the current matrix element let alone the vacuum persistence amplitude.

To obtain a gauge invariant definition of the current, we first notice that

$$\int_{x,C}^{y} d\xi^\mu \partial_\mu \Lambda(\xi) = \Lambda(y) - \Lambda(x).$$

Thus

$$\bar{S}_F(x,y;A,C) \equiv e^{iq\int_{x,C}^{y} d\xi^\mu A_\mu(\xi)} S_F(x,y;A)$$

is gauge invariant, albeit path dependent†. We can attempt to define the current matrix element as the coincident point limit of $\bar{S}_F$ averaged in some suitable way over C. The result will certainly be gauge invariant, and with a bit of luck might also be conserved. If so we can then define $\langle out|in\rangle$ from this construction through the variational equation.

If we consider the weak field expansion for $S_F(x,y;A)$, the divergence gets one degree less severe with each extra factor of A. Thus the field independent piece has a cubic divergence, the linear term a quadratic divergence, the quadratic term a linear divergence and the cubic term only a logarithmic divergence. Thereafter all the terms are finite. Since the modification to S in $\bar{S}$ vanishes linearly, one only needs to keep terms up to order A^3 in the modification factor. Actually, charge conjugation invariance makes alternate terms vanish. The electromagnetic current $\langle out|\, j^\mu \,|in\rangle$ should have only odd powers of A, whereas the axial current $j_5^\mu \equiv \bar{\psi}\gamma_5\gamma^\mu\psi$ should have only even powers of A in its *outin* matrix element‡. Thus we tentatively

---

† In the nonabelian case, we can construct the matrix $P\exp\{ig\int_{x,C}^{y} d\xi^\mu A_\mu(\xi)\}$, where the P denotes path ordering: matrices associated with "later" points on the path C always stand to the left of "earlier" ones. This matrix transforms under gauge transformations as

$$P\exp\{ig\int_{x,C}^{y} d\xi^\mu A_\mu(\xi)\} \to \Omega(y) P\exp\{ig\int_{x,C}^{y} d\xi^\mu A_\mu(\xi)\}\Omega^\dagger(x).$$

Then

$$\bar{S}_F(x,y;A,C) \equiv Pe^{ig\int_{x,C}^{y} d\xi^\mu A_\mu(\xi)} S_F(x,y;A)$$

transforms by the similarity transformation $\bar{S}_F(x,y;A,C) \to \Omega(y)\bar{S}_F(x,y;A,C)\Omega^\dagger(y)$ under gauge transformations

‡ The currents in a nonabelian gauge theory carry a gauge symmetry index, $j_a^\mu = \bar{\psi}\lambda_a\gamma^\mu\psi$, and $j_{5a}^\mu = \bar{\psi}\lambda_a\gamma_5\gamma^\mu\psi$, where λ_a is the generator of the gauge group in the representation carried by ψ. Their *outin* matrix elements can be defined by

$$J_a^\mu(x) = -\lim_{\epsilon\to 0} \mathrm{Tr}\,\lambda_a\gamma^\mu\bar{S}(x-\epsilon/2, x+\epsilon/2; A)$$
$$J_{5a}^\mu(x) = -\lim_{\epsilon\to 0} \mathrm{Tr}\,\lambda_a\gamma_5\gamma^\mu\bar{S}(x-\epsilon/2, x+\epsilon/2; A)$$

define

$$
\begin{aligned}
J^\mu(x) \equiv & \frac{\langle out|\, j^\mu(x)\, |in\rangle}{\langle out|in\rangle} \\
= & -q \lim_{\epsilon\to 0} \frac{1}{2} \operatorname{Tr}\{\gamma^\mu \bar{S}(x-\epsilon/2, x+\epsilon/2; A) - \gamma^\mu \bar{S}(x-\epsilon/2, x+\epsilon/2; -A)\} \\
J_5^\mu(x) \equiv & \frac{\langle out|\, j_5^\mu(x)\, |in\rangle}{\langle out|in\rangle} \\
= & -\lim_{\epsilon\to 0} \frac{1}{2} \operatorname{Tr}\{\gamma_5\gamma^\mu \bar{S}(x-\epsilon/2, x+\epsilon/2; A) + \gamma_5\gamma^\mu \bar{S}(x-\epsilon/2, x+\epsilon/2; -A)\}
\end{aligned}
$$

To check current conservation we need the Green function equations

$$
\begin{aligned}
(\frac{1}{i}\gamma\cdot\partial_x + m - q\gamma\cdot A(x))S_F(x,y;A) = & S_F(x,y;A)(-\frac{1}{i}\gamma\cdot\overleftarrow{\partial}_y + m - q\gamma\cdot A(y)) = \\
& -i\delta(x-y),
\end{aligned}
$$

from which we get[§]

$$
\begin{aligned}
\partial_\mu \operatorname{Tr}[\gamma^\mu S(x-\epsilon/2, x+\epsilon/2; A)] = & \\
& iq \operatorname{Tr}[(A_\mu(x-\epsilon/2) - A_\mu(x+\epsilon/2))\gamma^\mu S(x-\epsilon/2, x+\epsilon/2; A)].
\end{aligned}
$$

To calculate the gradient of the modification factor, we must first specify what happens to the integration contour under differentiation. Under the change $x \to x + \delta x$, we globally translate C parallel to itself by the same amount. We use Stokes theorem in the form

$$
\oint d\xi^\mu A_\mu(\xi) = \int d\sigma^{\mu\nu} F_{\mu\nu},
$$

where the integral on the r.h.s. is over a surface spanning the closed curve implicit in the line integral on the l.h.s[¶]. To be completely explicit about the conventions in this identity parametrize the surface by $\xi^\mu(\sigma,\tau)$

---

§ In the nonabelian case the gauge fields are of course matrices and the appropriate equation is

$$
\begin{aligned}
\partial_\mu \operatorname{Tr}[\lambda_a\gamma^\mu S(x-\epsilon/2, x+\epsilon/2; A)] = & \\
& iq \operatorname{Tr}[\lambda_a(A_\mu(x-\epsilon/2) - A_\mu(x+\epsilon/2)\lambda_a)\gamma^\mu S(x-\epsilon/2, x+\epsilon/2; A)].
\end{aligned}
$$

Then the coincident point limit gives formal *covariant* conservation rather than ordinary conservation.

¶ A gauge covariant form of Stokes theorem can be given in the nonabelian case for an *infinitesimal* closed loop

$$
Pe^{ig\oint d\xi^\mu A_\mu(\xi)} \approx I + ig\int d\sigma^{\mu\nu} F_{\mu\nu}, \qquad \text{infinitesimal } \textit{loop},
$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]$ is the nonabelian field strength. (The quadratic term in F comes from $(ig)^2 \oint d\xi^\mu \oint d\xi'^\nu P[A_\mu(\xi)A_\nu(\xi')]$ which for an infinitesimal loop at z over which A is constant is just $(ig)^2 A_\mu(z)A_\nu(z)\oint d\xi^\mu \xi^\nu = -(ig)^2[A_\mu, A_\nu]\delta\sigma^{\mu\nu}$.) Another version of this is a statement of how the path ordered phase changes under an infinitesimal deformation of the curve C spanning an infinitesimal surface element $\delta\sigma^{\mu\nu}$ at the point z:

$$
\delta_z Pe^{ig\int d\xi^\mu A_\mu(\xi)} = igP[\delta\sigma^{\mu\nu} F_{\mu\nu}(z) e^{ig\int d\xi^\mu A_\mu(\xi)}]
$$

where it is understood that $F(z)$ is included in the path ordering.

with σ and τ each ranging from 0 to 1, with σ labelling the abscissa and τ the ordinate. Then

$$d\sigma^{\mu\nu} = \frac{1}{2} d\sigma d\tau \left(\frac{\partial \xi^\mu}{\partial \sigma} \frac{\partial \xi^\nu}{\partial \tau} - \frac{\partial \xi^\nu}{\partial \sigma} \frac{\partial \xi^\mu}{\partial \tau} \right), \tag{16.1}$$

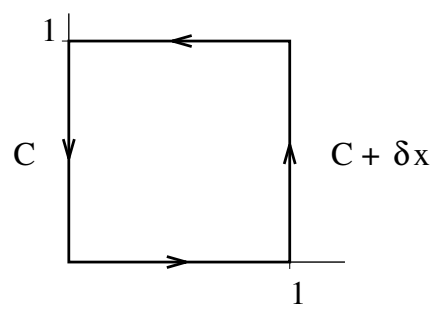
and with these parameters, the line integral runs around the boundary of the unit square of parameter space in a counter-clockwise direction. The factor in parentheses in (16.1) is of course a total divergence:

$$\begin{aligned} 2\left(\frac{\partial \xi^\mu}{\partial \sigma} \frac{\partial \xi^\nu}{\partial \tau} - \frac{\partial \xi^\nu}{\partial \sigma} \frac{\partial \xi^\mu}{\partial \tau} \right) = & \frac{\partial}{\partial \sigma} \left(\xi^\mu \frac{\partial \xi^\nu}{\partial \tau} - \xi^\nu \frac{\partial \xi^\mu}{\partial \tau} \right) \\ & + \frac{\partial}{\partial \tau} \left(\xi^\nu \frac{\partial \xi^\mu}{\partial \sigma} - \xi^\mu \frac{\partial \xi^\nu}{\partial \sigma} \right), \end{aligned}$$

which leads to the useful identity

$$\int d\sigma^{\mu\nu} = \frac{1}{2} \oint \xi^\mu d\xi^\nu = -\frac{1}{2} \oint \xi^\nu d\xi^\mu, \tag{16.2}$$

with the contour for the r.h.s. enclosing the surface on the l.h.s. in a counterclockwise sense.



To define the derivative of the modification factor, take the left vertical boundary of parameter space to map onto the curve C, the right vertical boundary to map onto $C + \delta x$, the curve rigidly translated by the amount δx. The top and bottom boundaries then give the displacements of the end points of the curve $x + \epsilon/2$ and $x - \epsilon/2$ respectively. Using Stokes theorem on this closed contour then gives*

$$\partial_\mu \int_{C_x} d\xi \cdot A = A_\mu(x + \epsilon/2) - A_\mu(x - \epsilon/2) + \int_C d\xi^\nu F_{\mu\nu}(\xi).$$

---

* For the nonabelian case we get

$$\partial_\mu P e^{ig \int d\xi^\mu A_\mu} = ig(A_\mu(x + \epsilon/2) P e^{ig \int d\xi^\mu A_\mu} - P e^{ig \int d\xi^\mu A_\mu} A_\mu(x - \epsilon/2) + P[\int d\xi^\nu F_{\mu\nu}(\xi) e^{ig \int d\xi^\mu A_\mu}]).$$

Combining these results then gives the conservation laws**

$$
\begin{aligned}
\partial_\mu J^\mu(x) =& -iq^2 \lim_{\epsilon\to 0} \frac{1}{2} \int_C d\xi^\nu F_{\mu\nu}(\xi)\, \mathrm{Tr}\{\gamma^\mu(\bar{S}(x-\epsilon/2, x+\epsilon/2; A) \\
&+ \bar{S}(x-\epsilon/2, x+\epsilon/2; -A))\} \\
\partial_\mu J_5^\mu(x) =& im \lim_{\epsilon\to 0} \mathrm{Tr}\,\gamma_5(\bar{S}(x-\epsilon/2, x+\epsilon/2; A) + \bar{S}(x-\epsilon/2, x+\epsilon/2; -A)) \\
&- iq \lim_{\epsilon\to 0} \frac{1}{2} \int_C d\xi^\nu F_{\mu\nu}(\xi)\, \mathrm{Tr}\{\gamma_5\gamma^\mu(\bar{S}(x-\epsilon/2, x+\epsilon/2; A) \\
&- \bar{S}(x-\epsilon/2, x+\epsilon/2; -A))\}
\end{aligned}
\tag{16.3}
$$

Now we consider whether we can choose the contour to give conserved currents. We need to consider the behavior of the factors multiplying $\int F$. Since the latter is of order ϵ only the singular parts of these factors need be retained. First consider $\mathrm{Tr}\,\gamma^\mu \bar{S}_F$. Since it is gauge invariant it must be a vector formed from ϵ^μ, $F_{\mu\nu}$ and derivatives of F. The field independent term must be of the form $\epsilon^\mu f(\epsilon^2)$ where f behaves as $1/\epsilon^4$. Because of the antisymmetry of F, it will drop out if we specify the integration contour to be a straight line connecting the two endpoints ($\epsilon^\mu\epsilon^\nu F_{\mu\nu} = 0$). The term linear in A was *a priori* quadratically divergent, but in $\bar{S}$ this is reduced by one power of ϵ because a factor of momentum must be provided to form F out of A. Technically the quadratic and linear divergences in the unadjusted vacuum polarization $T^{\mu\nu}(k)$ turn out to be proportional to $(1 - i\epsilon\cdot k)\ \epsilon^\mu\epsilon^\nu/\epsilon^4$ in point splitting regulariation and therefore give zero when contracted with $F_{\mu\nu}\epsilon^\nu$. Thus the linear term in the "vector" current is at worst logarithmically divergent.

The quadratic term in A is *a priori* linearly divergent but gauge invariance in the abelian case requires two factors of momentum to be used in forming F^2 so it and all higher powers must be finite. In the nonabelian case a quadratic term with no derivatives could conceivably be needed to complete the nonabelian field strength that might appear in the linear term in A. But as we note below, the divergence in the "vector" linear term is only logarithmic and so does not contribute. The divergence in the "axial" linear term is linear and then the quadratic term provides the rest of the nonabelian field strength.

In summary, provided we take a straight line contour, the only term of relevance in the abelian case is the linear one. It doesn't contribute in the divergence of the electromagnetic current because of odd charge conjugation. Even if we don't explicitly enforce charge conjugation by making J manifestly odd in A, we know from our vacuum polarization calculation that the linear term is only logarithmically divergent since the adjustments for gauge invariance are included when we use $\bar{S}$. With point splitting regularization the linear divergence can be present, but its coefficient vanishes for "vector" currents. This is important in the

** We quote here the nonabelian results:

$$
\begin{aligned}
\partial_\mu J_a^\mu(x) =& -ig \lim_{\epsilon\to 0} \mathrm{Tr}[\lambda_a, A_\mu(x+\epsilon/2)]\gamma^\mu \bar{S}(x-\epsilon/2, x+\epsilon/2; A) \\
&- ig \lim_{\epsilon\to 0} \mathrm{Tr}\,\lambda_a\gamma^\mu \int_C d\xi^\nu P[F_{\mu\nu}(\xi) e^{ig\int d\xi^\mu A_\mu}] S(x-\epsilon/2, x+\epsilon/2; A) \\
\partial_\mu J_{5a}^\mu(x) =& 2im \lim_{\epsilon\to 0} \mathrm{Tr}\,\lambda_a\gamma_5 \bar{S}(x-\epsilon/2, x+\epsilon/2; A) \\
&- ig \lim_{\epsilon\to 0} \mathrm{Tr}[\lambda_a, A_\mu(x+\epsilon/2)]\gamma_5\gamma^\mu \bar{S}(x-\epsilon/2, x+\epsilon/2; A) \\
&- ig \lim_{\epsilon\to 0} \mathrm{Tr}\,\lambda_a\gamma_5\gamma^\mu \int_C d\xi^\nu P[F_{\mu\nu}(\xi) e^{ig\int d\xi^\mu A_\mu}] S(x-\epsilon/2, x+\epsilon/2; A)
\end{aligned}
$$

nonabelian case because then the charge conjugation properties are no longer so simple. It is sufficient to know that the term is only log divergent to prove that it won't contribute to the "vector" conservation law. Thus the e.m. current and, more generally, the "vector" nonabelian currents, defined through $\bar{S}$ with a straightline contour, are conserved after $\epsilon \to 0$, and $\bar{S}$ therefore determines a gauge invariant persistence amplitude.

The same cannot be said about the axial current which is even under charge conjugation. In that case the linear term of $\mathrm{Tr}\,\gamma_5\gamma^\mu\bar{S}$ does contribute a linearly divergent factor to cancel the factor of ϵ implicit in $\int F$, leaving a finite anomalous contribution. This is the celebrated axial anomaly to which we shall return below.

We have shown that choice of a straight line contour suffices to make $\bar{S}_F$ yield a conserved e.m. current. However a direction ϵ^μ is singled out breaking Lorentz invariance. Thus, it is convenient to average over all directions. This averaging procedure preserves gauge invariance and current conservation, since those features hold for each fixed direction. Since all terms of order A^4 and higher are independent of ϵ as $\epsilon \to 0$ we only need to apply this averaging procedure for the linear and cubic terms. (The constant and quadratic terms vanish by charge conjugation (Furry's theorem).) The linear term is the one relevant to vacuum polarization, so let's look at that one in detail.

The linear terms that come from the modification factor

$$iq \int_{x-\epsilon/2}^{x+\epsilon/2} d\xi^\mu A_\mu(\xi) \approx iq\epsilon\cdot A + \frac{iq}{24}(\epsilon\cdot\partial)^2\epsilon\cdot A$$

are multiplied by

$$\begin{aligned}\mathrm{Tr}[\gamma^\mu S_F(\epsilon)] =& -4i\int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot\epsilon}\frac{p^\mu}{p^2+m^2-i\epsilon}\\ =&4\frac{\partial}{\partial\epsilon}\int\frac{d^4p}{(2\pi)^4}\frac{e^{-ip\cdot\epsilon}}{p^2+m^2-i\epsilon} \equiv \frac{\epsilon^\mu}{\epsilon^4}f(m^2\epsilon^2),\end{aligned}$$

where f is regular and nonvanishing at zero argument. In the average over directions of ϵ the following replacements take place

$$\begin{aligned}\epsilon^\mu\epsilon^\nu \to& \frac{1}{4}\eta^{\mu\nu}\epsilon^2\\ \epsilon^\mu\epsilon^\nu\epsilon^\rho\epsilon^\sigma \to& \frac{\epsilon^4}{24}(\eta^{\mu\nu}\eta^{\rho\sigma}+\eta^{\mu\rho}\eta^{\nu\sigma}+\eta^{\nu\rho}\eta^{\mu\sigma})\end{aligned}$$

Thus

$$\left\langle iq\int_{x-\epsilon/2}^{x+\epsilon/2} d\xi^\mu A_\mu(\xi)\,\mathrm{Tr}\,\gamma^\mu S_F\right\rangle \approx iq\frac{f(m^2\epsilon^2)}{\epsilon^4}\left(\frac{\epsilon^2}{4}A^\mu + \frac{\epsilon^4}{24^2}(\partial^2A^\mu + 2\partial^\mu\partial\cdot A)\right).$$

These terms represent the adjustment that should be made to the linear term in A (whose coefficient is related to the vacuum polarization $T^{\mu\nu}(k)$). We see that they have the qualitative appearance of the adjustment we actually had to make in our original calculation: a quadratically divergent constant times $\eta^{\mu\nu}$ and a finite second order polynomial in k. The detailed coefficients cannot be compared because the cutoff procedure was different in that calculation.

The term of order A^2 is absent in J^μ because of charge conjugation (Furry's theorem). Finally there will be a finite adjustment to the term of order A^3 that arises from the order k term in the modification factor times the order $3-k$ term in $S_F(A)$ for $k = 1, 2, 3$. This modification enters the fourth order term in $\langle out|in\rangle$.

The unmodified value of this term turns out to be uv finite but not gauge invariant. In summary, we have seen that a careful gauge invariant definition of the e.m. current has led to modifications in the calculation of only the first few terms in the weak field perturbation series for $\langle out|in\rangle$, specifically the quadratic and order four terms. All higher terms are gauge invariant without modification.

Gauge Invariant Regulation Procedures For practical calculations in quantum gauge field theories, it is unnecessary to carry out this detailed procedure. It is sufficient to establish a systematic procedure for regulating divergences compatible with gauge invariance. Such a procedure will automatically supply the polynomial modifications needed for gauge invariance. We mention briefly the more popular procedures.

Pauli-Villars Method. This method is particularly suited to Abelian Gauge theories since it relies on the fact that the violations in gauge invariance are independent of the mass of the charged fields. Thus if we introduce extra charged fields of large mass M_i which contribute to divergent diagrams with negative signs we can adjust the coefficients of their contributions to render the loop integrals finite. Then the gauge violating pieces of each contribution will cancel, and the regulated calculation will be gauge invariant. One then lets $M_i \to \infty$ at the end of the calculation. This may still leave uv divergences, but only those compatible with gauge invariance as we found in the vacuum polarization calculation.

Dimensional Regularization. The idea here is that the severity of uv divergences depends on the space-time dimension. In particular, the nature of uv gauge invariance violations is different in each dimension. Thus if we can carry out the calculation in a way that applies to general dimension, the violations of gauge invariance must disappear. This method will obviously not work in theories that can be defined only in particular dimensions. For example the alternating symbol $\epsilon^{\mu_1\cdots\mu_D}$ has a different tensor structure in each dimension. A theory in which it appears in a fundamental way can not be formulated in a general dimension.

To illustrate how dimensional regularization is used, we reconsider the vacuum polarization calculation. Let us first make some general remarks. In D dimensions the charge has units $[mass]^{2-D/2}$. Thus it is convenient to introduce a mass parameter μ to define a dimensionless coupling q_I in general D via

$$q^2 \equiv q_I^2 \mu^{4-D}.$$

The identities for calculating traces of products of gamma matrices carry over to D dimensions except that the constant $4 = \mathrm{Tr}\, I$ is replaced by $2^{D/2}$. The Feynman trick for handling denominators of propagators is unchanged in general dimension, and the shift of integration variables to make the denominators depend only on the squares of the loop momenta is still applicable. But then averaging over directions of the loop momenta gives a D dependent factor: For example $\langle p^\mu p^\nu\rangle_{angles} = \eta^{\mu\nu}p^2/D$. After taking all this into account, the vacuum polarization calculation for general D, becomes

$$T_D^{\mu\nu}(k) = -\,2^{D/2}\frac{q_I^2\mu^{4-D}}{(2\pi)^D}\int_0^1 dx \int d^Dp \frac{\eta^{\mu\nu}[p^2((2/D)-1) - m^2 + x(1-x)k^2] - 2x(1-x)k^\mu k^\nu}{[m^2+p^2+x(1-x)k^2]^2}$$

Clearly we need to be able to do the integral

$$\begin{aligned}\int \frac{d^Dp p^m}{[p^2+A^2]^2} =& A^{D+m-4}\frac{2\pi^{D/2}}{\Gamma(D/2)}\int_0^\infty \frac{p^{D+m-1}dp}{[p^2+1]^2}\\ =& \pi^{D/2}A^{D+m-4}\frac{\Gamma((D+m)/2)\Gamma(2-(D+m)/2)}{\Gamma(D/2)}\end{aligned}$$

where we used the identity

$$\int_0^\infty \frac{p^{m+D-1}dp}{[p^2+1]^2} = \frac{1}{2}\int_0^1 dx x^{1-(D+m)/2}(1-x)^{((D+m)/2)-1}$$
$$= \frac{\Gamma(2-(D+m)/2)\Gamma((D+m)/2)}{2\Gamma(2)}.$$

Note that the quantity $\Omega_D \equiv 2\pi^{D/2}/\Gamma(D/2)$ is just the value of the integral over all angles in D dimensions. The following table lists Ω_D for the $2 \leq D \leq 8$.

Angular Integral for Various Dimensions							
D	2	3	4	5	6	7	8
Ω_D	2π	4π	$2\pi^2$	$8\pi^2/3$	π^3	$16\pi^3/15$	$\pi^4/3$

Using these results to do the integrals, we obtain

$$T_D^{\mu\nu}(k) = \frac{q_I^2}{2\pi^2}\Gamma(2-D/2)\int_0^1 dx x(1-x)(k^\mu k^\nu - k^2\eta^{\mu\nu})\left[\frac{m^2+x(1-x)k^2}{2\pi\mu^2}\right]^{(D-4)/2}.$$

As advertised the result is gauge invariant. The answer is finite as long as $D < 4$, but we see that the gamma function has a pole as $D \to 4$. This is how divergences appear in dimensional regularization. To regain the result for 4 dimensions we have to write $A^{(D-4)/2} \approx 1 + \frac{D-4}{2}\ln A$ and the second term must be retained since $\Gamma(2-D/2)(D-4)/2 \to -1$ as $D \to 4$:

$$T_D^{\mu\nu}(k) \to \frac{q_I^2}{2\pi^2}(k^\mu k^\nu - k^2\eta^{\mu\nu})\int_0^1 dx x(1-x)\left[\Gamma(2-D/2) - \ln\frac{m^2+x(1-x)k^2}{2\pi\mu^2}\right].$$

The pole at $D = 4$ represents an infinity which has after renormalization the same fate as the cutoff dependence did in our earlier calculation, namely it disappears after expressing measurable results in terms of measured parameters.

Since the pole at $D = 4$ corresponds to the logarithmic divergence of a direct cutoff procedure, it is useful to establish the relation between the residue of the pole and the coefficient of $\ln(\Lambda^2)$. This follows from the simple integral

$$\int_\mu^\Lambda dp p^{D-5} = \frac{\Lambda^{D-4}-\mu^{D-4}}{D-4}$$
$$\sim \begin{cases} -\frac{1}{D-4} & \Lambda \to \infty, D < 4 \\ \ln\frac{\Lambda}{\mu} & D \to 4, \Lambda \text{ fixed}, \end{cases}$$

from which we see that the coefficient of $\ln(\Lambda^2)$ is $-$(residue of pole)/2. This is of course in agreement with our two calculations of vacuum polarization.

The existence of a gauge invariant regularization scheme such as dimensional regularization assures the absence of anomalies (violations) of gauge invariance. Our careful analysis of the e.m. current showed that it is indeed possible to define $\langle out|in\rangle$ in a gauge invariant way for the case of the electromagnetic field. But our parallel discussion of the axial current $j_5^\mu = [\bar\psi, \gamma_5\gamma^\mu\psi]$ showed that it is not conserved for $m = 0$, contrary to what the c-number Dirac equation would lead us to believe. The significance of this is that whereas we can consistently couple the quantum Dirac field to electromagnetism through $j^\mu A_\mu$, the gauge coupling $j_5^\mu A_\mu$ violates gauge invariance, and would lead to inconsistencies. Our point splitting definition of the currents shows the presence of the anomaly and allows us to compute it. One can also understand why the popular regularization schemes I have described fail to forbid anomalies. The Pauli-Villars scheme requires the addition of massive fermions which explicitly violate chiral invariance, and dimensional regularization gives no method for defining $\epsilon^{\mu\nu\rho\sigma}$ or γ_5.

The possibility of anomalies in axial gauge couplings puts constraints on viable theories of the weak interactions which violate parity conservation through just such couplings. The way parity violation enters the standard electroweak theory is by assigning left and right handed fermions to different representations of the electro-weak gauge group $SU(2) \times U(1)$. This is of course possible only if explicit mass terms are not included in the Hamiltonian. Thus the $I \pm \gamma_5$ projections of the Dirac field for each fermion couple in different ways to the gauge fields. What our discussion shows is that such a scheme would be inconsistent for a single fermion. The way the electroweak theory escapes this difficulty is by a cancellation of the anomalies between the contributions of different fermions.

Let us return to the axial current conservation law (16.3).

$$\partial_\mu J_5^\mu(x) = -2mJ_5 - iq\lim_{\epsilon\to 0}\frac{1}{2}\int_C d\xi^\nu F_{\mu\nu}(\xi)$$
$$\mathrm{Tr}\{\gamma_5\gamma^\mu(\bar S(x-\epsilon/2, x+\epsilon/2; A) - \bar S(x-\epsilon/2, x+\epsilon/2; -A))\},$$

where we have defined J_5 as the suitably regularized version of

$$\langle out|\,[\bar\psi, i\gamma_5\psi]\,|in\rangle \,/\, \langle out|in\rangle$$

given by the first term on the r.h.s. of (16.3). We would like to extract the explicit contribution of the anomaly which arises from the linearly divergent term in $\mathrm{Tr}\,\gamma_5\gamma^\mu\bar S(x-\epsilon/2, x+\epsilon/2; A)$, which resides (in the abelian case) in the term with only one power of A. When we expand $S(x-\epsilon/2, x+\epsilon/2; A)$ in A, the order zero term vanishes because one can't form an axial vector from the only available four vector ϵ^μ. The linear term in A would appear to be quadratically divergent, but a momentum factor must be used along with ϵ^μ to form a pseudo two index tensor, so the divergence is only linear. The quadratic term in A is also linearly divergent, but in the case of an abelian gauge field Bose symmetry kills this leading linear divergence. (We have already noted that since the linear divergence in this term would not involve derivatives of A it would be inconsistent with abelian gauge invariance. It does give a contribution to the anomaly in the nonabelian case, where it is needed to complete the nonabelian field strength whose derivatve terms come from the linear term. Since we only need keep the linearly divergent term, the modification factor can be dropped. Thus, for the abelian case we only need to extract the linearly divergent piece of

$$\begin{aligned}
&\mathrm{Tr}\,\gamma_5\gamma^\mu S(x-\epsilon/2, x+\epsilon/2; A)\\
&\quad\approx iq\int d^4y\,\mathrm{Tr}\,\gamma_5\gamma^\mu S_F(x-y-\epsilon/2)\gamma\cdot A(y)S_F(y-x-\epsilon/2)\\
&\quad\approx -iq\int\frac{d^4k}{(2\pi)^4}\tilde A(k)_\nu e^{i(x+\epsilon/2)\cdot k}\int\frac{d^4p}{(2\pi)^4}e^{-i\epsilon\cdot p}\frac{\mathrm{Tr}\,\gamma_5\gamma^\mu(m-\gamma\cdot p)\gamma^\nu(m-\gamma\cdot(p-k))}{(m^2+p^2-i\epsilon)(m^2+(p-k)^2-i\epsilon)}.
\end{aligned}$$

The trace of γ_5 times fewer than 4 gamma matrices vanishes and

$$\mathrm{Tr}\,\gamma_5\gamma^\mu\gamma^\rho\gamma^\nu\gamma^\sigma = -4i\epsilon^{\mu\rho\nu\sigma}.$$

Thus the trace in the numerator gives simply $+4i\epsilon^{\mu\rho\nu\sigma}p_\rho k_\sigma$. Furthermore, the leading divergence as $\epsilon \to 0$ coming from the integral over p is independent of k and m so the latter can be set to zero, and we only need to evaluate

$$\begin{aligned}
\int \frac{d^4p}{(2\pi)^4}\frac{p_\rho}{(p^2-i\epsilon)^2}e^{-i\epsilon\cdot p} =& \frac{-i\epsilon_\rho}{2}\int\frac{d^4p}{(2\pi)^4}\frac{1}{p^2-i\epsilon}\\
=&\frac{\epsilon_\rho}{2}\int\frac{d^4p_E}{(2\pi)^4}\frac{1}{p^2}\\
=&\frac{\epsilon_\rho}{2}\int\frac{d^4p_E}{(2\pi)^4}\int_0^\infty dT e^{-Tp^2-i\epsilon\cdot p}\\
=&\frac{\epsilon_\rho}{32\pi^4}\int_0^\infty dT\frac{\pi^2}{T^2}e^{-\epsilon^2/4T}\\
=&\frac{\epsilon_\rho}{8\pi^2\epsilon^2}
\end{aligned}$$

where we did a Wick rotation in the second line and used a simple representation for $1/p^2$ in the third line.

Collecting these results we obtain

$$\begin{aligned}
&\mathrm{Tr}\,\gamma_5\gamma^\mu S(x-\epsilon/2, x+\epsilon/2; A)\\
&\quad\sim 4q\int\frac{d^4k}{(2\pi)^4}\tilde{A}(k)_\nu e^{ix\cdot k}\epsilon^{\mu\rho\nu\sigma}k_\sigma\frac{\epsilon_\rho}{8\pi^2\epsilon^2}\\
&\quad= -iq\epsilon^{\mu\rho\nu\sigma}\partial_\sigma A_\nu(x)\frac{\epsilon_\rho}{2\pi^2\epsilon^2}\\
&\quad= iq\epsilon^{\mu\rho\nu\sigma}F_{\nu\sigma}(x)\frac{\epsilon_\rho}{4\pi^2\epsilon^2}
\end{aligned}$$

This can now be substituted into our expression for $\partial_\mu J^\mu(x)$, which gives after averaging over directions of ϵ

$$\begin{aligned}
\partial_\mu J_5^\mu(x) =& -2mJ_5 + \frac{q^2}{16\pi^2}\epsilon^{\mu\rho\nu\sigma}F_{\nu\sigma}(x)F_{\mu\rho}(x)\\
=& -2mJ_5 + \frac{\alpha_0}{4\pi}\epsilon^{\mu\rho\nu\sigma}F_{\mu\rho}(x)F_{\nu\sigma}(x)
\end{aligned}$$

More generally it is clear from our general remarks that these results generalize in the nonabelian case to

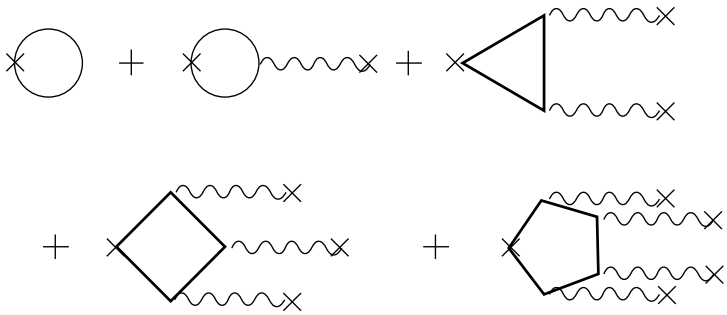
$$D_\mu J_{5a}^\mu(x) = -2mJ_{5a} + \frac{g^2}{16\pi^2}\epsilon^{\mu\rho\nu\sigma}\,\mathrm{Tr}[\lambda_a F_{\mu\rho}(x)F_{\nu\sigma}(x)].$$

<u>Ambiguities</u>

Having obtained the anomalous divergence law for a gauge invariant current, we now examine other definitions of the current useful in certain contexts. First consider the expansion of an unmodified current

in powers of the gauge potential described by the series of diagrams.

$$\frac{\langle out|\, j^\mu \,|in\rangle}{\langle out|in\rangle} =$$



which have degrees of divergence $3, 2, 1, 0, -1$ respectively.

Note that the terms that are involved in the anomaly are parity violating containing an odd number of γ_5's and start at 3rd order. Considering the l.h.s. of the divergence law order by order in A, we see that it is a sum of the divergence of the nth order term and, in the nonabelian case, terms with A multiplying the $n-1$th order term. Each of these diagrams naturally has a cyclic symmetry in the labels a, μ, x of the n vertices. Our gauge invariant construction of the current evidently does not reflect this symmetry in the terms with an odd number of γ_5's, because the anomaly is present in only one of the vertices. This difference reflects a fundamental ambiguity in the definition of divergent diagrams, which in four dimensions includes only those of order ≤ 4. In momentum space, a diagram with degree of divergence $\mathcal{D}$ is ambiguous up to the addition of a polynomial in the external momenta of order $\mathcal{D}$. Accordingly the parity violating ambiguities are in the triangle and square diagrams and are $\epsilon^{\mu\nu\rho\sigma}(\alpha p_1 + \beta p_2)_\rho$ and $\alpha\epsilon^{\mu\nu\rho\sigma}$ respectively. Our construction, with the anomlies absent from all but one vertex, is related to the cyclic symmetric definition by the addition of such polynomials. A potentially confusing point is that the pentagon diagram is finite and unambiguous but the r.h.s. of the divergence law contains terms quartic in A. This is explained by the fact that the l.h.s. contains a contribution from the square diagram. Changing the square diagram by a term proportional to $\alpha\epsilon^{\mu\nu\rho\sigma}$ adds a term of exactly the structure of the quartic term on the r.h.s. of the divergence law. Thus depending on the resolution of the ambiguities in the triangle and square, there may or may not be a quartic term on the r.h.s. Similarly the cubic term on the r.h.s. is influenced by the ambiguity resolutions in both the square and triangle diagram.

Physical Consequences of the Chiral Anomaly

The existence of the chiral anomaly has two sorts of ramifications. The more fundamental is the constraints it puts on the sorts of gauge fields that can be consistently coupled to fermions. But even if the chiral current is not coupled to a gauge field, it is still an observable of the theory, which would be a conserved current for massless fermions in the absence of the anomaly. The anomaly breaks this conservation law in a way that becomes experimentally significant for very light fermions. The classic example is the decay $\pi_0 \to \gamma + \gamma$ which would be oversuppressed by the small up and down quark masses were it not for the anomaly.

We first consider the limitations imposed on gauge couplings. The gauge fields mediating the weak interactions must couple differently to left and right handed fermions

$$\psi_R \equiv \frac{I+\gamma_5}{2}\psi \qquad\qquad \psi_L \equiv \frac{I-\gamma_5}{2}\psi.$$

This is an experimental necessity. The standard electroweak theory, based on the nonabelian gauge group $SU(2) \times U(1)$, achieves this requirement by assigning left handed fermions to doubets under $SU(2)$ while

the right-handed fermions are singlets under $SU(2)$. The two types of fermion have different nonzero weak hypercharges y under the $U(1)$. The ordinary electric charge is related to y and the weak isospin I_3, one of the three generators of the $SU(2)$, according to the formula

$$Q = I_3 + \frac{y}{2}.$$

For example, the first generation of fermions consists of the electron e, the electron neutrino ν_e the up quark u and the down quark d. The neutrino and left handed electron form a doublet $l_{1L} = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}$. Since the neutrino is neutral with $I_3 = 1/2$ and the electron has charge -1 with $I_3 = -1/2$, we have $y_l = -1$. In the standard model there is no right-handed neutrino and the right handed electron, being an $SU(2)$ singlet has $I_3 = 0$ and hence $y_{e_R} = -2$. The up and down quarks have charge $+2/3$ and $-1/3$ respectively. Their lefthanded components are an $SU(2)$ doublet with $I_3 = +1/2, -1/2$ respectively, and accordingly carry weak hypercharge $y_{qL} = +1/3$. Their right handed components are singlets and hence have $y_{u_R} = +4/3$ and $y_{d_R} = -2/3$. There are at least two more generations which seem to repeat the pattern of the first only differing in masses, which of course cannot arise from explicit mass terms which would violate gauge invariance.

Now we consider the limitations imposed by the anomaly. Since the couplings to left and right handed fermions are different the currents that must be conserved for gauge invariance are $\bar{\psi}\lambda_a(I \pm \gamma_5)\gamma_\mu\psi$, separately for $\lambda_a = I$. When $\lambda_a = \tau_a \in SU(2)$, only the left handed current couples to the gauge fields, and that's the one that must be conserved. When defined in terms of the Green's functions for the Dirac equation, these are just $J_a^\mu \pm J_{5a}^\mu$. Since the gauge couplings conserve handedness for massless fermions, it is not necessary to include the $(I \pm \gamma_5)/2$ in the coupling of each gauge field. The anomaly is thus proportional to $\mathrm{Tr}[\lambda_a \epsilon^{\mu\rho\nu\sigma} F_{\mu\rho}(x) F_{\nu\sigma}(x)]$. The field strengths can be expanded in terms of the matrices $F = \sum_a F_a \lambda_a$, so

the vanishing of the anomaly requires*

$$\sum \mathrm{Tr}\, \lambda_a \{\lambda_b, \lambda_c\} = 0$$

where the sum is over the contribution of all fermions coupling to the gauge field under examination.

We first notice that when a, b, c all refer to the $SU(2)$ matrices the contribution vanishes: $\mathrm{Tr}\, \tau_a \{\tau_b, \tau_c\} = 2\delta_{bc}\, \mathrm{Tr}\, \tau_a = 0$. When two refer to $SU(2)$ and one to $U(1)$ there is a potential anomaly proportional to $\sum_L y_L\, \mathrm{Tr}\, \tau_a \tau_b = 2\delta_{ab} \sum_L y_L$ where the sum is over all the hypercharges of the left handed doublets. The contribution where only one index refers to $SU(2)$ is clearly zero and we are left with the case where all indices refer to $U(1)$. Then both left and right handed fermions contribute, but with opposite signs, so this contribution is proportional to

$$\sum_R y_R^3 - \sum_L y_L^3 = 0.$$

It is fortunate that the fermion content of the standard model required by experiment satisfies the constraints on hypercharges we have just obtained. If we substitute the relation between electric charge and the weak hypercharges into the constraints, the first just requires that the charges of all the components of the lefthanded doublets sum to zero. Thus for the first generation this is realized because there are three "colors" for each quark: $-1 + 3(2/3 - 1/3) = 0$. The second generation consisting of the muon, muon neutrino, charmed quark and strange quark, has gauge couplings identical to the first and so the contribution to the anomaly from them also cancels. The third generation, follows the same pattern. the last member, the charged 2/3 top quark, has recently (Spring 1995) been discovered at the TeVatron at Fermilab. There

---

* The cancellation condition is certainly enough for the vanishing of the anomaly in the conservation of the currents constructed to be gauge covariant as we have done. This construction is not quite suited for the derivation of $\langle out|in\rangle$ when anomalies are present. To see this note that in the expansion of J in powers of the gauge field, the coefficient Γ_n of $n-1$ powers of A is an n current amplitude in which the vertex of J has been singled out. Because the anomaly is present only in this vertex the Γ_n is evidently not symmetric in the n currents as the coefficient of n powers of A in the expansion of $\ln \langle out|in\rangle$ must be. Therefore to construct $\langle out|in\rangle$ from its variational equation, one must first symmetrize each Γ_n in the labels of the n currents. Since different powers of A are symmetrized differently, this process destroys the gauge covariance of the current. The term in the anomaly quadratic in A retains its structure but is now distributed equally among the 3 vertices, so is multiplied by a factor of 1/3. The symmetrization process modifies the structure of the higher terms: in particular the quartic term disappears and the anomaly then reads

$$D_\mu J^\mu_{L,Ra}(x) = \pm \frac{g^2}{48\pi^2} \epsilon^{\mu\rho\nu\sigma}\, \mathrm{Tr}[\lambda_a (2\partial_\mu A_\rho(x) \partial_\nu A_\sigma(x) - ig\partial_\mu (A_\rho(x) A_\nu(x) A_\sigma(x)))].$$

The cancellation condition would seem to only insure the vanishing of the first term. The second term would seem to require the additional condition

$$\mathrm{Tr}\, \lambda_a \lambda_{[b} \lambda_c \lambda_{d]} = 0,$$

where the indices enclosed in square brackets are completely antisymmetrized. However this condition is automatically satisfied if the first is:

$$\begin{aligned} \lambda_{[b}\lambda_c\lambda_{d]} =& \lambda_b[\lambda_c, \lambda_d] + \lambda_d[\lambda_b, \lambda_c] + \lambda_c[\lambda_d, \lambda_b] \\ =& \frac{1}{2}\{\lambda_b, [\lambda_c, \lambda_d]\} + \frac{1}{2}\{\lambda_d[\lambda_b, \lambda_c]\} + \frac{1}{2}\{\lambda_c[\lambda_d, \lambda_b]\}, \end{aligned}$$

by virtue of the Jacobi identity.

has long been evidence for the τ lepton, its neutrino, and the bottom quark (with charge $-1/3$). We can regard the required cancellation of anomalies as a prediction of the existence of the top quark, which has now been confirmed. The recent Tevatron experiments measure its mass to be 176 ± 13 GeV, almost two orders of magnitude larger than all other quarks and leptons.

We have yet to consider the "cubic" constraint from anomaly cancellation, $\sum_R y_R^3 - \sum_L y_L^3 = 0$. It is helpful to express this constraint also in terms of the ordinary electric charge.

$$\begin{aligned} \sum_R (2Q_R)^3 - \sum_L (2Q_L - 2I_3)^3 =& 0 \\ 8(\sum_R Q_R^3 - \sum_L Q_L^3) + 8\sum_L (3Q_L^2 I_{3L} - 3Q_L I_{3L}^2 + I_{3L}^3) =& 0. \end{aligned}$$

The last term in the third sum vanishes within each doublet and the second term is proportional to $\sum_L Q_L$ because $(I_3)^2 = 1/4$, a constant for all terms. The contribution to the first term in the third sum from each doublet is $(Q^{\uparrow 2} - (Q^{\uparrow} - 1)^2)/2$ which is just equal to the sum of the two charges of the doublet. Thus the whole third sum is proportional to $\sum_L Q_L$ which vanishes by the first "linear" constraint:

$$\sum_R y_R^3 - \sum_L y_L^3 = 8(\sum_R Q_R^3 - \sum_L Q_L^3) + 6\sum_L Q_L.$$

Thus the new information in the "cubic" constraint reduces to

$$\sum_R (Q_R)^3 = \sum_L (Q_L)^3.$$

In the standard model each charged particle state has both a left and right handed component (this means it is possible for all charged particles to gain a (Dirac) mass), and this constraint is automatically satisfied. This left-right symmetry of nonzero charge assignments is an example of a vector-like Q. To define this "vector-like" property, first enumerate all of the fields according to their L components. Thus we think of the right-handed fields as the charge conjugates of left handed fields: $R = L'^c$, so that $Q_{L'} = -Q_R$. In this new labeling the anomaly cancellation conditions read

$$\begin{aligned} \sum_{L \in Doublet} Q_L =& 0 \\ \sum Q_L^3 =& 0 \end{aligned}$$

Then the charge operator Q is vector-like if for each non-vanishing charge $Q > 0$ there are an equal number of left-handed fields with charge Q and $-Q$. Thus the cubic equation above is automatically satisfied if Q is vector-like. In the new notation the most general mass term is of the form

$$\sum_{k,n} m_{k,n} L_k^T i\gamma^2\gamma^0 L_n + h.c.$$

If Q is vector-like it is therefore possible to have a charge conserving mass for every field of non-zero charge. If Q is not vector-like, there must remain at least one massless charged field.

But notice that the anomaly constraint could also have been satisfied in other more intricate ways, which would necessarily mean that Q is not vector-like, entailing the prediction of massless charged fermions, which experiment strongly contradicts. The fact that the standard model fails to rule out such a possibility *a priori*, has led many theorists to the idea of grand unification: that the gauge group is a simple or semi-simple one

(no $U(1)$ factors) and $SU(2) \times U(1)$ is just the remnant "low" energy group. One consequence of this hypothesis is the quantization of charge. In particular the hypercharge assignments will be determined by the representations to which each fermion is assigned. These will vary with the unification group and representation choice. Since the electric charge is proportional to a generator of a semi-simple group, its trace must be zero which means in our context that $\sum_L Q_L = 0$, a constraint satisfied in all unification schemes and satisfied by the fermions of the standard model (again because Q is vector-like). It does not rule out massless charged particles completely, but it does put an additional restriction on how they can arise. But as pointed out by Alvarez-Gaume and Witten, this constraint also follows from the requirement that the fermions can consistently couple to (classical) gravity. We won't show this, but there is a chiral anomaly for gravitons completely analogous to that for gauge particles. We don't need its explicit form since the graviton couples "universally". Thus this anomaly is just proportional to $\mathrm{Tr}\,\lambda_a$ and vanishes for the $SU(2)$ current and for the $U(1)$ current gives $\sum_L y_L - \sum_R y_R = 2(\sum_L Q_L - \sum_R Q_R) = \sum_{L'} Q_{L'} = 0$, where the last sum is over all left-handed fields in the new labelling scheme. Thus *this* argument for grand unification is weakened. Similarly, the consistent coupling of the electroweak $U(1)$ current to QCD requires an anomaly cancellation which holds if and only if $\sum_{L\in quarks} Q_L - \sum_{L\in quarks} Q_R = \sum_{quarks} Q_{L'} = 0$. These three constraints, viz. the cubic electroweak, gravitational, and QCD anomaly cancellation go some distance to forcing the vector-like character of Q.

In fact, if they are applied within a single "generation" ($(\nu, e)_L$, $(u, d)_L$, e_R, u_R, d_R), the vector-like character of Q follows from the single further assumption that precisely one of the leptons is neutral. If the neutrino's charge is fixed to be zero, the charge assignment to every other member of the generation is uniquely fixed by anomaly cancellation (Geng and Marshak, Minahan, Ramond and Warner) to be the standard one. More generally, with $Q_\nu \neq 0$ anomaly cancellation implies the following assignments:

Anomaly Free Charge Assignments: $Q_\nu \neq 1/2$						
ν	e_L	e_R	u_L	d_L	u_R	d_R
Q_ν	$Q_\nu - 1$	$2Q_\nu - 1$	$\frac{1}{2} + \frac{1-2Q_\nu}{2N_c}$	$-\frac{1}{2} + \frac{1-2Q_\nu}{2N_c}$	$\pm\frac{1-2Q_\nu}{2} + \frac{1-2Q_\nu}{2N_c}$	$\mp\frac{1-2Q_\nu}{2} + \frac{1-2Q_\nu}{2N_c}$

Anomaly Free Charge Assignments: $Q_\nu = 1/2$						
ν	e_L	e_R	u_L	d_L	u_R	d_R
$+1/2$	$-1/2$	0	$1/2$	$-1/2$	Q_{u_R}	$-Q_{u_R}$

In the case $Q_\nu \neq 1/2, 0, 1$ *none* of the charge assignments allows a mass for even a single member of the generation: the entire generation must be massless! $Q_\nu = 0$ gives the standard assignments and $Q_\nu = 1$ the charge conjugate standard assignments, both cases allowing a $\Delta I_W = 1/2$ mass for e, u, and d. The case $Q_\nu = 1/2$ forbids a $\Delta I_W = 1/2$ mass for the electron, but the up and down quarks can have a $\Delta I_W = 1/2$ mass if $Q_{u_R} = 1/2$. In this case a $\Delta I_W = 1$ mass for the (ν, e) doublet is possible. All of these possible mass terms violate the gauge symmetry and are forbidden unless the gauge symmetry is spontaneously broken (Higgs mechanism). Mass terms in the quark sector must be $\Delta I_W = 1/2$ (to be color singlets) and can arise if an $I_W = 1/2$ Higgs scalar develops a vacuum expectation value and has a Yukawa coupling to the quarks. In the cases $Q_\nu = 0, 1$, a mass for the electron can arise from the same mechanism. However in the case $Q_\nu = 1/2$, one would also need an $I_W = 1$ Higgs scalar to give a mass to a lepton.

Consequences of Anomalies that Don't Cancel

We have discussed above the restrictions arising when anomalies must cancel. Now we turn to their consequence when they are allowed. This is when the axial currents under examination are not coupled to gauge fields. For example the quark part of the axial current is not coupled to a gauge field. (Parts of it contribute to the electroweak currents but those also include the leptons and their anomaly is cancelled between quarks and leptons.) Thus the neutral component

$$j^\mu_{35} = \bar{q}\frac{\tau_3}{2}\gamma_5\gamma^\mu q = \frac{1}{2}(\bar{u}\gamma_5\gamma^\mu u - \bar{d}\gamma_5\gamma^\mu d)$$

of the axial isospin current for up and down quarks has an axial anomaly, **in the approximation of massless quarks**,

$$\partial_\mu j^\mu_{35} = N_c(\frac{4}{9} - \frac{1}{9})\frac{\alpha_0}{8\pi}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = \frac{N_c\alpha_0}{24\pi}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma},$$

where F is the e.m. field strength. The numerical factors are explained as follows: The up and down quarks couple with opposite signs to j_{35} and the contribution of each to the anomaly is the square of the charge. The overall factor of $N_c = 3$ is for the number of colors of each quark. This equation implies a nonvanishing matrix element of j^μ_{35} between the vacuum and a two photon state. Now neglect the weak interactions, keeping strong interactions to all orders. Since the strong and electromagnetic interactions conserve parity, this matrix element, $\langle 0| j^\mu_{35} |\gamma\gamma\rangle$, must be a pseudo three index tensor $X^{\mu\rho\sigma}$ where ρ and σ are the Lorentz indices describing the polarization of the photons, which carry momenta k_1, k_2 respectively. EM gauge invariance then requires $k_{1\rho}X^{\mu\rho\sigma} = k_{2\sigma}X^{\mu\rho\sigma} = 0$, and Bose statistics for photons requires symmetry under $\rho, k_1 \leftrightarrow \sigma, k_2$. There is essentially only one such pseudotensor (up to terms proportional to k_1^ρ or to k_2^σ which decouple from physical photons) that can be formed from the epsilon symbol and k_1, k_2 since*:

$$\epsilon^{\mu\rho\sigma\tau}(k_1 - k_2)_\tau + \frac{2(k_1+k_2)^\rho k_{1\tau}k_{2\lambda}\epsilon^{\mu\sigma\lambda\tau}}{(k_1+k_2)^2} - \frac{2(k_1+k_2)^\sigma k_{1\tau}k_{2\lambda}\epsilon^{\mu\rho\lambda\tau}}{(k_1+k_2)^2} = \frac{2(k_1+k_2)^\mu k_{1\tau}k_{2\lambda}\epsilon^{\rho\sigma\lambda\tau}}{(k_1+k_2)^2}.$$

Note that any contribution of $O(k)$ at small k is of necessity nonanalytic at zero k. The anomaly equation fixes uniquely the coefficient of this nonanalytic term†. Thus we can write

$$\langle k_1,\lambda_1;k_2,\lambda_2| j^\mu_{35} |0\rangle = i\frac{N_c\alpha_0}{24\pi}\epsilon^{\kappa\nu\rho\sigma}\langle k_1,\lambda_1;k_2,\lambda_2| F_{\kappa\nu}F_{\rho\sigma} |0\rangle \frac{(k_1+k_2)^\mu}{(k_1+k_2)^2},$$

where we stress we have approximated the quark masses as zero. The presence of the singularity at $(k_1 + k_2)^2 = 0$ is a striking consequence of the anomaly. The source of the singularity can be traced to the

---

⋆ The following equation is a special case for $k_1^2 = k_2^2$ of the identity

$$(k_1+k_2)^\mu k_{1\tau}k_{2\lambda}\epsilon^{\rho\sigma\lambda\tau} = \epsilon^{\mu\rho\sigma\tau}(k_1-k_2)_\tau\frac{(k_1+k_2)^2}{2} + \epsilon^{\mu\rho\sigma\tau}(k_1+k_2)_\tau\frac{(k_2^2-k_1^2)}{2} + (k_1+k_2)^\rho k_{1\tau}k_{2\lambda}\epsilon^{\mu\sigma\lambda\tau} - (k_1+k_2)^\sigma k_{1\tau}k_{2\lambda}\epsilon^{\mu\rho\lambda\tau}.$$

This identity can be proved by checking it in the center of mass frame of the two photons $\mathbf{k_1} = -\mathbf{k_2}$.

† This is an essential aspect of the anomaly, reflecting the fact that it really can't be removed by polynomial adjustments to the definition of the current. If an analytic piece of the axial current could produce the anomaly, one could make an analytic adjustment to the defintion of the current to remove it.

masslessness of the quarks. But if quarks are confined they can't be responsible for the singularity in the exact amplitude. There are therefore two possibilities ('t Hooft): either some of the physical baryons are massless or there is a massless scalar (to be identified with a Goldstone boson) coupling to the axial isospin current. The latter possibility seems to be the one realized in Nature, with the pion playing the role of the Goldstone boson:

$$\langle \mathbf{q}, \pi_0 | \, j^\mu_{35} \, | 0 \rangle = \frac{i q^\mu f_\pi}{(2\pi)^{3/2}\sqrt{2\omega}}.$$

(The π^0 is related by strong isospin to the $\pi^\pm$. In the limit of exact isospin ($m_u = m_d$ and EM turned off) $f_{\pi^0} = f_{\pi^\pm}$, and the latter can be independently measured in the weak decay process $\pi^- \to \mu^- + \bar{\nu}_\mu$.) In that case the residue of the pole is $(2\pi)^{3/2}\sqrt{2\omega} f_\pi q^\mu$ times the transition amplitude for the π_0 to decay into two photons. In perturbation theory, this latter transition amplitude is $(-i)$ times the matrix element of the perturbation in the Hamiltonian density describing the electromagnetic interactions of hadrons, from which we conclude:

$$\langle k_1, \lambda_1; k_2, \lambda_2 | \, \mathcal{H}_I^{e.m.}(0) \, | \mathbf{q}, \pi_0 \rangle \approx -\frac{1}{(2\pi)^{3/2}\sqrt{2\omega}} \frac{N_c \alpha_0}{24\pi f_\pi} \epsilon^{\mu\nu\rho\sigma} \, \langle k_1, \lambda_1; k_2, \lambda_2 | \, F_{\mu\nu} F_{\rho\sigma} \, | 0 \rangle \, ,$$

where the approximation is due only to the fact that the quarks and the pion are not exactly massless. This corresponds to a term in the effective lagrangian:

$$\frac{N_c \alpha_0}{24\pi f_\pi} \pi_0 \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}$$

To the extent that this is a good approximation, we see that the anomaly controls the decay $\pi_0 \to 2\gamma$. In fact, this approximation gives a good account of the experimental rate to within 20%. This success may be regarded as evidence for the three colors of quarks. Incidentally, the anomaly breaks the apparent chiral $U(1)$ invariance remaining after electromagnetic interactions have broken $SU(2) \times SU(2)$ by virtue of unequal up and down charges. In particular the π_0 will have a small squared mass of order α^2 even if all quark masses are zero. This shift however is very small compared to the order α shift given to the π^+ and does not disturb the derivation of the Gell-Mann Okubo relation in the $SU(3) \times SU(3)$ case.

Mathematical Consequences of the Anomaly: Index Theorems

The chiral anomaly puts constraints on the eigenvalues of the Euclidean space Dirac operator $(1/i)\gamma \cdot D$ in the presence of sufficiently "nice" gauge fields. This differential operator is antihermitian, provided it acts on functions for which one can integrate by parts without keeping surface terms. We assume that A is such that the Dirac operator has a complete set of eigenfunctions with this property. Then since it is antihermitian, the eigenvalues will be purely imaginary. For each nonzero eigenvalue $i\lambda_r$ there is another eigenvalue $-i\lambda_r$, because γ_5 anticommutes with γ^μ: if ψ_r is the eigenvector for $i\lambda_r$ then $\gamma_5 \psi_r$ is the eigenvector for $-i\lambda_r$.

Now introduce the Euclidean Green function for $m + \frac{1}{i}\gamma \cdot D$:

$$(m + (1/i)\gamma \cdot D) S_E(x, y; A) = \delta(x - y).$$

We can construct the gauge invariant $\bar{S}_E$ in Euclidean space exactly as in Minkowski space. Then the quantity

$$J^\mu_{5E} = -\operatorname{Tr} \gamma_5 \gamma^\mu \lim_{y \to x} \bar{S}_E(x, y; A)$$

has the anomalous conservation law

$$\partial_\mu J^\mu_{5E} = 2im \operatorname{Tr} \gamma_5 \lim \bar{S} - i \frac{g^2}{16\pi^2} \epsilon^{\mu\nu\rho\sigma} \operatorname{Tr} F_{\mu\nu} F_{\rho\sigma}.$$

The extra factor $-i$ in the second term on the r.h.s. appears because we have continued the Minkowski

result to Euclidean space: $x^0 = -ix^4$, and every contravariant time index gets this same $-i$. We could have used a "Euclidean" ϵ_E defined as the continuation of the Minkowski one so that $\epsilon_E^{1234} \equiv i\epsilon^{1230} = -i$, and with such a definition no $-i$ would appear. But the epsilon symbol is conventionally always understood to be real, and with the convention $\epsilon^{1234} = +1$ the $-i$ must be explicitly included as written. Because $m \neq 0$ there are no infrared singularities so if we integrate both sides over x, the l.h.s. will vanish and we get the identity

$$2m \int d^4x \, \mathrm{Tr}\, \gamma_5 \lim \bar{S} - \frac{g^2}{16\pi^2} \int d^4x \epsilon^{\mu\nu\rho\sigma} \, \mathrm{Tr}\, F_{\mu\nu} F_{\rho\sigma} = 0.$$

Now the quantity $\mathrm{Tr}\, \gamma_5 \bar{S}$ potentially has u.v. divergences, which would make the limit delicate. However the γ_5 requires multiplication by at least 4 gamma matrices to give a nonvanishing trace. To get this many one has to go to the order A^2 term in the weak field expansion which is a priori linearly divergent. But this is the term with 5 gamma matrices and that trace vanishes: the term with 4 gamma matrices is only log divergent. Finally to saturate the epsilon tensor one needs at least two vectors: one could be ϵ but the other must be an external momentum, which gives one further power of convergence, enough to make it finite. Similarly, all higher terms are convergent. Thus $\bar{S}$ can be replaced by S, and the limit $y \to x$ safely taken.

If the eigenfunctions of $(1/i)\gamma \cdot D$ are complete, we can represent the Green function as

$$S_E(x, y; A) = \sum_r \frac{\psi_r(x)\psi_r^\dagger(y)}{m + i\lambda_r}$$

and thus

$$\int d^4x \, \mathrm{Tr}\, \gamma_5 S = \sum_r \frac{\int d^4x \psi_r^\dagger(x)\gamma_5\psi_r(x)}{m + i\lambda_r}$$

But all terms for which $\lambda_r \neq 0$ vanish because ψ_r and $\gamma_5\psi_r$ then have different eigenvalues and so are orthogonal. Thus the sum is just over the values of r for which $\lambda_r = 0$. We can organize the zero eigenfunctions according to the eigenvalues of γ_5 which are $+1$ and -1 Let $n_\pm$ be the number of zero eigenvalues with ± 1 eigenvalue of γ_5. Then we have $\int d^4x \, \mathrm{Tr}\, \gamma_5 S = \frac{n_+ - n_-}{m}$ and finally the Atiyah-Singer index theorem

$$\frac{g^2}{32\pi^2} \int d^4x \epsilon^{\mu\nu\rho\sigma} \, \mathrm{Tr}\, F_{\mu\nu} F_{\rho\sigma} = (n_+ - n_-).$$

We have not been precise about the conditions on A except to say that the Dirac operator must possess a complete set of eigenfunctions. It is not hard to show that ϵF^2 is a total derivative of a gauge noninvariant function. Thus if the A falls off at infinity sufficiently rapidly we would expect the l.h.s. to vanish, so $n_+ = n_-$, with no conclusion about whether there are any zero eigenvalues of the Dirac operator. The proper condition is not that A vanish at infinity, but rather that it approach a pure gauge there. Then the l.h.s. just measures the number of times this gauge function "winds around" the three-sphere at infinity. The manifold of $SU(2)$ is the three sphere, so nontrivial windings are possible in that case, but not in the $U(1)$ case, when the manifold is a circle. (You can't lasso a sphere.) For such topologically nontrivial gauge fields the index theorem implies at least one vanishing eigenvalue for the Dirac operator. This has the nontrivial physical implication that the vacuum persistence amplitude, $\det \gamma \cdot D$, in the presence of such a field *vanishes* when the mass of the field vanishes.

17. Scattering Amplitudes in Quantum Field Theory

The observables of quantum field theory are local quantum fields. For example the energy momentum can be expressed as an integral of the local energy momentum tensor.

$$P^\mu = (P^0, \mathbf{P}) = \int d^3x \theta^{\mu 0}(\mathbf{x}, t)$$

where $\theta^{\mu\nu}(x)$, the energy momentum tensor, is a local operator:

$$[\theta^{\mu\nu}(x), \theta^{\rho\sigma}(y)] = 0 \qquad\qquad (x-y)^2 > 0\,.$$

<u>Single Particle States</u> $|\mathbf{p}, a\rangle$, with $\langle \mathbf{p}, a|\mathbf{p}', b\rangle = \delta_{ab}\delta^3(\mathbf{p}-\mathbf{p}')$ are characterized by possessing a unique energy for each momentum:

$$P^\mu|\mathbf{p}, a\rangle = \left(\sqrt{m_a^2 + \mathbf{p}^2}, \mathbf{p}\right)|\mathbf{p}, a\rangle.$$

The space-time picture of scattering processes requires the use of wave packet single particle states:

$$|f, a\rangle \equiv \int d^3p f(\mathbf{p})|\mathbf{p}, a\rangle.$$

Their physical properties follow by considering for any local operator $\Omega(x)$,

$$\langle f, a|\Omega(x)|f, a\rangle = \int d^3p' d^3p f^*(\mathbf{p}')f(\mathbf{p})e^{i(p-p')\cdot x}\langle \mathbf{p}', a|\Omega(0)|\mathbf{p}, a\rangle,$$

where we used translation invariance in space and time $\Omega(x) = e^{-iP\cdot x}\Omega(0)e^{iP\cdot x}$. This means we are restricting consideration to vanishing external fields. Assuming f is peaked at $\mathbf{p}_0$, over an interval small compared to important variations of $\langle \mathbf{p}', a|\Omega(0)|\mathbf{p}, a\rangle$, we have

$$\begin{aligned}\langle f, a|\Omega(x)|f, a\rangle &\approx \langle \mathbf{p}_0', a|\Omega(0)|\mathbf{p}_0, a\rangle \left[\int d^3p' f^*(\mathbf{p}')e^{-i\left(\mathbf{p}'\cdot\mathbf{x} - \omega_a(\mathbf{p}')t\right)}\right] \\ &\quad \left[\int d^3p f(\mathbf{p})e^{i\left(\mathbf{p}\cdot\mathbf{x} - \omega_a(\mathbf{p})t\right)}\right] \\ &\approx |f(\mathbf{x}, t)|^2\langle \mathbf{p}_0', a|\Omega(0)|\mathbf{p}_0, a\rangle.\end{aligned}$$

We can select $f(\mathbf{p})$ so that at some particular time, t,

$$f(\mathbf{x}, t) = \int d^3p f(\mathbf{p})e^{i\left(\mathbf{p}\cdot\mathbf{x} - \omega_a(\mathbf{p})t\right)}$$

is confined to some spatial volume $V(t)$. As time evolves this volume will move and spread. The center of the packet will follow a straight line trajectory at the group velocity

$$\mathbf{v}_g = \left.\frac{d\omega_a(\mathbf{p})}{d\mathbf{p}}\right|_{\mathbf{p}=\mathbf{p}_0} = \frac{\mathbf{p}_0}{\omega_a(\mathbf{p}_0)} = \mathbf{v}_0$$

and spreading will be negligible over a time interval $\ll \omega_a(\mathbf{p}_0)/(\Delta p)^2 < \Delta x/\Delta v$ where Δp is the width of the peak in f_a at $\mathbf{p}_0$.

<u>Creation Operators for Single Particle States</u>

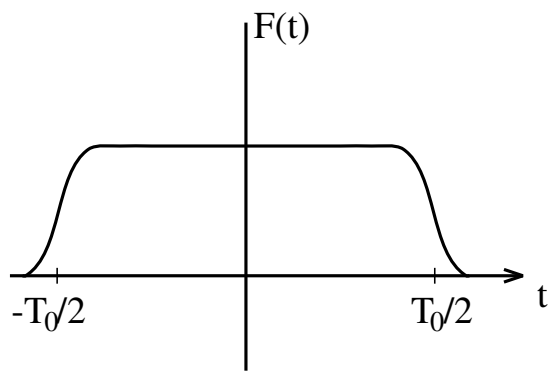
We *assume* that for each single particle state, a, there is a local operator $\Omega_a(x)$ such that

$$\langle 0|\Omega_a(0)|\mathbf{p}, a\rangle \neq 0$$

Ω_a must carry the quantum numbers of a, but is otherwise unspecified. A given Ω may couple to several particles. Introduce wave packets $f_a(\mathbf{p})$ peaked about some momentum with a narrow width Δp assumed much smaller than any variation in matrix elements or $\omega_a(\mathbf{p})$. These packets are selected so that

$$f_a(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi)^3} f_a(\mathbf{p}) e^{i(\mathbf{p}\cdot x - \omega_a(\mathbf{p})t)}$$

is confined to a volume $V_a(t) = 0\left(1/\Delta p^3\right)$ with negligible spreading in the interval $-T/2 < t < +T/2$. Introduce also a switching function $F(t)$ with shape



and normalized by

$$\int_{-\infty}^{\infty} dt F(t) = \tilde{F}(0) = 1$$

We stipulate that $1/T_0$ is much smaller than any important momentum variation in matrix elements or the function $\omega_a(\mathbf{p})$. In particular, $1/T_0$ is much smaller than any mass differences. This means that $\tilde{F}(\omega)$ the Fourier transform of F can be chosen sharply peaked in ω about $\omega = 0$. The width of this peak is only limited by $\Delta\omega > 1/T_0$. Then, we define (for $-T/2 < t_0 < T/2$)

$$A_a^\dagger(t_0) \equiv \int d^4x f_a(\mathbf{x}, t) F(t - t_0) \Omega_a^\dagger(x)$$

The essential point here is that $\int d^4x$ forces $A_a^\dagger(t_0)$ to create energy-momentum in a narrow range about $\left(\mathbf{p}_0, \omega_a(\mathbf{p}_0)\right)$. It then follows:

(1) $A_a^\dagger(t_0)$ creates a particle, a, in a wave packet f_a at t_0.

(2) $A_a(t_0)$ destroys a particle, a, in a wave packet f_a at t_0.

Proof:

$$
\begin{aligned}
&\langle 0|A_a^f(t_0)|\mathbf{p},b\rangle \\
&= \int d^4x f_a^*(\mathbf{x},t)F(t-t_0)\langle 0|\Omega_a(0)|\mathbf{p},b\rangle e^{ip_b\cdot x} \\
&= \int_{-\infty}^{\infty} dt f_a^*(\mathbf{p})e^{i\left(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\right)t}F(t-t_0)\langle 0|\Omega_a(0)|\mathbf{p},b\rangle \\
&= f_a^*(\mathbf{p})\langle 0|\Omega_a(0)|\mathbf{p},b\rangle e^{i\left(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\right)t_0}\underbrace{\tilde{F}\big(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\big)}_{\neq 0 \text{ only when } m_a=m_b} \\
&\approx \delta_{ab} f_a^*(\mathbf{p})\langle 0|\Omega_a(0)|\mathbf{p}_0,b\rangle
\end{aligned}
$$

The last approximate equality assumes each mass value is non degenerate, and is based on the narrow peaking of $\tilde{F}$. The equality is exact as long as the support of $\tilde{F}$ lies entirely within the gap between ω_a and any other energy available to the system. The fundamental assumption is that each single particle for which such a creation operator can be defined has a mass separated by a finite gap from any other mass. If the state b is a multiparticle state, its energy is not discrete but must in any case be greater than the sum of the masses of the particles it describes. This minimum multiparticle energy is also assumed to be separated by a gap from m_a. One can handle degeneracies in mass by carefully choosing the operators Ω_a so that they decouple from the single particle states degenerate with particle a. Clearly the construction does not completely succeed if there are massless particles in the theory because then a multiparticle state containing particle a and several low energy massless particles can have an energy arbitrarily close to m_a. In fact, in this case it is experimentally impossible to give a definite meaning to "single particle state" so the limitations of the theory are appropriate.

Normalization:

$$
\begin{aligned}
&\langle 0|A_a(t_0)A_a^\dagger(t_0)|0\rangle \\
&\approx \int d^3p\langle 0|A_a(t_0)|\mathbf{p},a\rangle\langle \mathbf{p},a|A_a^\dagger(t_0)|0\rangle = \int d^3p\big|f_a(\mathbf{p})\big|^2\big|\langle 0|\Omega_a(0)|\mathbf{p},a\rangle\big|^2 \\
&\approx \big|\langle 0|\Omega_a(0)|\mathbf{p}_0,b\rangle\big|^2 \qquad \text{since } f_a \text{ is sharply peaked.}
\end{aligned}
$$

The first approximate equality follows because after insertion of a complete set of states between A and $A^\dagger$, the smearing functions in the definition of the latter focus the energy and momentum of the intermediate states to be (within the allowed windows $\sim (1/T_0, \Delta p)$) those of the single particle state a. Thus only those give a significant contribution. The last approximate equality is based on the narrow peaking of the wave packet f_a.

Multiparticle States

We can use our creation operators to construct multiparticle states. For example

$$A_a^\dagger(t_0)A_b^\dagger(t_0)|0\rangle$$

creates a two particle state. Of course, there may be interactions between the two particles, so *this interpretation is only meaningful when packet, a, is spacelike separated from packet*, b. In this case, we have

$$\Big[A_a^\dagger(t_0), A_b^\dagger(t_0)\Big] = 0 \qquad (\text{Since } [\Omega_a(x),\Omega_b(y)] = 0 \text{ for } (x-y)^2 > 0)$$

from which it follows that the norm of the two particle state factorizes into the product of single particle

norms:

$$\begin{aligned}
&\langle 0|A_a(t_0)A_b(t_0)A_a^\dagger(t_0)A_b^\dagger(t_0)|)\rangle \\
&\quad = \langle 0|A_a(t_0)A_a^\dagger(t_0)A_b(t_0)A_b^\dagger(t_0)|0\rangle \\
&\quad \approx \langle 0|A_a(t_0)A_a^\dagger(t_0)|0\rangle\langle 0|A_b(t_0)A_b^\dagger(t_0)|0\rangle
\end{aligned}$$

Furthermore, for any operator of the form $\Omega = \int d^3x f(\mathbf{x})\Omega(\mathbf{x}, t_0)$, we have

$$\begin{aligned}
&\langle 0|A_a(t_0)A_b(t_0)\Omega A_a^\dagger(t_0)A_b^\dagger(t_0)|0\rangle \\
&\quad \approx \langle 0|A_a(t_0)\Omega A_a^\dagger(t_0)|0\rangle\langle 0|A_b(t_0)A_b^\dagger(t_0)|0\rangle \\
&\qquad + \langle 0|A_a(t_0)A_a^\dagger(t_0)|0\rangle\langle 0|A_b(t_0)\Omega A_b^\dagger(t_0)|0\rangle
\end{aligned}$$

This follows since we can write effectively

$$\Omega \sim \int\limits_{V_a(t_0)} d^3x f(\mathbf{x})\Omega(\mathbf{x}, t_0) + \int\limits_{V_b(t_0)} d^3x f(\mathbf{x})\Omega(\mathbf{x}, t_0)$$

and the two pieces act independently on particles a and b. Since all the observables of a quantum field theory are local fields such as $\Omega(x)$, this justifies the multiparticle interpretation. It should be clear that one can extend this construction to states with any number of spatially separated particles.

<u>Reduction Formulae</u> $\quad -\frac{T}{2} < t_1, t_2, \ldots, t_n < \frac{T}{2}$

$$\begin{aligned}
&T\left\{A_a\left(\frac{T}{2}\right)\Omega_1(x_1)\ldots\Omega_n(x_n)\right\} - T\left\{\Omega_1(x_1)\ldots\Omega_n(x_n)A_a\left(-\frac{T}{2}\right)\right\} \\
&= \int\limits_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 \frac{d}{dt_0} T\{A_a(t_0)\Omega_1(x_1)\ldots\Omega_n(x_n)\} \\
&= \int\limits_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 \int d^4x F(t-t_0)\frac{\partial}{\partial t} f_a^*(\mathbf{x}, t) T\{\Omega_a(x)\Omega_1(x_1)\ldots\Omega_n(x_n\} \\
&= \int\limits_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 \int d^4x F(t-t_0) f_a^*(\mathbf{x}, t)\left(i\omega_a(-i\nabla) + \frac{\partial}{\partial t}\right) T\{\Omega_a(x)\Omega_1(x_1)\ldots\Omega_n(x_n)\}
\end{aligned}$$

$$\begin{aligned}
&T\left\{A_a^\dagger\left(\frac{T}{2}\right)\Omega_1(x_1)\ldots\Omega_n(x_n)\right\} - T\left\{\Omega_1(x_1)\ldots\Omega_n(x_n)A_a^\dagger\left(-\frac{T}{2}\right)\right\} \\
&= \int\limits_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 \int d^4x F(t-t_0)\frac{\partial}{\partial t} f_a(\mathbf{x}, t) T\{\Omega_a^\dagger(x)\Omega_1(x_1\ldots\Omega_n(x_n)\} \\
&= \int\limits_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 \int d^4x F(t-t_0) f_a(\mathbf{x}, t)\left(-i\omega_a(i\nabla) + \frac{\partial}{\partial t}\right) T\{\Omega_a^\dagger(x)\Omega_1(x_1)\ldots\Omega_n(x_n)\}\,.
\end{aligned}$$

In these reduction formulae we have used the fact that by construction f_a satisfies a Schrödinger equation with hamiltonian $\omega_a(-i\nabla) = \sqrt{m_a^2 - \nabla^2}$. Spatial integration by parts then allows the spatial derivatives to be transferred to $\Omega(x)$.

Single Particle States are Handled Consistently

A single particle state prepared at early times should remain a single particle state for all times. This follows from the reduction formulae:

$$\langle 0|A_a\left(\frac{T}{2}\right)A_b^\dagger\left(-\frac{T}{2}\right)|0\rangle = \langle 0|A_a\left(-\frac{T}{2}\right)A_b^\dagger\left(-\frac{T}{2}\right)|0\rangle$$

$$+\int_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 \int d^4x F(t-t_0) f_a^*(\mathbf{x},t)\left(\frac{\partial}{\partial t}+i\omega_a(\nabla)\right)\langle 0|\Omega_a A_b^\dagger\left(-\frac{T}{2}\right)|0\rangle$$

But $\langle 0|\Omega_a(x)A_b^\dagger\left(-T/2\right)|0\rangle \approx \int d^3p\langle 0|\Omega_a(0)|\mathbf{p},b\rangle\langle \mathbf{p},b|A_b^\dagger\left(-\frac{T}{2}\right)|0\rangle e^{+i\left(p\cdot\mathbf{x}-\omega_b(\mathbf{p})t\right)}$
So the second term becomes

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 \int_{-\infty}^{\infty} dt \int d^3p f_a^*(\mathbf{p})F(t-t_0)i\left(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\right)e^{i\left(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\right)t}$$

$$\langle 0|\Omega_a(0)|\mathbf{p},b\rangle\langle \mathbf{p},b|A_b^\dagger\left(-\frac{T}{2}\right)|0\rangle$$

$$= \int d^3p \int_{-\frac{T}{2}}^{\frac{T}{2}} dt_0 e^{i\left(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\right)t_0} f_a^*(\mathbf{p})\tilde{F}\left(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\right)i\left(\omega_a(\mathbf{p})-\omega_b(\mathbf{p})\right)$$

$$\langle 0|\Omega_a|\mathbf{p},b\rangle\langle \mathbf{p},b|A_b^\dagger|0\rangle$$

$$\approx 0$$

Since $\tilde{F}\approx 0$ if $b\neq a$ and the r.h.s. is identically zero if $a=b$ because of the factor $\omega_a(\mathbf{p})-\omega_b(\mathbf{p})$. The approximate equalities become exact in the limit

$$T, T_0 \to \infty \qquad \text{with} \qquad \frac{1}{\delta m} \ll T_0 \ll T \ll T_0^2 m$$

(The latter inequality insures negligible wave packet spread.)

We therefore obtain the required result:

$$\langle 0|A_a\left(\frac{T}{2}\right)A_b^\dagger\left(-\frac{T}{2}\right)|0\rangle \approx \lim\langle 0|A_a\left(-\frac{T}{2}\right)A_b^\dagger\left(-\frac{T}{2}\right)|0\rangle$$
$$\approx \left|\langle 0|\Omega_a(0)|\mathbf{p}_0,a\rangle\right|^2\delta_{ab}$$

where the approximate equality becomes exact as $T_0, T\to\infty$.

Two Particle Scattering Amplitudes

The scattering process will consist of preparation at a very early time $-T'/2$ of two separated single particle wave packets with well defined momentum, aimed to collide at roughly time $t\approx 0$, and the subsequent observation at a much later time $+\frac{T}{2}$ of two well separated wave packets.

Let f_a, f_b describe the incoming packets with momentum $\mathbf{p}_a$ and $\mathbf{p}_b$ respectively; f_c, f_d the outgoing packets with momentum $\mathbf{p}_c, \mathbf{p}_d$. The information we want is contained in the matrix element

$$\begin{aligned}
&\langle 0|A_c\left(\frac{T}{2}\right)A_d\left(\frac{T}{2}\right)A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle \\
&\quad\approx \langle 0|\Omega_c|\mathbf{p}_c\rangle\langle 0|\Omega_d(0)|\mathbf{p}_d\rangle\langle\mathbf{p}_a|\Omega_a^\dagger(0)|0\rangle\langle\mathbf{p}_b|\Omega_b^\dagger(0)|0\rangle \\
&\qquad\int d^3p'\int d^3q'\int d^3p\int d^3q f_c^*(\mathbf{p}')f_d^*(\mathbf{q}')f_a(\mathbf{p})f_b(\mathbf{q})\langle\mathbf{p}',c;\mathbf{q}',d|S|\mathbf{p},a;\mathbf{q},b\rangle,
\end{aligned}$$

which can be taken as the definition of the S-matrix. Now, we apply the Reduction Procedure.

$$\begin{aligned}
&\langle 0|A_c\left(\frac{T}{2}\right)A_d\left(\frac{T}{2}\right)A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle \\
&\quad=\langle 0|A_c\left(\frac{T}{2}\right)A_d\left(-\frac{T}{2}\right)A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle \\
&\qquad+\int_{-\frac{T}{2}}^{\frac{T}{2}}dt_0^1\int d^4x_1F(t_1-t_0^1)f_d^*(x_1)\left(\frac{\partial}{\partial t_1}+i\omega_d(i\nabla_1)\right) \\
&\qquad\qquad\langle 0|A_c\left(\frac{T}{2}\right)\Omega_d(x_1)A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle
\end{aligned}$$

The first term on the RHS contains the amplitude that nothing happens. The final packet d has been extrapolated back as a free particle to time $-T/2$. If no scattering has occurred this extrapolated packet can overlap with initial packet a or packet b but not both since a and b are spatially separated. Thus $A_d(-T/2)$ will commute with $A_a(-T'/2)$ or $A_b(-T'/2)$ or both. Both possibilities are included by writing

$$\begin{aligned}
&\langle 0|A_c\left(\frac{T}{2}\right)A_d\left(-\frac{T}{2}\right)A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle \\
&\approx\delta_{m_a,m_d}\langle 0|A_c\left(\frac{T}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle\langle 0|\Omega_d(0)|\mathbf{p}_d\rangle\langle\mathbf{p}_a|\Omega_a^\dagger(0)|0\rangle\int d^3pf_d^*(\mathbf{p})f_a(\mathbf{p}) \\
&\quad+\delta_{m_b,m_d}\langle 0|A_c\left(\frac{T}{2}\right)A_a^\dagger\left(-\frac{T'}{2}\right)|0\rangle\langle 0|\Omega_d(0)|\mathbf{p}_d\rangle\langle\mathbf{p}_b|\Omega_b^\dagger(0)|0\rangle\int d^3pf_d^*(\mathbf{p})f_b(\mathbf{p}) \\
&\approx\Bigg[\delta_{m_a,m_d}\delta_{m_b,m_c}\int d^3pf_d^*(\mathbf{p})f_a(\mathbf{p})\int d^3pf_c^*(\mathbf{p})f_b(\mathbf{p}) \\
&\qquad+\delta_{m_a,m_c}\delta_{m_b,m_d}\int d^3pf_d^*(\mathbf{p})f_b(\mathbf{p})\int d^3pf_c^*(\mathbf{p})f_a(\mathbf{p})\Bigg] \\
&\qquad\times\langle 0|\Omega_d(0)|\mathbf{p}_d\rangle\langle 0|\Omega_c(0)|\mathbf{p}_c\rangle\langle\mathbf{p}_a|\Omega_a^\dagger(0)|0\rangle\langle\mathbf{p}_b|\Omega_b^\dagger(0)|0\rangle
\end{aligned}$$

where $\delta_{mm'}=0$ if $m\neq m'$, 1 if $m=m'$. Of course if the extrapolated packet overlaps with neither a nor b this term contributes nothing, as the above formula states because then f_d will be orthogonal to both f_a and f_b.

In the 2nd term we now reduce particle c:

$$\begin{aligned}
&\langle 0|A_c\left(\frac{T}{2}\right)\Omega_d(x_1)A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle \\
&\quad\approx\langle 0|\Omega_d(x_1)A_c\left(-\frac{T}{2}\right)A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle \\
&\qquad+\int_{-\frac{T}{2}}^{\frac{T}{2}}dt_0^{(2)}\int d^4x_2F(t_2-t_0^2)f_c^*(x_2)\left(\frac{\partial}{\partial t_2}+i\omega_c(i\nabla_2)\right) \\
&\qquad\times\langle 0|T[\Omega_c(x_2)\Omega_d(x_1)]A_a^\dagger\left(-\frac{T'}{2}\right)A_b^\dagger\left(-\frac{T'}{2}\right)|0\rangle
\end{aligned}$$

The first term will give negligible contribution because the factor $\left(\frac{\partial}{\partial t}+i\omega_d(-i\nabla_1)\right)$ will yield a factor $i\left(\omega_d+\omega_c-\omega_a-\omega_b\right)$ which will vanish within the support of the factor $\tilde{F}(\omega_d+\omega_c-\omega_a-\omega_b)$. (A nonvanishing contribution would require either $c=a$ or $c=b$, because A_c is a destruction operator and then the support of $\tilde{F}$ requires $d=b$ or respectively $d=a$.) In this second reduction we have tacitly written $A_c\left(\frac{T}{2}\right)\Omega_d(x_1)=T\left[A_c\left(\frac{T}{2}\right)\Omega_d(x_1)\right]$ which is not strictly true since $A_c\left(\frac{T}{2}\right)$ can involve $\Omega_c(y)$ at some times earlier than t_1. However the integrands include wave packets which are spacelike separated for times near $\frac{T}{2}$. Since Ω_c and Ω_d commute at spacelike separations no mistake is made by this procedure. Similarly the replacement $T\left[\Omega_d(x)A_c\left(-\frac{T}{2}\right)\right]=\Omega_d(x)A_c\left(-\frac{T}{2}\right)$ is validated because the wave packets extrapolated back to times near $-\frac{T}{2}$ will again be space-like separated. (We assume packets are aimed to overlap at times near 0.)

To continue reducing, we need to replace $T[\Omega_c(x_2)\Omega_d(x_1)]\ A_a^\dagger(-T'/2)$ by $T[\Omega_c(x_2)\ \Omega_d(x_1)A_a^\dagger(-T'/2)]$. The above argument could be applied for nonforward but not quite for forward scattering because then one of the final packets extrapolated to time $-T/2$ could well overlap one of the initial packets. However, by choosing $T'/2>T/2+T_0$, we force $-T'/2$ to be earlier than either t_1 or t_2 and likewise $T'/2$ to be later than t_1 or t_2. Then the required replacements are valid. As long as $-T/2$ is well before and $T/2$ well after the collision, such a choice is completely satisfactory.

With this in mind we can reduce particles a and b. We drop immediately terms where $A^\dagger$ stands next to $\langle 0|$, because they could only produce a negative energy state which we assume does not exist. The final result is:

$$\langle 0|\Omega_c(0)|\mathbf{p}_c\rangle\langle 0|\Omega_d(0)|\mathbf{p}_d\rangle\langle\mathbf{p}_a|\Omega_a^\dagger(0)|0\rangle\langle\mathbf{p}_b|\Omega_b^\dagger(0)|0\rangle\int d^3p'\,d^3q'\int d^3p\,d^3q$$
$$\times f_c^*(\mathbf{p}')f_d^*(\mathbf{q}')f_a(\mathbf{p})f_b(\mathbf{q})\langle\mathbf{p}',c;\mathbf{q}',d|(S-I)|\mathbf{p},a;\mathbf{q},b\rangle$$

$$\approx\int_{-T/2}^{T/2}dt_0^{(1)}dt_0^{(2)}\int_{-T'/2}^{T'/2}dt_0^{(3)}dt_0^{(4)}\int d^4x_1d^4x_2d^4x_3d^4x_4F(t_1-t_0^1)F(t_2-t_0^2)F(t_3-t_0^3)F(t_4-t_0^4)$$
$$f_d^*(x_1)f_c^*(x_2)f_a(x_3)f_b(x_4)\left(\frac{\partial}{\partial t_1}+i\omega_d\right)\left(\frac{\partial}{\partial t_2}+i\omega_c\right)\left(\frac{\partial}{\partial t_3}-i\omega_a\right)\left(\frac{\partial}{\partial t_4}-i\omega_b\right)$$
$$\langle 0|T[\Omega_c(x_2)\Omega_d(x_1)\Omega_a^\dagger(x_3)\Omega_b^\dagger(x_4)]|0\rangle$$

Now define

$$T[q_2q_1;q_3q_4]\equiv\int d^4x_1d^4x_2d^4x_3d^4x_4e^{-i(q_2\cdot x_2+q_1\cdot x_1-q_3\cdot x_3-q_4\cdot x_4)}$$
$$\langle 0|T[\Omega_c(x_2)\Omega_d(x_1)\Omega_a^\dagger(x_3)\Omega_b^\dagger(x_4)]|0\rangle$$

and the r.h.s. becomes

$$\int_{-T/2}^{T/2}dt_0^1dt_0^2\int_{-T'/2}^{T'/2}dt_0^3dt_0^4\int\prod_m\frac{d^4q_m}{(2\pi)^4}\tilde{F}(\omega_d-q_0^1)\tilde{F}(\omega_c-q_0^2)\tilde{F}(q_0^3-\omega_a)\tilde{F}(q_0^4-\omega_b)$$
$$\tilde{f}_d^*(\mathbf{q}_1)\tilde{f}_c^*(\mathbf{q}_2)f_a(\mathbf{q}_3)f_b(\mathbf{q}_4)i^4(\omega_d-q_1^0)(\omega_c-q_2^0)(\omega_a-q_3^0)(\omega_b-q_4^0)$$
$$e^{i(\omega_d-q_1^0)t_0^1}e^{i(\omega_c-q_2^0)t_0^2}e^{-i(\omega_a-q_3^0)t_0^3}e^{-i(\omega_b-q_4^0)t_0^4}T[q_2,q_1;q_3,q_4]$$
$$\widetilde{T,T'\to\infty}\int\frac{d^3q_1d^3q_2d^3q_3d^3q_4}{(2\pi)^3(2\pi)^3(2\pi)^3(2\pi)^3}f_d^*(\mathbf{q}_1)f_c^*(\mathbf{q}_2)f_a(\mathbf{q}_3)f_b(\mathbf{q}_4)$$
$$\lim\ (\omega_d-q_1^0)(\omega_c-q_2^0)(\omega_a-q_3^0)(\omega_b-q_4^0)T[q_2,q_1;q_3,q_4],$$

where the limit taken is $q_1^0\to\omega_d$, $q_2^0\to\omega_c$, $q_3^0\to\omega_a$, $q_4^0\to\omega_b$, and is forced by delta functions arising from the large T,T' limits of the integrals over the t_0^k. We see that there is nonzero scattering only if T has poles in all of the q_i^0 at the energies of the respective incoming and outgoing particles.

Taking the limit of infinitely narrow packets we may express our result as follows

$$
\begin{aligned}
T[p',q';p,q] \underset{\substack{p^{0\prime}\to\omega_c \\ q^{0\prime}\to\omega_d \\ p^0\to\omega_a \\ q^0\to\omega_b}}{\longrightarrow} & [(2\pi)^3]^4(2\omega_a)2\omega_b)(2\omega_c)(2\omega_d)\langle 0|\Omega_c(0)|\mathbf{p}'\rangle\langle 0|\Omega_d(0)|\mathbf{q}'\rangle \\
& \times \langle \mathbf{p}|\Omega_a^\dagger(0)|0\rangle\langle \mathbf{q}|\Omega_b^\dagger(0)|0\rangle \\
& \times \frac{-i}{p'^2+m_c^2}\frac{-i}{q'^2+m_d^2}\frac{-i}{p^2+m_a^2}\frac{-i}{q^2+m_b^2} \\
& \times \big[\langle \mathbf{p}',c;\mathbf{q}',d|S|\mathbf{p},a;\mathbf{q},b\rangle - \delta_{ac}\delta_{bd}\delta^3(\mathbf{p}'-\mathbf{p})\delta^3(\mathbf{q}'-\mathbf{q}) \\
& \qquad -\delta_{ad}\delta_{bc}\delta^3(\mathbf{p}-\mathbf{q}')\delta^3(\mathbf{p}'-\mathbf{q})\big]
\end{aligned}
$$

Note that translation invariance implies that T has an overall factor of $(2\pi)^4\delta^4(p'+q'-p-q)$ expressing energy momentum conservation. It should now be clear how the generalization to an arbitrary number of incoming and outgoing particles should be expressed. The F.T. of the T.O.P. of interpolating fields will have a pole factor $-i/(q_a^2+m_a^2)$ associated with each external line. The coefficient of all these pole factors is proportional to the desired scattering amplitude. Furthermore, the factors of proportionality are clear: a factor $2\omega_a(2\pi)^3\langle \mathbf{p}_a|\,\Omega_a(0)\,|0\rangle$ for each incoming line and a factor $2\omega_a(2\pi)^3\langle 0|\,\Omega_a^\dagger(0)\,|\mathbf{p}_a\rangle$ for each outgoing line.

We are left with the task of computing the matrix elements of the interpolating fields between the vacuum and one particle states. This information is contained in the F.T. of the two point functions

$$
\begin{aligned}
\int d^4x e^{-ix\cdot q}\,\langle 0|T[\Omega_a(x)\Omega_b^\dagger(0)]\,|0\rangle = & \\
& \sum_{\lambda,p}\langle 0|\,\Omega_a(0)\,|\lambda,p\rangle\,\langle \lambda,p|\,\Omega_b^\dagger(0)]\,|0\rangle\,(2\pi)^3\delta(\mathbf{p}-\mathbf{q})\frac{-i}{p^0-q^0-i\epsilon} \\
& \pm\sum_{\bar\lambda,p}\langle 0|\,\Omega_b^\dagger(0)\,|\bar\lambda,p\rangle\,\langle \bar\lambda,p|\,\Omega_a(0)]\,|0\rangle\,(2\pi)^3\delta(\mathbf{p}+\mathbf{q})\frac{-i}{p^0+q^0-i\epsilon}
\end{aligned}
$$

The sums over states include all states with any number of particles. But the single particle states are singled out by their unique values of energy for fixed three momentum. Thus only these states will produce poles in the variable q^0. (Multiparticle states produce cuts). Single particle contributions in the first term produce poles at positive values $q^0=+\omega_c(\mathbf{q})$ whereas in the second term they produce poles at negative values $q^0=-\omega_c(\mathbf{q})$. The states contributing to the second term are the charge conjugates (antiparticles) of those contributing to the second term. Focusing on the positive energy poles, we see that the single particle c gives a pole with structure

$$
\frac{-i}{q^2+m_c^2-i\epsilon}(2\pi)^3 2\omega(\mathbf{q})\,\langle 0|\,\Omega_a(0)\,|c,\mathbf{q}\rangle\,\langle c,\mathbf{q}|\,\Omega_b^\dagger(0)]\,|0\rangle\,.
$$

18. Feynman Rules for QED

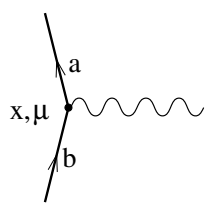
18.1. Rules in Coordinate Space

Propagators

$$S_F(x-y)_{ab} = -i\int \frac{d^4p}{(2\pi)^4}\frac{m\delta_{ab}-\gamma_{ab}\cdot p}{m^2+p^2-i\epsilon}e^{i(x-y)\cdot p}.$$

$$D_{F\mu\nu} = -i\int \frac{d^4k}{(2\pi)^4}\frac{\eta_{\mu\nu}-(1-\alpha)(k_\mu k_\nu/k^2)}{k^2-i\epsilon}e^{i(x-y)\cdot k}$$

Vertices



$$iQ_0\gamma^\mu_{ab}\int d^4x.$$

Rules for Calculating $\langle\psi(x_1)\cdots\psi(x_n)\bar\psi(y_n)\cdots\bar\psi(y_1)A(z_1)\cdots A(z_m)\rangle$

1. Draw all possible graphs connecting together the points (x_i, y_i, z_i). Drop all disconnected vacuum bubbles. Associate with each graph the product of propagators and vertices according to the above table. Integrate over all internal points.
2. Each distinct graph has a weight ± 1 as follows:
   (a) For each closed fermion loop include a factor (-1).
   (b) Two graphs of identical structure except for a permuatation of the x_i's or of the y_i's have a relative minus (plus) sign if the total permutation is odd (even).

Rules for Scattering Amplitudes

1. Drop all corrections to all external lines.
2. Replace propagators associated with external lines by the following factors:

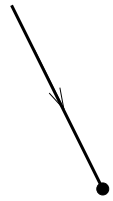
Outgoing Electron:

$$\frac{\sqrt{Z_2}}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}\bar u_\lambda(\mathbf{p})e^{-iy\cdot p}$$

Incoming Electron:

$$\frac{\sqrt{Z_2}}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}u_\lambda(\mathbf{p})e^{iy\cdot p}$$

Outgoing Positron

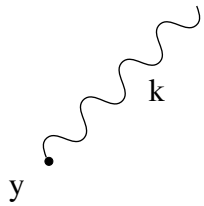


$$\frac{\sqrt{Z_2}}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}v_\lambda(\mathbf{p})e^{-iy\cdot p}$$

Incoming Positron:

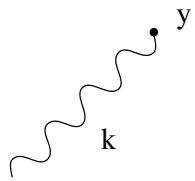
$$\frac{\sqrt{Z_2}}{(2\pi)^{3/2}\sqrt{2\omega(\mathbf{p})}}\bar{v}_\lambda(\mathbf{p})e^{iy\cdot p}$$

Outgoing Photon:



$$\frac{\sqrt{Z_3}}{(2\pi)^{3/2}2|\mathbf{k}|}\epsilon^*_\lambda(\mathbf{k})e^{-iy\cdot k}$$

Incoming Photon:



$$\frac{\sqrt{Z_3}}{(2\pi)^{3/2}2|\mathbf{k}|}\epsilon_\lambda(\mathbf{k})e^{iy\cdot k}$$

3. Z_2 and Z_3 are obtained from examining the poles in the F.T. of the two point functions:

$$\int d^4x e^{-iq\cdot x}\,\langle 0|\,T[\psi(x)\bar{\psi}(0)]\,|0\rangle \quad \underset{q^2\to -m^2}{\sim} Z_2\frac{-i(m-\gamma\cdot q)}{q^2+m^2}$$

$$\int d^4x e^{-iq\cdot x}\,\langle 0|\,T[A_\mu(x)A_\nu(0)]\,|0\rangle \quad \underset{q^2\to 0}{\sim} Z_3\frac{-i(\eta_{\mu\nu}-(1-\alpha')q_\mu q_\nu/q^2))}{q^2}$$

Note that $\alpha' \neq \alpha$ in general. Also Z_3 is gauge invariant whereas Z_2 is not.

<u>Cross Sections and Decay Rates</u>

Let $\mathcal{M}$ be the amplitude obtained from the above rules by dropping the factors $1/[(2\pi)^{3/2}\sqrt{2E}]$ associated with the external lines and also dropping the overall momentum conserving delta function factor $(2\pi)^4\delta(\sum_f p'_f - \sum_i p_i)$. Then the differential cross section for $2 \to N$ particle scattering is:

$$d\sigma = \frac{d^3p'_1\cdots d^3p'_N}{(2\pi)^3 2E'_1\cdots(2\pi)^3 2E'_N}(2\pi)^4\delta(\sum_f p'_f - \sum_i p_i)\frac{1}{4E_1E_2|\mathbf{v}_2-\mathbf{v}_1|}|\mathcal{M}|^2.$$

Here $\mathbf{v}_i$ are the velocities of the particles in the initial state. The initial state momenta are p_1, p_2 and the final state momenta all carry primes.

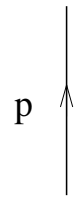
For a decay process $1 \to N$ the differential decay rate in the rest frame of the decaying particle is

$$d\Gamma = \frac{d^3p'_1\cdots d^3p'_N}{(2\pi)^3 2E'_1\cdots(2\pi)^3 2E'_N}(2\pi)^4\delta(\sum_f p'_f - p)\frac{1}{2M}|\mathcal{M}|^2,$$

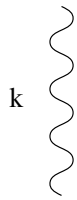
where $p = (M, \mathbf{0})$.

To obtain total cross sections and total rates one must integrate these expressions over all final state momenta and summing over all final spin states. In addition, if some of the final state particles are identical one must include the statistical factor $1/r_k!$ for each subset k of r_k identical particles to compensate for multiple counting of identical states when these final state sums are carried out independently.

Propagators

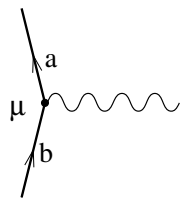


$$S_F(p)_{ab} = -i\frac{m\delta_{ab} - \gamma_{ab}\cdot p}{m^2 + p^2 - i\epsilon}.$$



$$D_{F\mu\nu}(k) = -i\frac{\eta_{\mu\nu} - (1-\alpha)(k_\mu k_\nu/k^2)}{k^2 - i\epsilon}$$

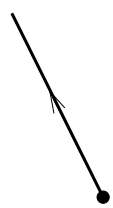
Vertices



$$iQ_0\gamma^\mu_{ab}.$$

1. Each line carries a momentum obeying the constraint that energy-momentum is conserved at each vertex.
2. The rules for drawing graphs are identical to those in coordinate space.
3. Each unconstrained internal momentum is integrated with weight $d^4p/(2\pi)^4$.
4. The external line factors for $\mathcal{M}$ are simply

Outgoing Electron:

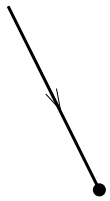


$$\sqrt{Z_2}\bar{u}_\lambda(\mathbf{p})$$

Incoming Electron:

$\sqrt{Z_2}u_\lambda(\mathbf{p})$

Outgoing Positron

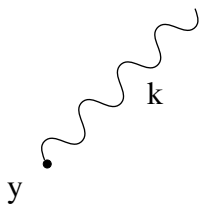


$\sqrt{Z_2}v_\lambda(\mathbf{p})$

Incoming Positron:

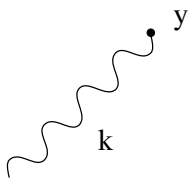
$\sqrt{Z_2}\bar{v}_\lambda(\mathbf{p})$

Outgoing Photon:



$\sqrt{Z_3}\epsilon^*_\lambda(\mathbf{k})$

Incoming Photon:

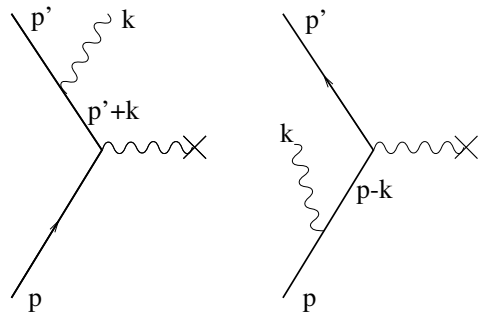


$\sqrt{Z_3}\epsilon_\lambda(\mathbf{k})$

19. Soft Bremstrahlung and Infrared divergences

19.1. One Soft Photon

Let us consider the emission of an extra photon in the scattering of an electron in an external potential. To regulate infrared divergences we temporarily introduce a small photon mass λ. To lowest order in the external potential the diagrams are:



Writing the Feynman amplitude for this process $\mathcal{M} = Q_0 \mathcal{M}^\nu \tilde{A}_\nu(p' + k - p)$, we have

$$\begin{aligned}\mathcal{M}^\nu =& iQ_0 \bar{u}' \left[\epsilon^* \cdot \gamma \frac{m - \gamma \cdot (p' + k)}{m^2 + (p' + k)^2} \gamma^\nu + \gamma^\nu \frac{m - \gamma \cdot (p - k)}{m^2 + (p - k)^2} \epsilon^* \cdot \gamma \right] u \\ =& iQ_0 \bar{u}' \left[\frac{2\epsilon^* \cdot p' - \epsilon^* \cdot \gamma \gamma \cdot k}{2p' \cdot k - \lambda^2} \gamma^\nu + \gamma^\nu \frac{2\epsilon^* \cdot p + \gamma \cdot k \epsilon^* \cdot \gamma)}{-2p \cdot k - \lambda^2} \right] u.\end{aligned}$$

In this chapter we are mainly interested in the case of soft photon emission, since that is intimately involved with the cancellation of infrared divergences in physical processes. In the limit of very small k, λ, the terms involving three gamma matrices are negligible and $\mathcal{M}^\nu$ becomes a numerical momentum dependent factor

$$Q_0 \left[\frac{2\epsilon^* \cdot p'}{2p' \cdot k - \lambda^2} - \frac{2\epsilon^* \cdot p}{2p \cdot k + \lambda^2} \right] \equiv Q_0 \epsilon^* \cdot J$$

times the lowest order elastic scattering amplitude. If the soft photon is unobserved and its energy is smaller than the energy resolution Δ, one must include the contribution of the emission when calculating the cross section for electron scattering. In this case we must obviously sum over polarization and integrate over photon angles. The soft photon contribution to the differential cross section is then approximately given by

$$\frac{d\sigma}{d\Omega}_{1\ soft\ photon} \approx \frac{d\sigma}{d\Omega}_{Elastic} Q_0^2 \int\limits_{|\mathbf{k}| \leq \Delta} \frac{d^3k}{(2\pi)^3 2\omega} \left(J \cdot J + \frac{(k_\mu \cdot J)^2}{\lambda^2} \right),$$

where we used the identity

$$\sum_{Pol} \epsilon_\mu \epsilon_\nu^* = \eta_{\mu\nu} + \frac{k_\mu k_\nu}{\lambda^2}$$

to evaluate the sum over polarizations.

Clearly $J \cdot J = O(1/k^2)$ for small k. On the other hand, $k \cdot J = O(\lambda^2/k)$, so the second term in the integrand is a factor of λ^2 smaller than the first, and can be dropped. Moreover, the smallest in magnitude

$k \cdot p$ or $k \cdot p'$ can be is $m\lambda$ so the λ^2 in the denominators of the expression for J can be dropped, leading to

$$J \cdot J \approx -\frac{2p' \cdot p}{p' \cdot kp \cdot k} - \frac{m^2}{(p' \cdot k)^2} - \frac{m^2}{(p \cdot k)^2}.$$

Since the last two terms are special cases of the first, we only need to evaluate one integral:

$$K \equiv \int\limits_{|\mathbf{k}| \leq \Delta} \frac{d^3k}{(2\pi)^3 2\omega} \frac{1}{p' \cdot kp \cdot k}.$$

To extract the divergent contribution of this integral, first scale $|\mathbf{k}| \to \lambda|\mathbf{q}|$. In the new variables $|\mathbf{q}| \leq \Delta/\lambda$. The divergence as $\lambda \to 0$ then comes from the integration region $q_0 \leq |\mathbf{q}| \leq \Delta/\lambda$, where $q_0 >> 1$ is fixed. For this region the q dependence is simply $1/q^3$ so the q integral just gives a factor $\ln(\Delta/q_0\lambda)$. Thus we have

$$K = \text{Finite} \quad + \ln(\Delta/\lambda)\frac{1}{16\pi^3}\int d\Omega_{\hat{\mathbf{k}}} \frac{1}{(\mathbf{p}' \cdot \hat{\mathbf{k}} - E')(\mathbf{p} \cdot \hat{\mathbf{k}} - E)}.$$

Define

$$I(p', p) \equiv \int d\Omega_{\hat{\mathbf{k}}} \frac{1}{(\mathbf{p}' \cdot \hat{\mathbf{k}} - E')(\mathbf{p} \cdot \hat{\mathbf{k}} - E)}.$$

Then

$$\frac{d\sigma}{d\Omega}_{1\ soft\ photon} \approx \text{Finite}$$
$$+ \ln(\Delta/\lambda)\frac{d\sigma}{d\Omega}_{Elastic}\frac{Q_0^2}{16\pi^3}[-2p' \cdot pI(p', p) - m^2 I(p', p') - m^2 I(p, p)].$$

The integral defining I can be simplified by combining the denominators with the Feynman trick, and then choosing the z-axis for the polar angles to be parallel to the vector $x\mathbf{p}' + (1-x)\mathbf{p}$. Then I reduces to

$$I(p', p) = 4\pi \int\limits_0^1 dx \frac{1}{(xE' + (1-x)E)^2 - (x\mathbf{p}' + (1-x)\mathbf{p})^2}.$$

The denominator is the Lorentz scalar $-(xp' + (1-x)p)^2$ which evaluates to $(x^2 + (1-x)^2)m^2 - 2x(1-x)p' \cdot p$. Using $q^2 \equiv (p' - p)^2 = -2m^2 - 2p' \cdot p$, we then obtain

$$I(p', p) = 4\pi \int\limits_0^1 dx \frac{1}{m^2 + x(1-x)q^2},$$

from which we see that I depends only on q^2. Taking account of this information, we can write

$$\frac{d\sigma}{d\Omega}_{1\ soft\ photon} \approx \text{Finite} \quad + \ln(\Delta/\lambda)\frac{d\sigma}{d\Omega}_{Elastic}\frac{\alpha}{4\pi^2}[q^2 I(q^2) + 2m^2(I(q^2) - I(0))].$$

What we have computed is the contribution to the scattering cross section for electron scattering with the emission of an unobservable soft photon. Adding this to the contribution with no photon emission gives to this order the total electron scattering cross section

$$\frac{d\sigma}{d\Omega} \approx \alpha \times \text{Finite} \quad + \frac{d\sigma}{d\Omega}_{Elastic}\left(1 + \ln(\Delta/\lambda)\frac{\alpha}{4\pi^2}[q^2 I(q^2) + 2m^2(I(q^2) - I(0))]\right).$$

The unobserved photon emission contribution seems to make this cross section diverge for $\lambda \to 0$. However, in this limit the elastic cross section is not directly measurable, since soft photon emission can not even in principle be experimentally vetoed. It is only a physical quantity (in principle) for a non zero photon mass. We shall find that the one loop radiative correction to the elastic amplitude has a divergence as $\lambda \to 0$ which precisely cancels the one in the above formula.

To see this cancellation, let's examine the one loop vertex function in the presence of a finite photon mass.

$$\Gamma^\mu(p',p) - \gamma^\mu = -iQ_0^2 \int \frac{d^4k}{(2\pi)^4} \frac{\gamma_\rho(m-(p'-k)\cdot\gamma)\gamma^\mu(m-(p-k)\cdot\gamma)\gamma^\rho}{(k^2+\lambda^2-i\epsilon)(k^2-2k\cdot p-i\epsilon)(k^2-2k\cdot p'-i\epsilon)}.$$

We are interested in the infrared divergence in this expression when $\lambda \to 0$. Since we want to present the result in a form suitable for comparison to the photon emission amplitude, it is appropriate to first evaluate the k^0 integration by contours. Examination of the denominators reveals six simple poles at the following values of k^0:

$$\pm(\sqrt{\mathbf{k}^2+\lambda^2} - i\epsilon) \qquad E \pm \sqrt{E^2+\mathbf{k}^2-2\mathbf{k}\cdot\mathbf{p}-i\epsilon} \qquad E' \pm \sqrt{E'^2+\mathbf{k}^2-2\mathbf{k}\cdot\mathbf{p}'-i\epsilon}.$$

Three poles are in the upper half plane and three in the lower half plane. We can close the contour up or down, but the best choice is down in the lower half plane. The reason is that only one of the poles in the lower plane is near the origin for small k, the region responsible for the infrared divergence. The other two poles are far from the origin at small k and it is simple to see that their residues do not contribute to the divergence. Thus closing in the lower half plane we only need to consider the residue of the pole at $k^0 = +\sqrt{\mathbf{k}^2+\lambda^2} - i\epsilon$.

$$\Gamma^\mu(p',p) - \gamma^\mu \approx Q_0^2 \int \frac{d^3k}{(2\pi)^3 2\omega} \frac{\gamma_\rho(m-p'\cdot\gamma)\gamma^\mu(m-p\cdot\gamma)\gamma^\rho}{(-\lambda^2-2k\cdot p)(-\lambda^2-2k\cdot p')},$$

where we have dropped terms linear in k in the numerator, since they won't contribute to the infrared divergence. When the vertex is sandwiched between on-shell spinors, the numerator simplifies to $4p'\cdot p$ by moving $\gamma\cdot p'$ to the left and $\gamma\cdot p$ to the right. To extract the divergent part as $\lambda \to 0$ we compute the integral over the region $|\mathbf{k}| \le k_0$ for some fixed k_0. This integral is identical to the one encountered in the soft photon calculation, so by comparison, we find

$$\Gamma^\mu(p',p) \approx \gamma^\mu \left(1 + \frac{Q_0^2}{2(2\pi)^3} p\cdot p' \ln\frac{k_0}{\lambda} I(q^2)\right).$$

Finally, we have to separate from the vertex correction the part that should be absorbed into $1/Z_1$. This is just the value of the correction at $p' = p$, so we write (Recall $2p\cdot p' = -q^2 - 2m^2$.)

$$\Gamma^\mu(p',p) = \frac{1}{Z_1}\gamma^\mu \left(1 - \frac{\alpha}{8\pi^2}\ln\frac{k_0}{\lambda}[q^2 I(q^2) + 2m^2(I(q^2)-I(0))]\right) + \alpha \times \text{Finite } + O(\alpha^2).$$

Of course, defining Z_1 this way makes it depend on the infrared cutoff. In any regulation scheme respecting the Ward Identity, it will turn out that $Z_1 = Z_2$ so that charge renormalization will **not** be infrared sensitive, even though Z_1, Z_2 separately are. The elastic electron cross section involves the square of Γ, and since the infrared divergence occurs at order α, it is correct to this order to extract the divergent λ dependence of the elastic differential cross section as an overall factor:

$$\begin{aligned}\frac{d\sigma}{d\Omega}_{Elastic} &= \frac{d\sigma}{d\Omega}_{IR\ Finite} \left((1 - \frac{\alpha}{8\pi^2}\ln\frac{k_0}{\lambda}[q^2 I(q^2) + 2m^2(I(q^2)-I(0))]\right)^2 + O(\alpha^2) \\ &= \frac{d\sigma}{d\Omega}_{IR\ Finite} \left((1 - \frac{\alpha}{4\pi^2}\ln\frac{k_0}{\lambda}[q^2 I(q^2) + 2m^2(I(q^2)-I(0))]\right) + O(\alpha^2).\end{aligned}$$

Inserting this information into the expression for the electron scattering cross section, we see that the sensitivity to the small photon mass disappears:

$$\frac{d\sigma}{d\Omega} \approx \alpha \times \text{Finite} \quad + \frac{d\sigma}{d\Omega}_{IR\ Finite} \left(1 + \ln(\Delta/k_0)\frac{\alpha}{4\pi^2}[q^2 I(q^2) + 2m^2(I(q^2)-I(0))]\right).$$

Notice, however that the physical effect of soft photons remains, signaled by the unavoidable sensitivity

of the electron scattering cross section to the experimental energy resolution Δ. There is no "infrared catastrophe," *i.e.* the theory gives perfectly finite predictions for all physical measurements.