8.323: QFT1 Lecture Notes

Joseph A. Minahan ©

MIT, Spring 2011

Preface

This volume is a compilation of eight installments of notes that I provided for the students who took Relativistic Quantum Field Theory 1 (8.323) at MIT during the spring of 2011. This is a first semester course in quantum field theory for beginning graduate students or advanced undergraduates. There were 26 lectures and the topics covered were free scalar fields, path integrals scalar ϕ^4-theory, free Dirac fields, scattering and QED. Each chapter of this volume corresponds to one of the installments and covers roughly 3 or 4 lectures worth of material.

As for any set of lecture notes, the reader has to take them as is. While I have located many typos, I am sure that many more still exist. Probably some things could have been better explained and if I ever teach such a course again I will likely go back and reformulate many things.

I would like to thank Prof. Krishna Rajagopal for giving me the opportunity to teach this course. I would also like to thank the 8.323 students who provided many helpful comments about the notes.

Contents

Chapter 1

Free scalar fields

1.1 Why Quantum Field Theory

Traditionally, quantum field theory (QFT) has been defined as the combination of special relativity with quantum mechanics. This is not wrong, but it slightly misses the point. QFT has become the essential tool of particle physics, but it is also highly relevant in condensed matter physics which is nonrelativistic.

By combining special relativity with quantum mechanics, one ends up with a quantum mechanical system with an infinite number of degrees of freedom. It is the infinite degrees of freedom which is responsible for QFT's features. This is is the more modern understanding of QFT and it is relevant for condensed matter physics because here the number of degrees of freedom is very large and can be more or less treated as infinite.

Having an infinite number of degrees of freedom turns out to be very tricky and it took several decades before the situation was well understood. The problem is that many calculations that you can do, and which we will explicitly perform in this class, are infinite. One needs a way to consistently get rid of the infinities so that we are left with finite physical quantities. In this course we will show how to do this at the "one-loop" level. A more detailed explanation will be given in the second semester course.

Even though QFT is applicable to nonrelativistic systems, our emphasis will be on scalar field theory and quantum electrodynamics (QED), which are explicitly relativisitic. QFT as applied to QED is a theoretical triumph, where theory has been shown to match with experiment up to 12 significant digits. But QFT remains incomplete and is still a very active field of theoretical research. Here we list several issues still to be resolved in QFT

- While QFT "works" extremely well, there are parameters in the theory that need to be set by hand, namely coupling strengths and masses. In QFT 2 & 3 you will learn how the masses are themselves determined by couplings to the so-called "Higgs" field, but what determines these couplings is not well understood. Physicists are always looking for a more complete theory where the couplings are somehow determined.

- At high enough energies QED breaks down and some other theory must take over. Actually, at a scale well below this breakdown, QED unifies with the weak force to form the electroweak force, but even this theory must break down at high enough energies. In principle, quantum chromodynamics (QCD), the theory of the strong force, could be consistent at an arbitrarily high scale because of a property called "asymptotic freedom", where the coupling becomes weaker at higher energy scales. One possibility to prevent a breakdown of the electroweak theory is that it is unified with QCD to make a grand unified theory that is also asymptotically free.
- Even if there is a grand unified theory, this theory is still missing gravity. It turns out that merging gravity with quantum mechanics is a very difficult business and has not been truly resolved, because of the aforementioned infinities. It is generally accepted that string theory can get around this problem and lead to a fully consistent theory.

1.2 Conventions

It is assumed that you have had an extensive course or subcourse in special relativity. In this section we give our conventions which we use throughout the course. They match those found in Peskin.

We will use units where the speed of light is $c = 1$. Hence, time has units of length. Likewise, from the relation

$$E^2 = p^2c^2 + m^2c^4\,, \tag{1.2.1}$$

we see that mass and momentum have the same units as energy. Space-time coordinates in a $3 + 1$ dimensional inertial frame $\mathbf{S}$ are given by x^μ where $\mu = 0, 1, 2, 3$ with $x^0 = t$. Now and then we will consider situations where there are $d + 1$ space-time dimensions with the μ labeled accordingly.

We will use the $(+---)$ convention for the metric $\eta_{\mu\nu}$, where

$$\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)\,. \tag{1.2.2}$$

The invariant length squared for a space-time displacement Δx^μ is

$$\Delta s^2 = -\Delta\tau^2 = -\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu = -\Delta x^\mu \Delta x_\mu \equiv -\Delta x^2\,, \tag{1.2.3}$$

where $\Delta\tau$ is the displaced proper time. Infinitesimal displacements dx^μ have the invariant length squared $ds^2 = -\eta_{\mu\nu}\, dx^\mu dx^\nu$. Δx^μ and dx^μ are examples of contravariant 4-vectors.

Under a Lorentz transformation from an inertial frame $\mathbf{S}$ to another inertial frame $\mathbf{S}'$ with coordinates $x^{\mu'}$, the coordinates transform as $\Delta x^{\mu'} = \Lambda^{\mu'}{}_\nu \Delta x^\nu$. Lorentz transformations have six independent generators given by boosts in the three spatial directions and three independent rotations. For example, a boost in the x^1 direction has Λ given by

$$\Lambda = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\,, \tag{1.2.4}$$

where $\gamma^{-1} = \sqrt{1-v^2}$. Δs^2 and ds^2 are invariant under these transformations. Note that for any Lorentz transformation, $\det(\Lambda) = 1$.

We can enlarge the set of transformations to also include constant shifts $x^\mu \to x^\mu + a^\mu$, where the a^μ are constants. The displacements and the differentials are clearly invariant under these transformations. The combination of space-time translations and Lorentz transformations are called Poincaré transformations.

The Lorentz transformations on contravariant vectors can be generalized to transformations on tensors. An $\left(\begin{array}{c} n \\ m \end{array}\right)$ tensor has the form $T^{\mu_1\mu_2\dots\mu_n}_{\nu_1\nu_2\dots\nu_m}$ and transforms under a Lorentz transformation as

$$T^{\mu'_1\mu'_2\dots\mu'_n}_{\nu'_1\nu'_2\dots\nu'_m} = \Lambda^{\mu'_1}{}_{\mu_1}\Lambda^{\mu'_2}{}_{\mu_2}\dots\Lambda^{\mu'_n}{}_{\mu_n}\Lambda^{\nu_1}{}_{\nu'_1}\Lambda^{\nu_2}{}_{\nu'_2}\dots\Lambda^{\nu_m}{}_{\nu'_m}T^{\mu_1\mu_2\dots\mu_n}_{\nu_1\nu_2\dots\nu_m}\,, \tag{1.2.5}$$

Indices can be raised and lowered with the metric, with

$$T^{\mu_1\mu_2\dots\mu_n\lambda}_{\nu_1\nu_2\dots\nu_m} = \eta^{\lambda\sigma}T^{\mu_1\mu_2\dots\mu_n}_{\nu_1\nu_2\dots\nu_m\sigma}\,. \tag{1.2.6}$$

We can also make an $\left(\begin{array}{c} n \\ m+1 \end{array}\right)$ tensor from $T^{\mu_1\mu_2\dots\mu_n}_{\nu_1\nu_2\dots\nu_m}$ by taking a derivative

$$T^{\mu_1\mu_2\dots\mu_n}_{\nu_1\nu_2\dots\nu_m,\sigma} \equiv \frac{\partial}{\partial x^\sigma}T^{\mu_1\mu_2\dots\mu_n}_{\nu_1\nu_2\dots\nu_m} \equiv \partial_\sigma T^{\mu_1\mu_2\dots\mu_n}_{\nu_1\nu_2\dots\nu_m}\,. \tag{1.2.7}$$

We will often consider the inner product of two contravariant 4-vectors, A^μ and B^μ,

$$A\cdot B \equiv A^\mu B_\mu\,, \tag{1.2.8}$$

which has no free indices and thus is a Lorentz invariant.

One particular tensor is the 4-momentum of a particle $p_\mu = (E, -\vec{p})$. The invariant made from this is

$$p^2 \equiv p_\mu p_\nu \eta^{\mu\nu} = p_\mu p^\mu = m^2\,. \tag{1.2.9}$$

Another tensor is the gauge field of an electromagnetic field $\mathcal{A}_\mu(x)$. The field strength $\mathcal{F}_{\mu\nu}$ is

$$\mathcal{F}_{\mu\nu}(x) = \partial_\mu\mathcal{A}_\nu(x) - \partial_\nu\mathcal{A}_\mu(x)\,. \tag{1.2.10}$$

Quantization relates the 4-momentum p_μ to a 4-wave vector $k_\mu = (\omega, -\vec{k})$ by $p_\mu = \hbar\, k_\mu$. We will choose units where $\hbar = 1$, although we will occasionally make the $\hbar$ explicit in our equations. With these units a length will have the dimension of an inverse momentum which translates into an inverse mass when $c = 1$. With our choice of units, all physical quantities have units which are mass to some power, D. We will call D the quantity's "dimension".

1.3 Free scalar fields

Consider a real classical scalar field $\phi(x)$ where 'x' in the argument refers to all 3+1 space-time coordinates. A scalar field is invariant under Poincaré transformations, meaning that an observer in an inertial frame $\mathbf{S}'$ will see the scalar field $\phi'(x')$ where $\phi'(x') = \phi(x)$. We will want our field to be dynamical, meaning that it should satisfy some second order differential equation with respect to the time coordinate. Furthermore, we will want the differential equation to be covariant, so that if $\phi(x)$ is a solution for an observer in $\mathbf{S}$ then $\phi'(x')$ is a solution for an observer in $\mathbf{S}'$. Therefore, the derivatives in the equation have to come with the combination $\partial^2 = \partial_\mu \partial^\mu = \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2$.

We will also assume that ϕ is a solution to a linear equation, which we will later see is relevant for free particles, that is, particles that don't interact with other particles. The simplest such equation we can write down is the *Klein-Gordon equation*,

$$\partial^2 \phi(x) + m^2 \phi(x) = 0\,. \tag{1.3.1}$$

Since a derivative has dimension 1, consistency requires that the parameter m is also dimension 1. For this reason we will refer to m as a mass.

A general solution to (1.3.1) is

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\,\omega(\vec{k})}} \left(A_{\vec{k}}\, e^{-ik\cdot x} + A^*_{\vec{k}}\, e^{+ik\cdot x} \right) \tag{1.3.2}$$

where $A_{\vec{k}}$ are constants with respect to the coordinates x^μ and $k^0 = \omega(\vec{k}) = \sqrt{\vec{k}\cdot\vec{k} + m^2}$. The factor of $\sqrt{2\omega(\vec{k})}$ has been inserted for later convenience. Let us now consider the Fourier transform of $\phi(x)$ with respect to the three spatial directions,

$$\widetilde{\phi}(\vec{k}, t) = \int d^3x\, e^{-i\vec{k}\cdot\vec{x}}\, \phi(x)\,. \tag{1.3.3}$$

Comparing with (1.3.2), we find that

$$\widetilde{\phi}(\vec{k}, t) = \frac{1}{\sqrt{2\,\omega(\vec{k})}} \left(A_{\vec{k}}\, e^{-i\omega(\vec{k})t} + A^*_{-\vec{k}}\, e^{+i\omega(\vec{k})t} \right)\,, \tag{1.3.4}$$

where $\omega(\vec{k})$ is defined as above. This does not look relativistic, but nonetheless, let us press on. Substituting $\tilde{\phi}(\vec{k}, t)$ into (1.3.1) we immediately find the equation

$$\frac{d^2}{dt^2}\, \widetilde{\phi}(\vec{k}, t) + \omega^2(\vec{k})\, \widetilde{\phi}(\vec{k}, t) = 0\,. \tag{1.3.5}$$

Hence, for every $\vec{k}$ we have an equation for an ordinary harmonic oscillator with frequency $\omega(\vec{k})$.

1.3.1 Quick review of the simple harmonic oscillator

Since our free classical field can be thought of as an infinite collection of harmonic oscillators, let us review how one quantizes the simple harmonic oscillator. Given an oscillator whose position is $x(t)$, its mass is m and its frequency is ω_0, the Hamiltonian is given by

$$H = \tfrac{1}{2}\, m\, \dot{x}^2 + \tfrac{1}{2}\, m\, \omega_0^2\, x^2\,. \tag{1.3.6}$$

Since the mass is a common factor, it is convenient to define a new variable $\phi(t) = \sqrt{m}\, x(t)$ so that the Hamiltonian becomes

$$H = \tfrac{1}{2}\dot{\phi}^2 + \tfrac{1}{2}\,\omega_0^2\,\phi^2\,, \tag{1.3.7}$$

and the equation of motion is

$$\frac{\partial^2}{\partial t^2}\phi(t) + \omega_0^2\,\phi(t) = 0\,. \tag{1.3.8}$$

Comparing to (1.3.1), we see that we can interpret the simple harmonic oscillator as a scalar field in $0+1$ dimensions. Notice further that the dimension of $\phi(t)$ is $D = -1/2$. A general solution to the equation of motion can be written as

$$\phi(t) = \frac{1}{\sqrt{2\omega_0}}\left(A\, e^{-i\omega_0 t} + A^*\, e^{+i\omega_0 t}\right)\,. \tag{1.3.9}$$

The action for this system is

$$\mathcal{S} = \int dt\, L(t)\,, \tag{1.3.10}$$

where $L(t)$ is the Lagrangian

$$L = \tfrac{1}{2}\dot{\phi}^2 - \tfrac{1}{2}\omega_0^2\,\phi^2\,. \tag{1.3.11}$$

The canonical momentum $\Pi(t)$ is then given by

$$\Pi(t) = \frac{\partial\, L}{\partial\dot{\phi}(t)} = \dot{\phi}(t)\,, \tag{1.3.12}$$

whose solutions are

$$\Pi(t) = -i\sqrt{\frac{\omega_0}{2}}\left(A\, e^{-i\omega_0 t} - A^*\, e^{+i\omega_0 t}\right)\,. \tag{1.3.13}$$

Now let us quantize this system. In this case ϕ becomes an operator that acts on the Hilbert space and whose time evolution is given by

$$\phi(t) = e^{iHt}\phi(0)\, e^{-iHt}\,. \tag{1.3.14}$$

The constants A and A^* become the annihilation and creation operators a and $a^\dagger$ which satisfy the commutation relation

$$[a, a^\dagger] = 1\,. \tag{1.3.15}$$

It then follows that the equal time commutation relations between $\phi(t)$ and $\Pi(t)$ have the standard form

$$[\phi(t), \Pi(t)] = i\,. \tag{1.3.16}$$

The normalized eigenstates of the system are built from the ground state $|0\rangle$, where $a|0\rangle = 0$. Thus, we have $|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle$.

1.3.2 2-point correlators for the harmonic oscillator

Of special interest are correlators of operators at different times[1]. In particular, the two-point function is given by

$$G_F(t, t') = \langle 0|T[\phi(t)\phi(t')]|0\rangle\,, \tag{1.3.17}$$

where the F subscript stands for Feynman and the symbol T stands for "time-ordered",

$$\begin{aligned} T[\phi(t)\phi(t')] &= \phi(t)\phi(t') \qquad t > t' \\ &= \phi(t')\phi(t) \qquad t < t'\,. \end{aligned} \tag{1.3.18}$$

Substituting (1.3.9) into (1.3.17) we obtain

$$G_F(t, t') = \frac{1}{2\omega_0} e^{-i\omega_0|t-t'|} = G_F(t-t')\,. \tag{1.3.19}$$

To see why this is interesting, let us consider the combination

$$\begin{aligned} \frac{\partial^2}{\partial t^2} G_F(t-t') + \omega_0^2\, G_F(t-t') &= -\omega_0^2\, G_F(t-t') + \omega_0^2\, G_F(t-t') - \frac{i}{2}\frac{\partial}{\partial t}\mathrm{sgn}(t-t') \\ &= -i\,\delta(t-t')\,, \end{aligned} \tag{1.3.20}$$

where $\mathrm{sgn}(t)$ is the sign of t. Therefore, $i\,G_F(t-t')$ is the Green's function for the differential operator $\frac{\partial^2}{\partial t^2} + \omega_0^2$.

It will often be convenient to Fourier transform the operator $\phi(t)$ and $G_F(t-t')$ into frequency space, where

$$\widetilde{\phi}(\omega) = \int dt\, e^{+i\omega t}\phi(t)\,. \tag{1.3.21}$$

For the two-point function, we Fourier transform both t and t' to get

$$\int dt\, dt'\, e^{+i\omega t + i\omega' t'} G_F(t-t') \;=\; \int dT d\tau\, e^{+i(\omega+i\omega')T + i(\omega-\omega')\tau/2} G_F(\tau)\,, \tag{1.3.22}$$

[1] Later in this course we will see that operator correlators are important for particle scattering.

where $T = \frac{1}{2}(t+t')$ and $\tau = t - t'$. Integrating over T then gives

$$\int dt\, dt'\, e^{+i\omega t + i\omega' t'} G_F(t-t') \;=\; 2\pi\,\delta(\omega+\omega') \int d\tau\, e^{+i\omega\tau} G_F(\tau) = 2\pi\,\delta(\omega+\omega')\,\widetilde{G}_F(\omega)\,, \tag{1.3.23}$$

where $\widetilde{G}_F(\omega)$ is the Fourier transform of $G_F(\tau)$,

$$\widetilde{G}_F(\omega) \equiv \int d\tau\, e^{+i\omega\tau} G_F(\tau) = \frac{1}{2\,\omega_0}\int d\tau\, e^{+i\omega\tau - i\omega_0|\tau|}\,. \tag{1.3.24}$$

The δ-function in (1.3.23) ensures energy conservation and is a consequence of the time translation invariance in the Green's function.

Strictly speaking this integral in (1.3.24) is not well-defined, but we can make it so by shifting ω_0 by a small negative imaginary part, $\omega_0 \to \omega_0 - i\,\epsilon$. Looking at $G_F(\tau)$ in (1.3.19) we see that $+i\epsilon$ gradually sends $G_F(\tau)$ to zero for $\tau \to \pm\infty$ which is a usual procedure for distributions: we assume that $G_F(\tau)$ is turned off in the infinite past and future. We then find

$$\begin{aligned}
\widetilde{G}_F(\omega) &= \frac{1}{2\,\omega_0}\left(\int_0^\infty d\tau\, e^{+i(\omega-\omega_0+i\epsilon)\tau} + \int_{-\infty}^0 d\tau\, e^{+i(\omega+\omega_0-i\epsilon)\tau}\right)\\
&= \frac{1}{2\,\omega_0}\left(\frac{i}{\omega-\omega_0+i\epsilon} - \frac{i}{\omega+\omega_0-i\epsilon}\right)\\
&= \frac{i}{\omega^2-\omega_0^2+i\epsilon}\,,
\end{aligned} \tag{1.3.25}$$

where we have absorbed a factor of ω_0 into ϵ (we are assuming that $\omega_0 > 0$).

We can go backward to obtain $G(t-t')$ from $\widetilde{G}(\omega)$, where

$$G_F(t-t') = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, e^{-i\omega(t-t')}\widetilde{G}_F(\omega)\,. \tag{1.3.26}$$

To do this integral we can close off the contour and pick off the contributions from the poles. Note that $\widetilde{G}(\omega)$ has simple poles at $\omega = \pm(\omega_0 - i\epsilon)$, hence one pole is just above the real line and the other is just below it (see figure 1). If $t > t'$ then we can close off the contour in the lower half plane, where we end up encircling the pole clockwise at $\omega_0 - i\epsilon$ to find

$$G_F(t-t') = -(2\pi i)\,\frac{1}{2\pi}\,\frac{i}{2\,\omega_0}\,e^{-i\omega_0(t-t')} = \frac{1}{2\,\omega_0}\,e^{-i\omega_0(t-t')}\,. \tag{1.3.27}$$

If $t < t'$ then we close off the contour in the upper half plane, finding

$$G_F(t-t') = (2\pi i)\,\frac{1}{2\pi}\,\frac{i}{-2\,\omega_0}\,e^{+i\omega_0(t-t')} = \frac{1}{2\,\omega_0}\,e^{+i\omega_0(t-t')}\,. \tag{1.3.28}$$

This is not the only way to get a Green's function. We could also consider the combination

$$G_R(t,t') \equiv \theta(t-t')\langle 0|[\phi(t),\phi(t')]|0\rangle = \frac{\theta(t-t')}{2\,\omega_0}\left(e^{-i\omega_0(t-t')} - e^{+i\omega_0(t-t')}\right) = G_R(t-t')\,, \tag{1.3.29}$$

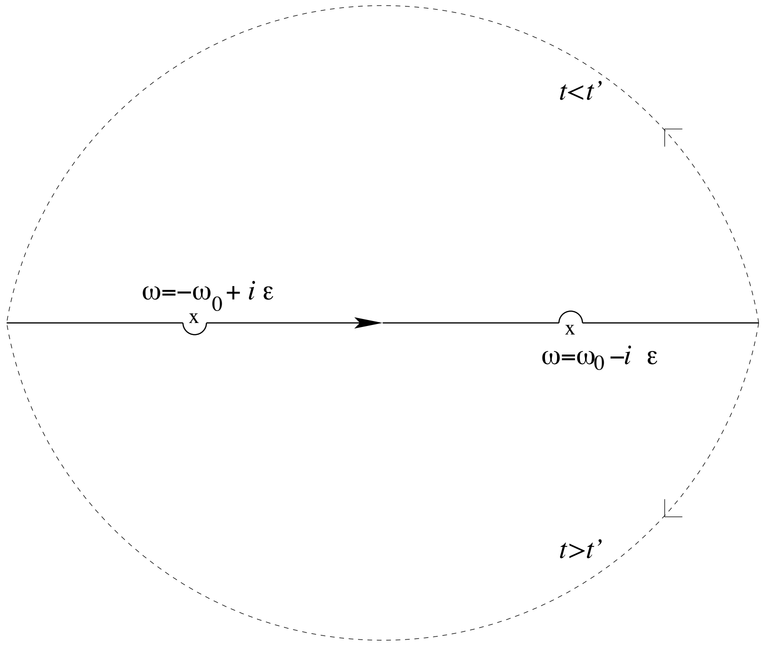


Figure 1.1: Integration contour for $G_F(t-t')$ in the case $t>t'$ (lower) and $t<t'$ (upper).

where $\theta(t)$ is the Heaviside function, $\theta(t)=1$, $t>0$, $\theta(t)=0$, $t<0$. Thus,

$$\begin{aligned}\frac{\partial^2}{\partial t^2}\,G_R(t-t')+\omega_0^2\,G_R(t-t') &= -\omega_0^2\,G_R(t-t')+\omega_0^2\,G_R(t-t')-i\,\frac{\partial}{\partial t}\theta(t-t') \\ &= -i\,\delta(t-t')\,. \end{aligned} \tag{1.3.30}$$

In frequency space we have that

$$\widetilde{G}_R(\omega)=\int_0^\infty d\tau e^{i\omega\tau}\,\frac{1}{2\,\omega_0}\left(e^{-i\omega_0\tau}-e^{+i\omega_0\tau}\right) \tag{1.3.31}$$

In order to do the integral we have to shift $\omega_0\to\omega_0-i\,\epsilon$ in the first term and $\omega_0\to\omega_0+i\,\epsilon$ in the second term. Hence we have

$$\widetilde{G}_R(\omega)=\frac{i}{(\omega-\omega_0+i\,\epsilon)(\omega+\omega_0+i\,\epsilon)}=\frac{i}{\omega^2-\omega_0^2+i\,\epsilon\,\mathrm{sgn}(\omega)}\,. \tag{1.3.32}$$

Clearly, both poles are below the real line (see figure 2).

The Green's function allows us to compute $\phi(t)$ in the presence of a source. Suppose we have a source $J(t)$ for our $0+1$ dimensional scalar field, such that

$$\frac{\partial^2}{\partial t^2}\phi(t)+\omega_0^2\,\phi(t)=J(t)\,. \tag{1.3.33}$$

This equation of motion can be derived from the action by including the term

$$\int dt\,J(t)\phi(t)\,. \tag{1.3.34}$$

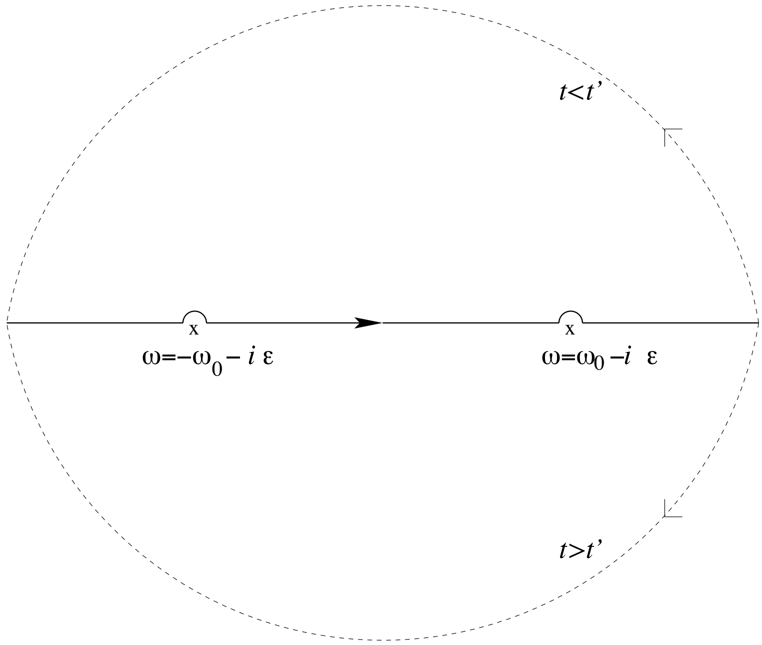


Figure 1.2: Integration contour for $G_R(t-t')$ in the case $t>t'$ (lower) and $t<t'$ (upper).

The solution to (1.3.33) is

$$\phi(t)=\phi_{\text{hom}}(t)+i\int dt'\,G(t-t')\,J(t')\,,\tag{1.3.35}$$

where the homogeneous part is the solution in (1.3.9) and $G(t-t')$ is one of the Green's functions.

By construction, $G_R(t-t')$ only has support when t' is in the past of t, hence this is a retarded Green's function. However, the Feynman Green's function has support both in the future and the past, so this is not retarded (nor is it advanced). To understand its significance, let us rewrite $G_F(t-t')$ as

$$G_F(t-t')=\frac{\theta(t-t')}{2\,\omega_0}\,e^{-i\omega_0(t-t')}-\frac{\theta(t'-t)}{2\,(-\omega_0)}e^{-i(-\omega_0)(t-t')}\,,\tag{1.3.36}$$

which has a form closer to $G_R(t-t')$. Here, however while the first term is the same as the first term in $G_R(t-t')$, the second term is advanced and the coefficient is that for a negative energy. Hence, sources in the past add positive energy modes, in other words they act as emitters. Sources in the future add negative energy modes, which is equivalent to removing positive energy modes, so they act as absorbers.

Let us look at this a different way. Suppose that the system is in the ground state until time t' when it is suddenly put in the first excited state via a source. As a function of t the amplitude is proportional to $e^{-i\omega_0(t-t')}$ if $t>t'$. This is the behavior of $\langle 0|\phi(t)\phi(t')|0\rangle$. But by the symmetry of this correlator there is a source term at t to take the system back down to the ground state. If $t<t'$ then the source at t can only add energy to the system and the source at t' removes it, in which case the correlator is $\langle 0|\phi(t')\phi(t)|0\rangle$. In other words, we should use $G_F(t-t')$ if sources can equally act as emitters or absorbers.

1.3.3 3+1 dimensional scalar field as an infinite set of harmonic oscillators

We now apply the ideas from a single harmonic oscillator to the scalar field in 3 + 1 dimensions. As we showed, the Fourier transformed components $\widetilde{\phi}(\vec{k},t)$ satisfy the equation of motion for an harmonic oscillator whose classical solution is expressed in (1.3.4). We should then proceed as we did for the single harmonic oscillator, although there a couple of wrinkles we need to iron out.

First, unlike $\phi(t)$, $\widetilde{\phi}(\vec{k},t)$ is not real. Hence, after quantization, $\widetilde{\phi}(\vec{k},t)$ becomes

$$\widetilde{\phi}(\vec{k},t)=\frac{1}{\sqrt{2\,\omega(\vec{k})}}\left(a_{\vec{k}}\,e^{-i\omega(\vec{k})t}+a^\dagger_{-\vec{k}}\,e^{+i\omega(\vec{k})t}\right)\,, \tag{1.3.37}$$

where $a_{\vec{k}}$ is the annihilation operator for a mode with momentum $\vec{k}$ and $a^\dagger_{-\vec{k}}$ is the creation operator for a different mode with momentum $-\vec{k}$. Here we can draw an analogy to a circularly symmetric simple harmonic oscillator in two dimensions. In this case we have two coordinates $\phi_1(t)$ and $\phi_2(t)$ which after quantization have the form

$$\begin{aligned}\phi_1(t) &= \frac{1}{\sqrt{2\omega_0}}\left(a_1\,e^{-i\omega_0 t}+a_1^\dagger\,e^{+i\omega_0 t}\right)\\ \phi_2(t) &= \frac{1}{\sqrt{2\omega_0}}\left(a_2\,e^{-i\omega_0 t}+a_2^\dagger\,e^{+i\omega_0 t}\right)\,.\end{aligned} \tag{1.3.38}$$

If we now let $\phi_\pm=\frac{1}{\sqrt{2}}(\phi_1(t)\pm i\,\phi_2(t))$, then we have that

$$\begin{aligned}\phi_+(t) &= \frac{1}{\sqrt{2\omega_0}}\left(a_+\,e^{-i\omega_0 t}+a_-^\dagger\,e^{+i\omega_0 t}\right)\\ \phi_-(t) &= \frac{1}{\sqrt{2\omega_0}}\left(a_-\,e^{-i\omega_0 t}+a_+^\dagger\,e^{+i\omega_0 t}\right)\,,\end{aligned} \tag{1.3.39}$$

where $a_\pm=\frac{1}{\sqrt{2}}(a_1\pm i\,a_2)$, $a_\pm^\dagger=\frac{1}{\sqrt{2}}(a_1^\dagger\mp i\,a_2^\dagger)$.

The other wrinkle is that we now have a continuous set of oscillators, so it will be necessary to modify the commutation relation in (1.3.15). Since the standard measure in three-dimensional momentum space is $\frac{d^3k}{(2\pi)^3}$, it is natural to normalize the relations to

$$[a_{\vec{k}},a^\dagger_{\vec{k}'}]=(2\pi)^3\delta^3(\vec{k}-\vec{k}')\,. \tag{1.3.40}$$

The ground state for a collection of oscillators, $|0\rangle$ satisfies $a_{\vec{k}}|0\rangle=0$ for all $\vec{k}$. The Hamiltonian should be a sum, or in this case an integral, over all the individual oscillators

$$H=\int\frac{d^3k}{(2\pi)^3}\,\tfrac{1}{2}\omega(\vec{k})\left(a^\dagger_{\vec{k}}\,a_{\vec{k}}+a_{\vec{k}}\,a^\dagger_{\vec{k}}\right)=\int\frac{d^3k}{(2\pi)^3}\,\omega(\vec{k})a^\dagger_{\vec{k}}\,a_{\vec{k}}+C\,. \tag{1.3.41}$$

The expression on the right hand side is the normal ordered version where all creation operators lie on the left of the annihilation operators. The constant C is an infinite constant which we can ignore since we will only be interested in relative energies between the states.

From now on we will call the ground state $|0\rangle$ the "vacuum". The states are then constructed by acting with the creation operators $a^\dagger_{\vec{k}}$ on $|0\rangle$. The full Hilbert space for this system is called a *Fock space*. We define the states as follows:

$$|\vec{k}_1, \vec{k}_2 \ldots \vec{k}_n\rangle = \prod_{i=1}^{n} \sqrt{2\,\omega(\vec{k}_i)}\, a^\dagger_{\vec{k}_i}|0\rangle\,, \tag{1.3.42}$$

where we assume that all $\vec{k}_i$ are different, which is a reasonable assumption since $\vec{k}$ is continous. The rather strange normailization factors are to ensure Lorentz invariant inner products, as we will demonstrate below. Acting with H on this state we find

$$H|\vec{k}_1, \vec{k}_2 \ldots \vec{k}_n\rangle = \sum_{i=1}^{n} \omega(\vec{k}_i)|\vec{k}_1, \vec{k}_2 \ldots \vec{k}_n\rangle\,, \tag{1.3.43}$$

while if we define the momentum operator $\vec{P}$

$$\vec{P} \equiv \int \frac{d^3k}{(2\pi)^3}\, \vec{k}\, a^\dagger_{\vec{k}}\, a_{\vec{k}} \tag{1.3.44}$$

we have that

$$\vec{P}|\vec{k}_1, \vec{k}_2 \ldots \vec{k}_n\rangle = \sum_{i=1}^{n} \vec{k}_i\, |\vec{k}_1, \vec{k}_2 \ldots \vec{k}_n\rangle\,. \tag{1.3.45}$$

Hence the state in the Fock space has the energy and momentum of n particles with rest mass m, where the individual momenta are $\vec{k}_i$ and their energies are $k^0 = \sqrt{\vec{k}^2 + m^2}$.

We can now point out several properties of these particles.

- They are free particles, meaning that they are noninteracting. The state $|\vec{k}_1, \vec{k}_2 \ldots \vec{k}_n\rangle$ is an eigenstate of H, hence the individual $\vec{k}_i$ are not changing in time. The only way that the individual momenta cannot change is that there are no forces acting on the particles. Furthermore, since $|\vec{k}_1, \vec{k}_2 \ldots \vec{k}_n\rangle$ is an eigenstate, no particles are being created or destroyed as time evolves.

- The particles have no internal quantum numbers. In particular, they must have spin 0. This is not too surprising since our particles originated from a scalar field which has no Lorentz indices, meaning that it is invariant under rotations.

- The state does not change sign under the exchange of $\vec{k}_i$ and $\vec{k}_j$. According to our interpretation, this switches the single particle states of two particles. Since there is no change in the state, the particles are identical bosons.

In our analysis we have chosen a particular inertial frame **S** and then Fourier transformed the spatial coordinates. An observer in a different frame of course would have a different basis of spatial momenta. However, certain things should be invariant. In particular, the vacuum should be invariant under a Lorentz transformation. Let's show that it is. The vacuum is distinguished by $a_{\vec{k}}|0\rangle = 0$ for all $\vec{k}$. We were able to distinguish between creation operators and annihilation operators because the annihilation operators in $\tilde{\phi}(\vec{k},t)$ came with a factor of $e^{-i\omega(\vec{k})t}$ while the creation operators came with a factor of $e^{+i\omega(\vec{k})t}$, where $\omega(\vec{k}) > 0$. Under a Lorentz transformation $\omega(\vec{k})$ can change, but its sign is invariant. Hence a Lorentz transformation will map annihilation operators to annihilation operators and creation operators to creation operators. Hence the state $|0\rangle$ will continue to be annihilated by all $a_{\vec{k}}$ and so continue to be defined as the vacuum.

Let us now consider the one particle states $|\vec{k}\rangle$. This is not invariant under a Lorentz transformation since the momentum will change. However, a one-particle state for an observer in $\mathbf{S}$ will be a one particle state for an observer in any other frame. Moreover, we would like the inner product between one particle states to be Lorentz invariant. We can quickly check that it is. Consider the Lorentz transformation of the 4-dimensional δ-function

$$(2\pi)^4\,\delta^4(k-q) \to (2\pi)^4\,\delta^4(k'-q') = (2\pi)^4\,\det(\Lambda^{-1})\delta^4(k-q) = (2\pi)^4\,\delta^4(k-q)\,, \tag{1.3.46}$$

where $\Lambda^{\mu'}{}_{\nu}$ is the Lorentz transformation matrix that takes $\mathbf{S}$ to $\mathbf{S}'$. The δ-function $2\pi\,\delta(k^2-q^2)$ is also clearly invariant. When $\vec{k}=\vec{q}$ then this δ-function becomes

$$(2\pi)\delta(k^2-q^2)\Big|_{\vec{k}=\vec{q}} = 2\pi(2k^0)^{-1}\delta(k^0-q^0)\,. \tag{1.3.47}$$

If q^μ is on the mass-shell, meaning that $q^2=m^2$, then the δ-function forces k^μ to also be on the mass-shell and so we can replace $k^0=\omega(\vec{k})$. Hence we have the Lorentz invariant combination

$$\frac{(2\pi)^4\,\delta^4(k-q)}{(2\pi)\,\delta(k^2-q^2)} = 2\omega(\vec{k})(2\pi)^3\delta^3(\vec{k}-\vec{q})\,. \tag{1.3.48}$$

One can quickly check that

$$\langle\vec{q}|\vec{k}\rangle = 2\omega(\vec{k})\,(2\pi)^3\delta^3(\vec{k}-\vec{q})\,, \tag{1.3.49}$$

and so is Lorentz invariant.

Let us now return to the scalar field in coordinate space, which by (1.3.2) is written in terms of the oscillators as

$$\phi(x) = \int\frac{d^3k}{(2\pi)^3}\,\frac{1}{\sqrt{2\,\omega(\vec{k})}}\left(a_{\vec{k}}\,e^{-ik\cdot x} + a^\dagger_{\vec{k}}\,e^{+ik\cdot x}\right)\,. \tag{1.3.50}$$

To find the canonical momentum for this field, we need to know the Lagrangian density $\mathcal{L}(x)$ that gives the Klein-Gordon equation in (1.3.1). Using that the functional derivative satisfies

$$\frac{\delta\phi(x)}{\delta\phi(y)} = \delta^4(x-y)\,, \tag{1.3.51}$$

and that the action is

$$\mathcal{S} = \int d^4x\, \mathcal{L}(x) \tag{1.3.52}$$

we have that the equation of motion comes from varying $\phi(x)$ to $\phi(x) + \delta\phi(x)$ so that

$$\delta\mathcal{S} = 0\,. \tag{1.3.53}$$

We assume that $\mathcal{L}$ is made up of $\phi(x)$ and its derivatives $\partial_\mu\phi(x)$. Hence the variation of $\mathcal{S}$ is

$$\delta\mathcal{S} = \int d^4x \left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\mu\delta\phi + \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi\right) = \int d^4x \left(-\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi} + \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi\right)\delta\phi\,, \tag{1.3.54}$$

where we integrated by parts in the second step. Hence for general $\delta\phi(x)$ we find that the equations of motion follow from the Euler-Lagrange equation

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi} - \frac{\partial\mathcal{L}}{\partial\phi} = 0\,. \tag{1.3.55}$$

Therefore, we obtain the Klein-Gordon equation if

$$\mathcal{L} = \tfrac{1}{2}\partial_\mu\phi\,\partial^\mu\phi - \tfrac{1}{2}\,m^2\,\phi^2\,. \tag{1.3.56}$$

Note that the action $\mathcal{S}$ is dimesionless in any number of space-time dimensions, therefore $\mathcal{L}$ has dimension 4 in 3+1 dimensions and so ϕ has dimension 1.

The canonical momentum $\Pi(x)$ is then

$$\Pi(x) = \frac{\partial\mathcal{L}}{\partial_0\phi}(x) = \partial^0\phi(x)\,. \tag{1.3.57}$$

Therefore, in terms of the oscillators $\Pi(x)$ is given by

$$\Pi(x) = -i\int\frac{d^3k}{(2\pi)^3}\sqrt{\frac{\omega(\vec{k})}{2}}\left(a_{\vec{k}}\,e^{-ik\cdot x} - a^\dagger_{\vec{k}}\,e^{+ik\cdot x}\right)\,. \tag{1.3.58}$$

The equal time commutators are then

$$[\phi(x^0,\vec{x}),\Pi(x^0,\vec{y})] = i\int\frac{d^3k}{(2\pi)^3}\,\frac{1}{2}\left(e^{-i\vec{k}\cdot(\vec{x}-\vec{y})} + e^{+i\vec{k}\cdot(\vec{x}-\vec{y})}\right) = i\,\delta^3(\vec{x}-\vec{y})\,. \tag{1.3.59}$$

Having $\mathcal{L}$ and the canonical momentum we can also construct the Hamiltonian density. Recall from your courses in classical mechanics that for a set of coordinates q^I and momentum p^I, the Hamiltonian H is

$$H = \left(\sum_I p^I\,\dot{q}^I\right) - L(p,q)\,, \tag{1.3.60}$$

where $L(p, q)$ is the Lagrangian as a function of the q^I and p^I. The equations of motion follow from the Poisson brackets with the Hamiltonian,

$$\dot{q}^I = \{q^I, H\} \qquad \dot{p}^I = \{p^I, H\}\,, \tag{1.3.61}$$

where the Poisson brackets are defined by

$$\{A, B\} = \sum_I \frac{\partial A}{\partial q^I}\frac{\partial B}{\partial p^I} - \frac{\partial A}{\partial p^I}\frac{\partial B}{\partial q^I}\,. \tag{1.3.62}$$

In the case of the scalar field we have an infinite number of "coordinates", namely $\phi(\vec{x})$ for each value of $\vec{x}$, and an infinite number or momenta $\Pi(\vec{x})$. The Hamiltonian is then

$$H = \int d^3x\, \Pi(\vec{x})\dot{\phi}(\vec{x}) - L = \int d^3x \left(\Pi(\vec{x})\dot{\phi}(\vec{x}) - \mathcal{L}\right) = \int d^3x\, \mathcal{H}(\vec{x})\,, \tag{1.3.63}$$

where $\mathcal{H}(\vec{x})$ is the Hamiltonian density. The equations of motion again follow from the Poisson brackets,

$$\dot{\phi}(\vec{x}) = \{\phi(\vec{x}), H\}\,, \qquad \dot{\Pi}(\vec{x}) = \{\Pi(\vec{x}), H\}\,, \tag{1.3.64}$$

where the Poisson brackets are now defined by

$$\{A, B\} = \int d^3x \left(\frac{\delta A}{\delta\phi(\vec{x})}\frac{\delta B}{\delta\Pi(\vec{x})} - \frac{\delta A}{\delta\Pi(\vec{x})}\frac{\delta B}{\delta\phi(\vec{x})}\right)\,, \tag{1.3.65}$$

and the functional derivatives satisfy

$$\frac{\delta\phi(\vec{y})}{\delta\phi(\vec{x})} = \delta^3(\vec{x} - \vec{y})\,, \qquad \frac{\delta\Pi(\vec{y})}{\delta\Pi(\vec{x})} = \delta^3(\vec{x} - \vec{y})\,. \tag{1.3.66}$$

The derivation of the Klein-Gordon equation from (1.3.64), (1.3.56) and (1.3.57) is left as an exercise

1.3.4 2-point correlators for the scalar field

Consider the 2-point correlator

$$\langle\phi(x)\phi(y)\rangle \equiv \langle 0|\phi(x)\phi(y)|0\rangle\,, \tag{1.3.67}$$

where no assumptions are made about x^μ and y^μ. Using the commuation relations in (1.3.40) it is straightforward to show that

$$\langle\phi(x)\phi(y)\rangle = \int \frac{d^3k}{(2\pi)^3 2\,\omega(\vec{k})}\, e^{-ik\cdot(x-y)}\,. \tag{1.3.68}$$

This should be a Lorentz invariant, since $\phi(x)$ and the vacuum are Lorentz invariant. In fact, one can show that the measure factor $\int \frac{d^3k}{(2\pi)^3\, 2\omega(\vec{k})}$ is Lorentz invariant since

$$\int \frac{d^3k}{(2\pi)^3 2\omega(\vec{k})}(2\pi)^3 2\omega(\vec{k})\delta^3(\vec{k} - \vec{q}) = 1\,. \tag{1.3.69}$$

The right hand side is obviously Lorentz invariant and the integrand is also Lorentz invariant as shown in (1.3.48), hence the measure factor is Lorentz invariant.

The time ordered correlator is

$$G_F(x-y) = \langle T\big[\phi(x)\phi(y)\big]\rangle = \int \frac{d^3k}{(2\pi)^3 2\,\omega(\vec{k})}\,\left(\theta(x^0-y^0)e^{-ik\cdot(x-y)} + \theta(y^0-x^0)e^{+ik\cdot(x-y)}\right)\,. \tag{1.3.70}$$

If we act on $G_F(x-y)$ with the Klein-Gordon operator ∂^2+m^2 we find

$$\begin{aligned}(\partial^2+m^2)G_F(x-y) &= -i\int \frac{d^3k}{(2\pi)^3}\,\frac{1}{2}\,\left(\partial_{x^0}\theta(x^0-y^0)e^{-ik\cdot(x-y)} - \partial_{x^0}\theta(y^0-x^0)e^{+ik\cdot(x-y)}\right)\\ &= -i\,\delta(x^0-y^0)\int\frac{d^3k}{(2\pi)^3}\,e^{+i\vec{k}\cdot(\vec{x}-\vec{y})} = -i\,\delta^4(x-y)\,. \end{aligned}\tag{1.3.71}$$

Hence $G_F(x-y)$ is a Green's function for the Klein-Gordon operator. In the presence of a source $J(x)$, the Klein-Gordon equation becomes

$$\partial^2\phi(x) + m^2\phi^2(x) = J(x) \tag{1.3.72}$$

where we get this equation by adding the source term

$$\mathcal{L}_{\mathrm{S}} = J(x)\phi(x) \tag{1.3.73}$$

to the Lagrangian. A general solution in the presence of the source is

$$\phi(x) = \phi_{\mathrm{hom}}(x) + \int d^4x'\,G_F(x-x')\,J(x')\,, \tag{1.3.74}$$

where ϕ_{hom} is the homogeneous solution in (1.3.2).

Fourier transforming $\phi(x)$ and $G_F(x-y)$ we have

$$\widetilde{\phi}(k) = \int d^4x e^{ik\cdot x}\phi(x)\,, \tag{1.3.75}$$

and

$$\begin{aligned}\int d^4x d^4y e^{ik\cdot x+iq\cdot y}\,G_F(x-y) &= (2\pi)^4\delta^4(k+q)\int d^4z\,e^{ik\cdot y}\,G_F(z)\\ &= (2\pi)^4\delta^4(k+q)\widetilde{G}_F(k)\,,\end{aligned}\tag{1.3.76}$$

where here the δ-function ensures conservation of 4-momentum. After integrating the spatial part we find

$$\begin{aligned}\widetilde{G}_F(k) &= \int dz^0\,e^{ik^0z^0}\,\frac{1}{2\,\omega(\vec{k})}\left(\theta(z_0)e^{-i\omega(\vec{k})z^0} + \theta(-z_0)e^{i\omega(-\vec{k})z^0}\right)\\ &= \frac{i}{(k^0)^2-\omega(\vec{k})^2+i\epsilon} = \frac{i}{k^2-m^2+i\epsilon}\,.\end{aligned}\tag{1.3.77}$$

Similar to the case of a single oscillator, (1.3.77) has poles at $k^0 = \pm(\omega(\vec{k})-i\epsilon)$

1.3.5 The physical interpretation of $G_F(x-y)$ and Causality

Suppose we have a single free particle at space-time position y^μ. The corresponding quantum state we write as $|\vec{y}, y^0\rangle$ where we have made the spatial and time components explicit. The probability amplitude to find the particle at a position $\vec{x}$ at time $x^0 > y^0$ is $\langle\vec{x}, x^0|\vec{y}, y^0\rangle$. This amplitude should be a Lorentz invariant. Now we have that

$$|\vec{y}, y^0\rangle = e^{iHy^0}|\vec{y}\rangle = \int \frac{d^3k}{(2\pi)^3(2\omega(\vec{k}))}\,|\vec{k}\rangle\langle\vec{k}|e^{iHy^0}|\vec{y}\rangle\,, \tag{1.3.78}$$

where $|\vec{k}\rangle$ is the single particle state defined in (1.3.42). Hence, we get

$$|\vec{y}, y^0\rangle = \int \frac{d^3k}{(2\pi)^3(2\omega(\vec{k}))}\,\sqrt{2\omega(\vec{k})}\,e^{i\omega(\vec{k})y^0 - i\vec{k}\cdot\vec{y}}\,a^\dagger_{\vec{k}}|0\rangle = \phi(y)|0\rangle\,, \tag{1.3.79}$$

where we used that $|\vec{y}\rangle$ has a normalization to cancel off the normalization factors in (1.3.42). Therefore, $\langle\vec{x}, x^0|\vec{y}, y^0\rangle = \theta(x_0 - y_0)\langle\phi(x)\phi(y)\rangle$ for $x^0 > y^0$. In other words the 2-point correlator is the amplitude for a free particle at y^μ to propagate to x^μ. For this reason the 2-point correlator is also called the propagator. $G_F(x-y)$ is usually called the Feynman propagator, where we allow for a particle to propagate from y^μ to x^μ if $x^0 > y^0$ and from x^μ to y^μ if $y^0 > x^0$.

The poles in $\widetilde{G}_F(k)$ tell us the physical mass of the particle. The presence of the poles at $k^2 = m^2 - i\epsilon$ is responsible for the $e^{-i\omega(\vec{k})x^0 + i\vec{k}\cdot\vec{x}}$ behavior in $G_F(x)$, which we expect for the amplitude of a single particle of mass m to go from y to x. Once we consider interactions the amplitude will be modified, but any pole the correlator has will still give us this same sort of behavior for the amplitude (there could be other contributions as well) letting us know that there is a physical particle with a mass at the pole.

The expression in (1.3.70) and the definition of a propagator relies on us being able to time order the fields. However if $x-y$ is spacelike, *i.e.*, $(x-y)^2 < 0$, then the time ordering is a relative concept. In other words there exists inertial frames where $x^0 - y^0 > 0$ and other inertial frames where $x^0 - y^0 < 0$. In order to avoid a contradiction, we must have that $\langle[\phi(x), \phi(y)]\rangle = [\phi(x), \phi(y)] = 0$ if $(x-y)^2 < 0$. To show that this is true, we first observe that $\langle[\phi(x), \phi(y)]\rangle$ is a Lorentz invariant. Hence, if it is zero it will be zero in all frames. So one strategy is to choose a particular frame where it is relatively easy to see if this commutator is zero. In particular, there exists a frame where the time coordinates are simultaneous, $x^0 - y^0 = 0$. It then follows from (1.3.68) that in this frame

$$\langle[\phi(x), \phi(y)]\rangle = \int \frac{d^3k}{(2\pi)^3 2\,\omega(\vec{k})}\,\left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})} - e^{-i\vec{k}\cdot(\vec{x}-\vec{y})}\right) = 0 \tag{1.3.80}$$

because of the symmetry of the measure under $\vec{k} \to -\vec{k}$.

If $\phi(x)$ did not commute with $\phi(y)$ when $(x-y)^2 < 0$ then causality would be violated. To see why, note that $\phi(x)$ is an Hermitian operator acting on a Hilbert space and corresponds to a physical quantity that is in principle measurable. Let us suppose that measurements are made of $\phi(x)$ and $\phi(y)$ where $(x-y)^2 < 0$. The space-time

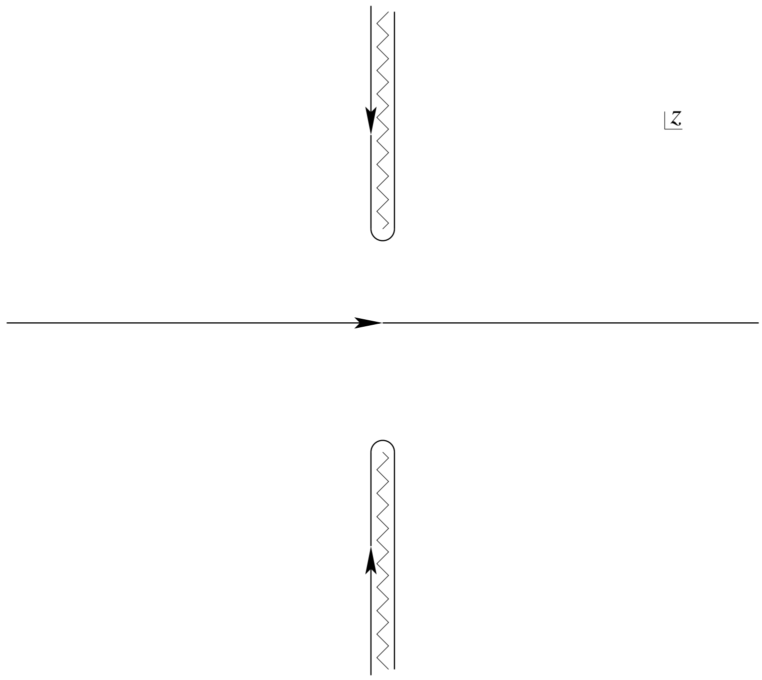


Figure 1.3: Integration contour with cuts in the z plane. The branch points are at $z = \pm i\, m|x-y|$.

positions x and y are out of causal contact, meaning that no physical signal can travel from one space-time point to the other. Hence the measurement of $\phi(x)$ cannot effect the measurement of $\phi(y)$ if causality is to hold. We know that in ordinary quantum mechanics two measurements do not affect each other only if the corresponding operators commute.

This is not to say that $\langle\phi(x)\phi(y)\rangle$ is zero if $(x-y)^2 < 0$. To see how the correlator behaves when $x^\mu - y^\mu$ is space-like, let us choose a frame where $x^0 - y^0 = 0$. In this case the propagator is

$$\begin{aligned}
\langle\phi(x)\phi(y)\rangle &= \int \frac{d^3k}{(2\pi)^3 2\,\omega(\vec{k})} e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \\
&= \frac{1}{2(2\pi)^3}\int_0^\infty k^2 dk \int_{-1}^{1} d(\cos\theta) \int_0^{2\pi} d\phi\, \frac{1}{\sqrt{k^2+m^2}}\, e^{ik\cos\theta|\vec{x}-\vec{y}|} \\
&= \frac{1}{4\pi^2}\int_0^\infty \frac{k\,dk}{\sqrt{k^2+m^2}}\, \frac{\sin(k|\vec{x}-\vec{y}|)}{|\vec{x}-\vec{y}|} = \frac{1}{8\pi^2|\vec{x}-\vec{y}|^2}\int_{-\infty}^{\infty} \frac{z\,dz}{\sqrt{z^2+m^2|\vec{x}-\vec{y}|^2}} \sin\, z\,.
\end{aligned} \tag{1.3.81}$$

We can now restore the correlator to the Lorentz invariant form by replacing $|\vec{x}-\vec{y}|^2$ with $-(x-y)^2 \equiv |x-y|^2$.

To do the last integral in (1.3.81), one observes that there are branch cuts running from $z = \pm i\, m|x-y|$ to $\pm i\,\infty$ (see figure 3). One can then write $\sin z = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right)$ and deform the contour around the upper branch cut for the e^{ikz} term and around the lower branch cut for the e^{-ikz} term. Both parts give the same contribution, so the correlator becomes

$$\langle\phi(x)\phi(y)\rangle = \frac{1}{4\pi^2|x-y|^2}\int_{m|x-y|}^{\infty}\frac{z\,dz}{\sqrt{z^2-m^2|x-y|^2}}\,e^{-z}\,. \qquad (1.3.82)$$

The integral can be done and gives a result involving special functions (a modified Bessel function of the second kind), but we are mainly interested in the general behavior of the correlator. In the limit where $m|x-y|\to 0$, the integral clearly approaches 1. If $m|x-y| >> 1$ then

$$\begin{aligned}\frac{1}{4\pi^2|x-y|^2}\int_{m|x-y|}^{\infty}\frac{z\,dz}{\sqrt{z^2-m^2|x-y|^2}}\,e^{-z} &\approx \frac{1}{4\pi^2|x-y|^2}\frac{(m|x-y|)^{1/2}}{\sqrt{2}}e^{-m|x-y|}\int_0^{\infty}z^{-1/2}e^{-z}\\ &= \frac{1}{4\pi^2}\left(\frac{m\pi}{2|x-y|^3}\right)^{1/2}e^{-m|x-y|}\,. \qquad (1.3.83)\end{aligned}$$

Hence, we find an exponential falloff if x^μ and y^μ are space-like separated. This should be expected: if the points are so separated, then a classical particle cannot propagate from one space-time point to the other. Quantum mechanically, there can be tunneling where we expect an exponentially small probability for the propagation to occur. Notice further that Compton wave-length of the particle is m^{-1}, which determines the falloff rate. To understand why we have this falloff, recall that in nonrelativistic quantum mechanics the exponential fall-off for tunneling between $\vec{y}$ and $\vec{x}$ can be computed using the WKB method,

$$\psi(x)\sim\exp\left(i\int_{\vec{y}}^{\vec{x}}\vec{p}\cdot d\vec{x}\right)\,, \qquad (1.3.84)$$

where $\vec{p}\cdot\vec{p}=2m(E-V(\vec{x}))$ is the classical trajectory for $\vec{p}$. In the classically forbidden region $\vec{p}\cdot\vec{p}$ is negative and so $\vec{p}\cdot d\vec{x}$ is imaginary, leading to the exponential suppression. In the case we are considering, let us choose a frame where $x^0-y^0=0$. The classical trajectory is determined by $p\cdot p=m^2$ and $m\,\dot{x}^\mu=p^\mu$, where "$\cdot$" refers to the derivative with respect to the proper-time. From this we see that $p^0=0$ and so $\vec{p}\cdot\vec{p}=-m^2$. Hence $\vec{p}$ is imaginary and so is the proper time. $\vec{p}$ is directed along $\vec{x}-\vec{y}$, from which it follows that $e^{-ip\cdot(x-y)}=e^{-m|\vec{x}-\vec{y}|}$.

If $x^\mu-y^\mu$ is time-like, then to evaluate the correlator we choose a frame where $\vec{x}-\vec{y}=0$. In this case, assuming $x^0-y^0>0$, we have

$$\langle\phi(x)\phi(y)\rangle = \frac{1}{4\pi^2}\int_0^{\infty}\frac{k^2dk}{\omega(k)}e^{-i\omega(k)(x^0-y^0)} = \frac{1}{4\pi^2}\int_m^{\infty}d\omega\,(\omega^2-m^2)^{1/2}\,e^{-i\omega(x^0-y^0)}\,. \qquad (1.3.85)$$

To make the this expression Lorentz invariant we replace x^0-y^0 with $\sqrt{(x-y)^2}\equiv|x-y|$. If $m|x-y| >> 1$ then we can approximate (1.3.85) to

$$\langle\phi(x)\phi(y)\rangle \approx \frac{1}{4\pi^2}\left(\frac{m\pi i}{2|x-y|^3}\right)^{1/2}e^{-im|x-y|}\,, \qquad (1.3.86)$$

thus here we find oscillatory behavior. Notice that (1.3.86) can also be obtained from (1.3.83) by replacing $|x-y|$ with $i|x-y|$, which is what one does to go from a space-like to a time-like separation.

As a final thought, note that correlator is is singular if $x^\mu - y^\mu$ is light-like, *i.e.* $(x-y)^2 = 0$, even if $x^\mu - y^\mu \neq 0$. This is basically a consequence of the Lorentz invariance and the singular behavior as $x^\mu - y^\mu \to 0$. If $(x-y)^2 = 0$ then we can keep boosting to a new frame, making $x^\mu - y^\mu$ arbitrarily small. Since the final result cannot depend on the frame, it must be that the correlator is singular in all frames.

1.4 Symmetries and Noether's theorem

Symmetries are an important tool in physics as they allow us to make practical statements about many physical quantities. For example, the mass of the proton is 938 MeV. The mass is almost entirely dependent on the physics of QCD, but actually computing this value is extremely difficult and requires a huge amount of computing power. On the other hand, because of a symmetry called isospin, one can definitively say that the neutron should have have a mass that is very close to the proton mass, which it is, 939 MeV. Isospin symmetry is not exact, that is why the masses are slightly different, but it is close to an exact symmetry. In this section we will explore some aspects of symmetries in scalar field theories.

If a *continuous* symmetry exists, then under the shift $\phi(x) + \epsilon f(\phi(x))$, where ϵ is an infinitesimal parameter and $f(\phi(x))$ is some function of the fields, the Lagrangian $\mathcal{L}$ is invariant up to order ϵ^2. The symmetries we consider here are global symmetries, meaning that ϵ has no space-time dependence. Later in the course we will see a different type of symmetry called a gauge symmetry, where the transformations are local, meaning that they can depend on the coordinate x^μ.

As an example, consider the massless Lagrangian $\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi$. This is invariant under $\phi(x) \to \phi(x) + \epsilon$. If there were a mass term then this would no longer be the case. As a second example, consider the Lagrangian for a complex scalar field,

$$\mathcal{L} = \partial_\mu\phi^*(x)\partial^\mu\phi - m^2\phi^*(x)\phi(x)\,. \tag{1.4.1}$$

Under the infinitesimal shift, $\phi(x) \to \phi(x) + i\epsilon\,\phi(x)$ the Lagrangian shifts to $\mathcal{L} \to \mathcal{L} + O(\epsilon^2)$ and so is invariant to leading approximation. This last transformation is the infinitesimal form of the transformation of $\phi(x) \to e^{i\theta}\,\phi(x)$ which clearly leaves $\mathcal{L}$ invariant.

Noether's theorem states that for every *continuous symmetry* there is a conserved current. Actually, we can relax our symmetry requirements a bit by allowing $\mathcal{L}$ to be invariant up to a total derivative term

$$\mathcal{L} \to \mathcal{L} + \epsilon\,\partial_\mu\mathcal{J}^\mu \tag{1.4.2}$$

since such a term will not affect the equations of motion. Since the symmetry transformation is infinitesimal, we can identify $\epsilon\, f(\phi(x))$ with $\delta\phi(x)$. If we assume that $\phi(x)$

satisfies the equations of motion then

$$\begin{aligned}
\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi}\partial_\mu\delta\phi + \frac{\partial\mathcal{L}}{\partial\,\phi}\delta\phi \\
&= \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi}\delta\phi\right) - \left(\partial_\mu\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi} - \frac{\partial\mathcal{L}}{\partial\,\phi}\right)\delta\phi \\
&= \epsilon\,\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi}f(\phi)\right) = \epsilon\,\partial_\mu\mathcal{J}^\mu\,. \qquad (1.4.3)
\end{aligned}$$

If we define the current j^μ

$$j^\mu = \frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi}f(\phi) - \mathcal{J}^\mu\,, \qquad (1.4.4)$$

it follows from (1.4.3) that $\partial_\mu j^\mu = 0$. Hence, j^μ is conserved.

In the case of the constant shift, $\phi(x) \to \phi(x) + \epsilon$, the current is $j^\mu = \partial^\mu\phi(x)$. In the case of the complex field the current is

$$j^\mu = i(\partial^\mu\phi^*)\phi - i\phi^*\partial^\mu\phi\,. \qquad (1.4.5)$$

If we write $\phi(x) = \rho(x)\,e^{i\theta(x)}$, then we have that $j^\mu(x) = 2\rho^2(x)\,\partial^\mu\theta(x)$.

From the time component of the current we can build the charge

$$Q = \int d^3x\,j^0(\vec{x}, x^0) \qquad (1.4.6)$$

which is constant in time assuming that $j^0(x)$ falls off sufficiently fast as $x \to \infty$:

$$\left(\frac{\partial}{\partial t}Q = \int d^3x\,\frac{\partial}{\partial t}j^0(\vec{x}, x^0) = -\int d^3x\,\vec{\nabla}\cdot\vec{j}(\vec{x}, x^0) = -\int_\infty d\vec{S}\cdot\vec{j} = 0\,.\right) \qquad (1.4.7)$$

After quantizing, Q becomes an operator that commutes with the Hamiltonian. Therefore, the eigenstates in the quantum field theory can also be eigenstates of Q and Q becomes a well-defined quantum number. Of course, if there several independent charges, then the individual charges don't necessarily commute with each other.

As an example, let us return to the complex scalar. The canonical momentum for $\phi(x)$ is $\Pi(x) = \partial^0\phi^*(x)$, while that for $\phi^*(x)$ is $\Pi^*(x) = \partial^0\phi(x)$. Hence, after quantization the equal time commutation relations are

$$\begin{aligned}
[\phi(\vec{x}, x^0), \partial^0\phi^*(\vec{y}, x^0)] &= i\,\delta^3(\vec{x}-\vec{y})\,, \qquad [\phi^*(\vec{x}, x^0), \partial^0\phi(\vec{y}, x^0)] = i\,\delta^3(\vec{x}-\vec{y})\,, \\
[\phi(\vec{x}, x^0), \partial^0\phi(\vec{y}, x^0)] &= [\phi^*(\vec{x}, x^0), \partial^0\phi^*(\vec{y}, x^0)] = 0\,. \qquad (1.4.8)
\end{aligned}$$

If we now commute Q with $\phi(\vec{x}, x^0)$ and $\phi^*(\vec{x}, x^0)$, using (1.4.5) and (1.4.8) as well as the fact that Q is constant in time (and so we can substitute $Q = Q(x^0)$), we find

$$[Q, \phi(\vec{x}, x^0)] = +\,\phi(\vec{x}, x^0)]\,, \qquad [Q, \phi^*(\vec{x}, x^0)] = -\,\phi^*(\vec{x}, x^0)]\,. \qquad (1.4.9)$$

When quantizing this Q there is an ambiguity about operator ordering, since $\phi(x)$ does not commute with $\partial^0\phi^*(x)$. A standard choice is to assume that the fields in Q are normal ordered. This means that all creation operators are to the left of annihilation operators. With this condition we have that $Q|0\rangle = 0$, that is, the vacuum has zero charge. If we create the single particle state out of the vacuum, $\phi(x)|0\rangle$, then it follows from the commutation relations (1.4.9) that this state has charge $+1$. On the other hand, we can create a different particle state, $\phi^*(x)|0\rangle$ which has charge -1. Once we include interactions the particle numbers can change. But if the transformation $\phi(x) \to e^{i\theta}\phi(x)$ continues to be a symmetry, then the charge Q cannot change in the interactions.

Noether's theorem can also be used for symmetries that do not directly change the fields, but instead are symmetries of the coordinates. For example, space-time translations are implemented by $x^\mu \to x^{\mu'} = x^\mu - \epsilon^\mu$. In a translationally invariant theory, any scalar is invariant under this transformation, in the sense that $\phi'(x') = \phi(x)$. To leading order in ϵ^μ this means that the transformation for $\phi(x)$ is

$$\phi(x) \to \phi'(x) = \phi(x+\epsilon) = \phi(x) + \epsilon^\mu \partial_\mu \phi(x)\,. \tag{1.4.10}$$

Since (1.4.10) applies for any scalar, the transformation for $\mathcal{L}(x)$ is

$$\mathcal{L}(x) \to \mathcal{L}(x) + \epsilon^\mu \partial_\mu \mathcal{L}(x)\,, \tag{1.4.11}$$

since it too is a scalar. If we let

$$T^{\mu\nu} = \frac{\mathcal{L}}{\partial\partial_\mu\phi}\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}\,, \tag{1.4.12}$$

then following the same argument as in (1.4.3), we find that $\partial_\mu T^{\mu\nu} = 0$. $T^{\mu\nu} = T^{\nu\mu}$ is the energy momentum tensor and its conservation is a statement about the local conservation of energy and momentum. T^{00} is the energy density and the "charge" is the total energy,

$$E = \int d^3x\, T^{00}. \tag{1.4.13}$$

The other charges are the components of the total momentum

$$P^i = \int d^3x\, T^{0i}\,. \tag{1.4.14}$$

We previously argued that the translational invariance of $G_F(x-y)$ led to a conservation of 4-momentum for the two-point function. Here we see that this is a consequence of Noether's theorem.

Chapter 2

Path Integrals

In this chapter of the notes we introduce the concept of a path integral. We first consider the path integral for a single nonrelativistic particle. We then specialize this to the harmonic oscillator, which is a noninteracting $0+1$ dimensional field theory. We then generalize to the case of path integrals for scalar fields in $3+1$ space-time dimensions.

2.1 Path integrals for nonrelativistic particles

2.1.1 Generalities

Suppose we pose the following problem in nonrelativistic quantum mechanics: given that a particle is at position $\vec{x}_1$ at time t_1, what is the probability that the particle is at position $\vec{x}_2$ at time t_2?

The wave-function evolves as[1]

$$\Psi(\vec{x},t) = e^{-\frac{i}{\hbar}H(t-t_0)}\Psi(\vec{x},t_0)\,, \tag{2.1.1}$$

where the Hamiltonian H is assumed to have the form

$$H = \frac{\vec{p}^2}{2m} + V(\vec{x}). \tag{2.1.2}$$

Writing $\Psi(\vec{x},t)$ as an inner product with position states $|\vec{x}\rangle$, we then have

$$\Psi(\vec{x},t) = \langle\vec{x}|\Psi,t\rangle = \langle\vec{x}|e^{-iHt/\hbar}|\Psi\rangle \equiv \langle\vec{x},t|\Psi\rangle\,. \tag{2.1.3}$$

(We have inserted an explicit factor of $\hbar$ for later pedagogical convenience.) $|\vec{x},t\rangle$ is the state where the particle is at $\vec{x}$ at time t . Hence, the amplitude that the particle starts at $\vec{x}_1$ at t_1 and ends up at $\vec{x}_2$ at t_2 is

$$\langle\vec{x}_2,t_2|\vec{x}_1,t_1\rangle = \langle\vec{x}_2|e^{-\frac{i}{\hbar}H(t_2-t_1)}|\vec{x}_1\rangle\,. \tag{2.1.4}$$

[1] For this section we will keep the factors of $\hbar$.

To evaluate the expression in (2.1.4), we split up the time interval between t_1 and t_2 into a large number of infinitesimally small time intervals Δt , so that

$$e^{-\frac{i}{\hbar}H(t_2-t_1)} = \prod_{\Delta t} e^{-\frac{i}{\hbar}H\Delta t}, \tag{2.1.5}$$

where the product is over all time intervals between t_1 and t_2. Since, $\vec{p}$ does not commute with $\vec{x}$, we see that

$$e^{-\frac{i}{\hbar}H\Delta t} \neq e^{-\frac{i}{2m\hbar}\vec{p}^2\Delta t}e^{-\frac{i}{\hbar}V(\vec{x})\Delta t}, \tag{2.1.6}$$

however, if Δt is very small, then it is approximately true, in that

$$e^{-\frac{i}{\hbar}H\Delta t} = e^{-\frac{i}{2m\hbar}\vec{p}^2\Delta t}e^{-\frac{i}{\hbar}V(\vec{x})\Delta t} + \mathrm{O}((\Delta \mathrm{t})^2). \tag{2.1.7}$$

Hence, in the limit that $\Delta t \to 0$, we have that

$$\langle\vec{x}_2, t_2|\vec{x}_1, t_1\rangle = \lim_{\Delta t\to 0}\langle\vec{x}_2|\prod_{\Delta t} e^{-\frac{i}{2m\hbar}\vec{p}^2\Delta t}e^{-\frac{i}{\hbar}V(\vec{x})\Delta t}|\vec{x}_1\rangle. \tag{2.1.8}$$

In order to evaluate the expression in (2.1.8), we need to insert a complete set of states between each term in the product, the states being either position or momentum eigenstates and normalized so that

$$\begin{aligned} &\int d^3x\, |\vec{x}\rangle\langle\vec{x}| = 1 && \langle\vec{x}_2|\vec{x}_1\rangle = \delta^3(\vec{x}_2-\vec{x}_1) \\ &\int \frac{d^3p}{(2\pi\hbar)^3}\, |\vec{p}\rangle\langle\vec{p}| = 1 && \langle\vec{p}_2|\vec{p}_1\rangle = (2\pi\hbar)^3\delta^3(\vec{p}_2-\vec{p}_1) \\ &\langle\vec{x}|\vec{p}\rangle = e^{i\vec{x}\cdot\vec{p}/\hbar} \end{aligned} \tag{2.1.9}$$

Inserting the position states first, we have that (2.1.8) is

$$\langle\vec{x}_2, t_2|\vec{x}_1, t_1\rangle = \int \prod_{t_1<t<t_2} d^3x(t)\left[\prod_{t_1\le t<t_2}\langle\vec{x}(t+\Delta t)|e^{-\frac{i}{2m\hbar}|\vec{p}|^2\Delta t}e^{-\frac{i}{\hbar}V(\vec{x})\Delta t}|\vec{x}(t)\rangle\right], \tag{2.1.10}$$

where the first product is over each time between $t_1+\Delta t$ and $t_2-\Delta t$ and the second product is over each time between t_1 and $t_2-\Delta t$. We also have that $\vec{x}(t_1) = \vec{x}_0$ and $\vec{x}(t_2) = \vec{x}_2$. Note that for each time interval, we have a position variable that we integrate over. You should think of the time variable in these integrals as a label for the different $\vec{x}$ variables.

We now need to insert a complete set of momentum states at each time t. In particular, we have that

$$\begin{aligned} &\langle\vec{x}(t+\Delta t)|e^{-\frac{i}{2m\hbar}|\vec{p}|^2\Delta t}e^{-\frac{i}{\hbar}V(\vec{x})\Delta t}|\vec{x}(t)\rangle \\ &= \int \frac{d^3p}{(2\pi\hbar)^3}\langle\vec{x}(t+\Delta t)|\vec{p}\rangle e^{-\frac{i}{2m\hbar}|\vec{p}|^2\Delta t}\langle\vec{p}|\vec{x}(t)\rangle e^{-\frac{i}{\hbar}V(\vec{x}(t))\Delta t} \\ &= \int \frac{d^3p}{(2\pi\hbar)^3} e^{i\Delta\vec{x}(t)\cdot\vec{p}/\hbar}e^{-\frac{i}{2m\hbar}|\vec{p}|^2\Delta t}e^{-\frac{i}{\hbar}V(\vec{x}(t))\Delta t} \end{aligned} \tag{2.1.11}$$

where $\Delta x(t) = x(t+\Delta t) - x(t)$. We can now do the gaussian integral in the last line in (2.1.11), which after completing the square gives

$$\langle \vec{x}(t+\Delta t)| e^{-\frac{i}{2m\hbar}|\vec{p}|^2\Delta t} e^{-\frac{i}{\hbar}V(\vec{x})\Delta t} |\vec{x}(t)\rangle = \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{3/2} \exp\left(\frac{i}{\hbar}\left[\frac{m}{2}\left(\frac{\Delta\vec{x}}{\Delta t}\right)^2 - V(\vec{x}(t))\right]\Delta t\right). \tag{2.1.12}$$

Strictly speaking, the integral in the last line of (2.1.11) is not a Gaussian, since the coefficient in front of $\vec{p}^2$ is imaginary. However, we can regulate this by assuming that the coefficient has a small negative real part and then let this part go to zero after doing the integral. Shortly we will give a more physical reason for this regularization. The last term in (2.1.12) can be written as

$$= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{3/2} \exp\left(\frac{i}{\hbar}\left[\frac{m}{2}\dot{\vec{x}}^2 - V(\vec{x}(t))\right]\Delta t\right) = \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{3/2} \exp\left(\frac{i}{\hbar}L(t)\Delta t\right), \tag{2.1.13}$$

where $L(t)$ is the lagrangian of the particle evaluated at time t. Hence the complete expression can be written as

$$\begin{aligned} \langle \vec{x}_2, t_2 | \vec{x}_1, t_1\rangle &= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{3N/2} \int \prod_{t_1<t<t_2} d^3x(t) \exp\left(\frac{i}{\hbar}S\right) \\ &= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{3N/2} \int \mathcal{D}\vec{x}\, \exp\left(\frac{i}{\hbar}S\right), \end{aligned} \tag{2.1.14}$$

where S is the action

$$S = \int_{t_1}^{t_2} L(t)dt. \tag{2.1.15}$$

N counts the number of time intervals in the path and $\mathcal{D}\vec{x}$ tells us to integrate over all $\vec{x}$ for every point in time t. Taking the limit $N \to \infty$ leads to an infinite dimensional integral, known as a functional integral. In this limit we see that the constant in front of the expression diverges. It is standard practice to drop it since it is essentially a normalization constant and can always be restored later.

The expression in (2.1.14) was first derived by Feynman and it gives a very intuitive way of looking at quantum mechanics. What the expression is telling us is that to compute the probability amplitude, we need to sum over all possible paths that the particle can take in getting from $\vec{x}_1$ to $\vec{x}_2$, weighted by $e^{\frac{i}{\hbar}S}$. For this reason the functional integral is also called a path integral.

It is natural to ask which path dominates the path integral. Since the argument of the exponential is purely imaginary, we see that the path integral is a sum over phases. In general, when integrating over the $\vec{x}(t)$, the phase varies and the phases coming from the different paths tend to cancel each other out. What is needed is a path where varying to a nearby path gives no phase change. Then the phases add constructively and we are left with a large contribution to the path integral from the path and its nearby neighbors.

Hence, we look for the path, given by a parameterization $\vec{x}(t)$, such that $\vec{x}(t_1) = \vec{x}_1$ and $\vec{x}(t_2) = \vec{x}_2$, and such that the nearby paths have the same phase, or close to the same phase. This means that if $\vec{x}(t)$ is shifted to $\vec{x}(t) + \delta\vec{x}(t)$, then the change to the action is very small. To find this path, note that under the shift, to lowest order in $\delta\vec{x}$, the action changes by

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial \dot{\vec{x}}}\delta\dot{\vec{x}} + \frac{\partial L}{\partial \vec{x}}\delta\vec{x}\right] = \int_{t_1}^{t_2} dt \left[-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{x}}}\right) + \frac{\partial L}{\partial \vec{x}}\right]\delta\vec{x}. \tag{2.1.16}$$

Hence there would be no phase change to lowest order in δx if the term inside the square brackets is zero. But of course, this is just the classical equation of motion. A generic path has a phase change of order $\delta\vec{x}$, but the classical path has a phase change of order $(\delta\vec{x})^2$.

Next consider what happens as $\hbar \to 0$. Then a small change in the action can lead to a big change in the phase. In fact, even a very small change in the action essentially wipes out any contribution to the path integral. In this case, the classical path is essentially the *only* contribution to the path integral. For nonzero $\hbar$, while the classical path is the dominant contributor to the path integral, the nonclassical paths also contribute, since the phase is finite.

The path integral need not be used only for a transition from one position state to another, we can also use it for other transition amplitudes. For example, the transition amplitude $\langle\psi_2, t_2|\psi_1, t_1\rangle$ can be converted to the form above by inserting the complete set of position states at t_1 and t_2, where afterward one finds

$$\langle\psi_2, t_2|\psi_1, t_1\rangle = \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{3N/2}\int\prod_{t_1\le t\le t_2} d^3x(t)\,\psi_2^*(x(t_2))\psi_1(x(t_1))\,\exp\left(\frac{i}{\hbar}S\right). \tag{2.1.17}$$

2.1.2 The simple harmonic oscillator again

Let us now return to the one dimensional harmonic oscillator. We again write the Lagrangian as

$$L = \tfrac{1}{2}\dot{\phi}^2 - \tfrac{1}{2}\omega_0^2\phi^2\,. \tag{2.1.18}$$

The path integral for the transition from the ground state at $t = -T$ to the ground state at $t = +T$ is then[2]

$$\begin{aligned}\mathcal{Z} &\equiv \langle 0, +T|0, -T\rangle \\ &= (2\pi i\,\Delta t)^{-N/2}\int\prod_{-T\le t\le +T} d\phi(t)\,\psi_0^*(\phi(T))\psi_0(\phi(-T))\,\exp\left(i\int dt\,\tfrac{1}{2}\dot{\phi}^2(t) - \tfrac{1}{2}\omega_0^2\phi^2(t)\right) \\ &= \mathcal{C}\int\prod_{-T\le t\le +T} d\phi(t)\,e^{-\frac{1}{2}\omega_0\phi^2(T)}\,e^{-\frac{1}{2}\omega_0\phi^2(-T)}\,\exp\left(i\int dt\,\tfrac{1}{2}\dot{\phi}^2(t) - \tfrac{1}{2}\omega_0^2\phi^2(t)\right)\,,\end{aligned} \tag{2.1.19}$$

[2] $\hbar$ is back to $\hbar = 1$.

where we have incorporated all constants into one constant $\mathcal{C}$, which will later drop out of our computations. The integral in (2.1.19) has the form of an infinite number of gaussians and hence is solvable. The gaussians are coupled because of the $\dot{\phi}^2$ term, so it is very helpful to choose new coordinates so that the argument of the exponent is diagonalized.

But before doing this, let us make a remark about the wave-function factor $e^{-\frac{1}{2}(\phi^2(T)+\phi^2(-T))}$. Let us suppose that we have not yet taken the limit $N\to\infty$ where N is the number of time intervals. In this case, the path integral has the form

$$\mathcal{Z} \;=\; \mathcal{C}\int\prod_{j=1}^{N} d\phi_j \exp\left(-\tfrac{1}{2}\phi_j M_{jk}\phi_k)\right)\,. \qquad (2.1.20)$$

Because all ϕ_j are indirectly coupled to $\phi(\pm T)$ through the $\dot{\phi}^2$ term, all eigenvalues of M_{jk} will have a positive real part. Hence, the integral is well defined and is given by

$$\mathcal{Z} \;=\; \mathcal{C}(2\pi)^{N/2}(\det M)^{-1/2}\,, \qquad (2.1.21)$$

where we have used that the determinant of a matrix is the product of its eigenvalues. In the limit that $N\to\infty$, the eigenvalues continue to have a positive real part which becomes vanishingly small, but continues to leave the integrals well defined. In the end we can ignore the effects of the ground-state wave-functions, except to include a small positive real part in the eigenvalues of M.

In the limit that $\Delta t\to 0$, M becomes the differential operator $i(\partial_t^2-\omega_0^2)$. The determinant of this might seem a little obscure, so let us go back to the path integral in (2.1.19) and Fourier transform to frequency space, where

$$\widetilde{\phi}(\omega)=\int dt\, e^{i\omega t}\,\phi(t)\,. \qquad (2.1.22)$$

Carrying out the Fourier transform (almost) diagonalizes M and the path integral becomes

$$\mathcal{Z} \;=\; \mathcal{C}\int\mathcal{D}\widetilde{\phi}\;\exp\left(i\int\frac{d\omega}{2\pi}\,\tfrac{1}{2}\,\widetilde{\phi}(-\omega)(\omega^2-\omega_0^2+i\,\epsilon)\widetilde{\phi}(\omega)\right)\,, \qquad (2.1.23)$$

where the $+i\,\epsilon$ term gives all eigenvalues a positive real part. In transforming the measure $\mathcal{D}\phi$ to $\mathcal{D}\widetilde{\phi}$ we used that $\mathcal{D}\widetilde{\phi}=\det(\Delta t\, e^{i\omega t})\mathcal{D}\phi$, where the determinant is for an infinite dimensional matrix (in the limit where $\Delta t\to 0$ and $T\to\infty$) where the columns are over the t space and the rows are over the ω space. The factor is independent of $\widetilde{\phi}(\omega)$ and so can be absorbed into $\mathcal{C}$. Note that $\widetilde{\phi}(-\omega)=\widetilde{\phi}^*(\omega)$ and that

$$\int d^2z\, e^{-b|z|^2}=\frac{2\pi}{b}\,,$$

if $\mathrm{Re}\, b>0$. Therefore, the path integral is given by the infinite product

$$\mathcal{Z} \;=\; \lim_{\Delta\omega\to 0}\mathcal{C}\prod_{\omega}\left(\frac{4\,i\,\pi^2/\Delta\omega}{\omega^2-\omega_0^2+i\epsilon}\right)^{1/2}\,. \qquad (2.1.24)$$

The path integral has a strong similarity to a partition function in statistical mechanics, so for this reason $\mathcal{Z}$ is often called a partition function.

2.2 n-point correlators and Wick's theorem

Our real interest is in the correlators, which will be of particular importance when we consider interactions. We can generate correlators for any number of fields using source terms.

Let us continue with our example of the simple harmonic oscillator. In this case the n-point time-ordered correlator is given by

$$\langle T[\phi(t_1)\phi(t_2)\dots\phi(t_n)]\rangle = \mathcal{Z}^{-1}\int\mathcal{D}\phi\,\phi(t_1)\phi(t_2)\dots\phi(t_n)\,\exp\left(i\int dt\,\tfrac{1}{2}\dot\phi^2(t)-\tfrac{1}{2}\omega_0^2\phi^2(t)\right)\,. \tag{2.2.1}$$

where $\mathcal{Z}^{-1}$ normalizes the expression. Notice that the path integral automatically time orders the fields, since the path integral itself is constructed by time ordering the vanishing small segments. Using that $\frac{\delta\phi(t)}{\delta\phi(t')}=\delta(t-t')$, we can then write (2.2.1) as

$$\langle T[\phi(t_1)\phi(t_2)\dots\phi(t_n)]\rangle = \mathcal{Z}^{-1}\prod_{j=1}^{n}-i\frac{\delta}{\delta J(t_j)}\,\mathcal{Z}(J)\Big|_{J=0}\,, \tag{2.2.2}$$

where $\mathcal{Z}(J)$ is given by

$$\mathcal{Z}(J)=\mathcal{C}\int\mathcal{D}\phi\,e^{iS(J)}=\mathcal{C}\int\mathcal{D}\phi\,\exp\left(i\int dt\,\tfrac{1}{2}\dot\phi^2(t)-\tfrac{1}{2}(\omega_0^2-i\,\epsilon)\phi^2(t)+J(t)\phi(t)\right)\,, \tag{2.2.3}$$

and where we have included the $i\,\epsilon$ term discussed in the previous section. Note that the $i\,\epsilon$ arises in the limit where $\Delta t\to 0$ but also $T\to\infty$. As long as ϵ is not exactly zero, then T is not exactly ∞. Hence the correlators will be zero if $t_j<-T$ or $t_j>+T$ and so they are strictly speaking turned off in the distant past or distant future.

$\mathcal{Z}(J)$ can be evaluated by completing the square. Doing an integration by parts we can rewrite $\mathcal{Z}(J)$ as

$$\begin{aligned}\mathcal{Z}(J) &= \mathcal{C}\int\mathcal{D}\phi\,\exp\left(-i\int dt\,\tfrac{1}{2}\big[\phi(t)-JM^{-1}(t)\big](\partial_t^2+\omega_0^2-i\,\epsilon)\big[\phi(t)-M^{-1}J(t)\big]\right)\\ &\qquad\times\exp\left(i\int dt\,\tfrac{1}{2}\,JM^{-1}(t)(\partial_t^2+\omega_0^2-i\,\epsilon)M^{-1}J(t)\right)\,,\end{aligned} \tag{2.2.4}$$

where

$$M^{-1}\phi(t)=\int dt' M^{-1}(t,t')\phi(t')\,, \tag{2.2.5}$$

and where

$$(\partial_t^2+\omega_0^2-i\epsilon)M^{-1}(t,t')=\delta(t-t')\,. \tag{2.2.6}$$

Hence, $M^{-1}(t,t') = i\,G_F(t-t')$ where $G_F(t-t')$ is the Feynman Green's function (it has the Feynman pole structure because of the $i\,\epsilon$ term). Shifting the integration variables, $\phi(t) \to \phi(t) + JM^{-1}(t)$, we get

$$\begin{aligned}
\mathcal{Z}(J) &= \mathcal{C}\int \mathcal{D}\phi\ \exp\left(i\int dt\,\tfrac{1}{2}\dot\phi^2(t) - \tfrac{1}{2}\omega_0^2\phi^2(t)\right)\exp\left(-\int dt\,dt'\,\tfrac{1}{2}\,J(t)G_F(t-t')J(t')\right)\right) \\
&= \mathcal{Z}(0)\ \exp\left(-\int dt\,dt'\,\tfrac{1}{2}\,J(t)G_F(t-t')J(t')\right)\,. \qquad (2.2.7)
\end{aligned}$$

It is now very simple to find the correlator. Since we will be setting $J(t)$ to zero after taking the derivatives, it is clear that we cannot leave any $J(t)$ terms outside of the exponential. Hence, we always have to do the functional derivatives in pairs. The n-point correlator will then be a sum over all inequivalent ways to make $n/2$ pairs. We thus find,

$$\begin{aligned}
\langle T[\phi(t_1)\phi(t_2)\ldots\phi(t_n)]\rangle &= \sum_{\text{pairs}}\prod_{k=1}^{n/2}\mathcal{Z}^{-1}(0)\frac{-i\,\delta}{\delta J(t_{I(k)})}\,\frac{-i\,\delta}{\delta J(t_{J(k)})}\,\mathcal{Z}(J)\Big|_{J(t)=0} \\
&= \sum_{\text{pairs}}\prod_{k=1}^{n/2}G_F(t_{I(k)}-t_{J(k)}) = \sum_{\text{pairs}}\prod_{k=1}^{n/2}\langle T[\phi(t_{I(k)})\phi(t_{J(k)})]\rangle\,.
\end{aligned} \qquad (2.2.8)$$

The $\{I(j), J(j)\}$ represent the indices in the j^{th} pair. Thus, the n-point correlator can be written as a sum over products of 2-point correlators. This reduction of correlators is known as Wick's theorem[3]. The two fields that appear in a 2-point correlator are said to be *Wick contracted*.

As an example, let us consider the 4-point correlator. Here there are three inequivalent ways to make pairs, $\{\{1,2\},\{3,4\}\}$, $\{\{1,3\},\{2,4\}\}$, and $\{\{1,4\},\{2,3\}\}$. Hence the correlator is

$$\begin{aligned}
\langle T[\phi(t_1)\phi(t_2)\phi(t_3)\phi(t_3)]\rangle &= G_F(t_1-t_2)G_F(t_3-t_4) + G_F(t_1-t_3)G_F(t_2-t_4) \\
&\quad + G_F(t_1-t_4)G_F(t_2-t_3)\,. \qquad (2.2.9)
\end{aligned}$$

For an n-point correlator the number of inequivalent pairs is

$$\prod_{j=1}^{n/2}(2j-1) = \frac{n!}{2^{n/2}(n/2)!}\,. \qquad (2.2.10)$$

To close this section, note that the correlator is zero if n is odd. This is obvious because there will be a left over $J(t)$ when taking functional derivatives. But this can also be understood by the symmetry in the theory. Note that $\mathcal{L}$ is invariant under the discrete transformation $\phi(t) \to -\phi(t)$. Since the path integral's functional integrand integrates over all $\phi(t)$, we have that $\mathcal{Z}(J) = \mathcal{Z}(-J)$. Hence, an odd number of functional derivatives on $\mathcal{Z}(J)$ at $J(t) = 0$ gives zero.

[3] Actually, Wick's theorem is a little stronger than what is stated here, but we don't need the stronger version. You can read more about this in Peskin.

2.3 Path integrals for the scalar field

It is relatively straightforward to generalize the path integral for the simple harmonic oscillator to the case of a free scalar field. If we Fourier transform the spatial part, so that

$$\widetilde{\phi}(\vec{k},t)=\int d^3x\, e^{-i\vec{k}\cdot\vec{x}}\,\phi(x)\,, \tag{2.3.1}$$

then the path integral for the vacuum to vacuum amplitude is that for an infinite number of independent oscillators. Therefore, we define $\mathcal{Z}$ to be

$$\mathcal{Z}\equiv\lim_{T\to\infty}\langle 0,+T|0,-T\rangle=\lim_{T\to\infty}\langle 0|e^{-2iTH}|0\rangle\,, \tag{2.3.2}$$

where

$$H=\int\frac{d^3k}{(2\pi)^3}\,H_{\vec{k}}=\int\frac{d^3k}{(2\pi)^3}\,\tfrac{1}{2}\omega(\vec{k})\left(a^\dagger_{\vec{k}}\,a_{\vec{k}}+a_{\vec{k}}\,a^\dagger_{\vec{k}}\right)\,. \tag{2.3.3}$$

If we think of all values of $\vec{k}$ as the limit of a discrete set, where each component of $\vec{k}$ is separated by Δk from its neighbor, then we can write $\mathcal{Z}$ as (dropping overall constants)

$$\mathcal{Z}=\lim_{\Delta k\to 0}\prod_{\vec{k}}\langle 0|\exp\left(-2iT\frac{(\Delta k)^3}{(2\pi)^3}H_{\vec{k}}\right)|0\rangle=\lim_{\Delta k\to 0}\prod_{\vec{k}}\mathcal{Z}_{\vec{k}}\,. \tag{2.3.4}$$

If we now take $\vec{k}\neq 0$ and consider the product $\mathcal{Z}_{\vec{k}}\mathcal{Z}_{-\vec{k}}$, we find using our results from a single harmonic oscillator (actually 2 harmonic oscillators)

$$\mathcal{Z}_{\vec{k}}\mathcal{Z}_{-\vec{k}}=\int\mathcal{D}\widetilde{\phi}_{\vec{k}}\mathcal{D}\widetilde{\phi}_{-\vec{k}},\exp\left(i\frac{(\Delta k)^3}{(2\pi)^3}\int dt\,\dot{\widetilde{\phi}}_{\vec{k}}(t)\dot{\widetilde{\phi}}_{-\vec{k}}(t)-\widetilde{\phi}_{\vec{k}}(t)(\vec{k}^{\,2}+m^2-i\,\epsilon)\widetilde{\phi}_{-\vec{k}}(t)\right)\,. \tag{2.3.5}$$

Hence, the full path integral is

$$\mathcal{Z}\;=\;\int\left(\prod_{\vec{k}}\mathcal{D}\widetilde{\phi}_{\vec{k}}\right)\,\exp\left(i\int\frac{d^3k}{(2\pi)^3}\int dt\,\tfrac{1}{2}\,\dot{\widetilde{\phi}}_{\vec{k}}(t)\dot{\widetilde{\phi}}_{-\vec{k}}(t)-\tfrac{1}{2}\,\widetilde{\phi}_{\vec{k}}(t)(\vec{k}^{\,2}+m^2-i\,\epsilon)\widetilde{\phi}_{-\vec{k}}(t)\right)\,. \tag{2.3.6}$$

Finally we fourier transform back to position space. The functional measure in momentum space is related to the functional measure in position space by

$$\lim_{\Delta k\to 0}\prod_{\vec{k}}\mathcal{D}\widetilde{\phi}_{\vec{k}}=\lim_{\Delta x\to 0}\det((\Delta x)^3\,e^{-i\vec{k}\cdot\vec{x}})\prod_{\vec{x}}\mathcal{D}\phi(\vec{x})\,, \tag{2.3.7}$$

where the determinant is of an infinite dimensional matrix where the rows are the different $\vec{k}$ and the columns are the different $\vec{x}$ and the matrix element for "row" $\vec{k}$ and column

$\vec{x}$ is $(\Delta x)^3\, e^{-i\vec{k}\cdot\vec{x}}$. The determinant is then some constant that is independent of $\phi(x)$. Dropping overall constants, we then find

$$\begin{aligned}\mathcal{Z} &= \int \mathcal{D}\phi \,\exp\left(i\int d^4x\,\tfrac{1}{2}\,\partial_\mu\phi(x)\partial^\mu\phi(x) - \tfrac{1}{2}\,(m^2 - i\,\epsilon)\phi^2(x)\right)\\ &= \int \mathcal{D}\phi\, e^{i\int d^4x\,\mathcal{L}(x)}\,, \end{aligned}\tag{2.3.8}$$

where now the functional integral signifies that we integrate over $\phi(x)$ for every space-time point x^μ.

We put in the source terms in the same way as we did before, namely

$$\mathcal{Z}(J) = \int \mathcal{D}\phi\,\exp\left(i\int d^4x\,\tfrac{1}{2}\,\partial_\mu\phi(x)\partial^\mu\phi(x) - \tfrac{1}{2}\,(m^2 - i\,\epsilon)\phi^2(x) + J(x)\phi(x)\right)\,.\tag{2.3.9}$$

We can again evaluate $\mathcal{Z}(J)$ by completing the square,

$$\begin{aligned}\mathcal{Z}(J) &= \mathcal{C}\int\mathcal{D}\phi\,\exp\left(-i\int d^4x\,\tfrac{1}{2}\big[\phi(x) - JM^{-1}(x)\big](\partial^2 + m^2 - i\,\epsilon)\big[\phi(t) - M^{-1}J(x)\big]\right)\\ &\qquad\times\exp\left(i\int d^4x\,\tfrac{1}{2}\,JM^{-1}(x)(\partial^2 + m^2 - i\,\epsilon)M^{-1}J(x)\right)\,,\end{aligned}\tag{2.3.10}$$

where now

$$M^{-1}J(x) = \int d^4y M^{-1}(x,y)J(y)\tag{2.3.11}$$

and

$$(\partial^2 + m^2 - i\,\epsilon)M^{-1}(x,y) = \delta^4(x-y)\,.\tag{2.3.12}$$

Thus, $M^{-1}(x,y) = i\,G_F(x-y)$ and $\mathcal{Z}(J)$ can be written as

$$\mathcal{Z}(J) = \mathcal{Z}(0)\exp\left(-\int d^4x d^4y\,\tfrac{1}{2}\,J(x)G_F(x-y)J(y)\right)\,.\tag{2.3.13}$$

Correlators work the same way as for the single oscillator, namely

$$\langle T[\phi(x_1)\phi(x_2)\dots\phi(x_n)]\rangle = \mathcal{Z}^{-1}\prod_{j=1}^{n} -i\frac{\delta}{\delta J(x_j)}\,\mathcal{Z}(J)\Big|_{J=0}\,,\tag{2.3.14}$$

where now we use that $\frac{\delta J(y)}{\delta J(x)} = \delta^4(x-y)$. Happily, there is no change to the structure of Wick's theorem either:

$$\begin{aligned}\langle T[\phi(x_1)\phi(x_2)\dots\phi(x_n)]\rangle &= \sum_{\text{pairs}}\prod_{k=1}^{n/2}\mathcal{Z}^{-1}(0)\frac{-i\,\delta}{\delta J(x_{I(k)})}\,\frac{-i\,\delta}{\delta J(x_{J(k)})}\,\mathcal{Z}(J)\Big|_{J(t)=0}\\ &= \sum_{\text{pairs}}\prod_{k=1}^{n/2}G_F(x_{I(k)} - x_{J(k)}) = \sum_{\text{pairs}}\prod_{k=1}^{n/2}\langle T[\phi(x_{I(k)})\phi(x_{J(k)})]\rangle\,.\end{aligned}\tag{2.3.15}$$

Chapter 3

Scalar Perturbation theory

3.1 Introduction

Up to now we have been considering the quantum field theory for a free scalar particle. But this is rather boring, and besides in the real world particles *do* interact with each other. In the next few lectures we will describe one particular interacting scalar field theory, which goes under the name of ϕ^4 theory ("phi fourth theory").

For our interacting theory we will include an extra term in the Lagrangian density,

$$\mathcal{L}_I(x) = -\frac{1}{4!}\,\lambda\,\phi^4(x)\,. \tag{3.1.1}$$

The parameter λ is set by hand and is called the coupling. The factor of $\frac{1}{4!}$ is for later convenience. The interaction term shifts the Hamiltonian density by $\mathcal{H}_I(x) = +\frac{1}{4!}\,\lambda\,\phi^4$, hence the energy is bounded from below for positive λ. The interaction term is local, meaning that all four fields are evaluated at the same space-time point x^μ. The locality is necessary to keep the time ordering superfluous when there is spacelike separation, $(x-y)^2 < 0$. If it were nonlocal then in general

$$[\mathcal{L}_I(x), \mathcal{L}_I(y)] \neq 0 \qquad \text{if } (x-y)^2 < 0\,. \tag{3.1.2}$$

An example of a nonlocal but Lorentz invariant interaction term is

$$\mathcal{L}_I(x) = -\frac{\lambda}{2}\,\overline{\phi^2}(x)\overline{\phi^2}(x)\,, \tag{3.1.3}$$

where $\overline{\phi^2}(x)$ is defined as

$$\overline{\phi^2}(x) \equiv \frac{1}{2}\int d^4y\,\phi(x)G_F(x-y)\phi(y)\,, \tag{3.1.4}$$

and where $G_F(x-y)$ is the free Feynman propagator.

In $3+1$ dimensions the dimension of $\phi(x)$ is 1. Since $\mathcal{L}_I$ has dimension 4, this means that λ has dimension 0. Theories with dimensionless couplings are called renormalizable for reasons to be explained later in the course.

3.2 Perturbation theory for the simple harmonic oscillator

3.2.1 Feynman diagrams

As we have done previously, it is useful to set the stage by considering the analogous theory for the simple harmonic oscillator. Since the harmonic oscillator can be understood as a $0+1$ dimensional scalar field theory, we will replace the frequency ω_0 with m_0 to put it in the form of a particle mass. The interaction term we add to the Lagrangian is

$$L_I(t) = -\frac{1}{4!}\,\lambda\,\phi^4(t)\,, \tag{3.2.1}$$

and so the Hamiltonian has the extra term $H_I = \frac{1}{4!}\,\lambda\,\phi^4(t)$. With this extra term, the harmonic oscillator becomes the anharmonic oscillator which cannot be solved analytically. However, if λ is small then one can get a good approximation to the solutions using perturbation theory. Recall that $\phi(t)$ has dimension $-1/2$, while L_I has to have dimension 1. Hence λ has dimension 3. Thus, when we say that λ is small, it means small compared to some dimension 3 scale in the theory. The only such available quantity is m_0^3, thus we assume $\lambda << m_0^3$.

Our main interest will be the correlators in the presence of the interaction term. Hence the normalized n-point time-ordered correlator is

$$\frac{\langle T[\phi(t_1)\phi(t_2)\dots\phi(t_n)]\rangle}{\langle 0|0\rangle} = \mathcal{Z}^{-1}(0)\prod_{j=1}^{n}\frac{-i\,\delta}{\delta J(t_j)}\,\mathcal{Z}(J)\Big|_{J=0}\,, \tag{3.2.2}$$

where now

$$\mathcal{Z}(J) = \mathcal{C}\int\mathcal{D}\phi\,\exp\left(i\int dt\,\left(\tfrac{1}{2}\,\dot\phi^2(t) - \tfrac{1}{2}\,m_0^2\,\phi^2(t) - \tfrac{1}{4!}\lambda\,\phi^4(t) + J(t)\phi(t)\right)\right)\,. \tag{3.2.3}$$

The vacuum state $|0\rangle$ is understood to be that of the interacting theory. Alas, with the interaction term $\mathcal{Z}(J)$ is non-Gaussian and cannot be solved exactly. Our strategy will be to find $\mathcal{Z}(J)$ perturbatively, by expanding about $\lambda = 0$. Hence, we have

$$\mathcal{Z}(J) = \mathcal{C}\int\mathcal{D}\phi\sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{i}{4!}\lambda\right)^n\prod_{j=1}^{n}\int dt_j\phi^4(t_j)\exp\left(i\int dt\,\left(\tfrac{1}{2}\,\dot\phi^2(t) - \tfrac{1}{2}\,m_0^2\,\phi^2(t) + J(t)\phi(t)\right)\right)\,. \tag{3.2.4}$$

Thus, the perturbative expansion of $\mathcal{Z}(J)$ can be written as a sum over correlators for the free theory. Each insertion of a $\int dt\phi^4(t)$ term we will call a vertex.

Let us start by evaluating the expansion for $\mathcal{Z}(0)$, which we will need anyhow since it appears in (3.2.2). Here we find

$$\mathcal{Z}(0) = \langle 1\rangle_{\text{free}} - \frac{i}{4!}\,\lambda\int dt\langle\phi^4(t)\rangle_{\text{free}} + \frac{1}{2}\frac{(-i)^2}{(4!)^2}\,\lambda^2\int dt_1 dt_2\langle T[\phi^4(t_1)\phi^4(t_2)]\rangle_{\text{free}} + \dots \tag{3.2.5}$$

We then use Wick's theorem to evaluate each term in the sum. Notice that the interaction terms automatically come out time ordered. In fact, we see that we can write $\mathcal{Z}(0)$ as

$$\mathcal{Z}(0) = \langle T[e^{i\int dt L_I(t)}]\rangle_{\rm free}\,, \tag{3.2.6}$$

and the full n-point correlator as

$$\frac{\langle T[\phi(t_1)\phi(t_2)\dots\phi(t_n)]\rangle}{\langle 0|0\rangle} = \frac{\langle T[\phi(t_1)\dots\phi(t_n)e^{i\int dt L_I(t)}]\rangle_{\rm free}}{\langle T[e^{i\int dt L_I(t)}]\rangle_{\rm free}}\,, \tag{3.2.7}$$

where everything is evaluated using the free theory.

Assuming that we choose an overall constant so that $\langle 1\rangle_{\rm free}=1$, and we further let t run from $-T/2$ to $T/2$, we then have that the term with one vertex is

$$V_1 \equiv -\frac{i}{4!}\lambda\int dt\langle\phi^4(t)\rangle_{\rm free} = -\frac{i\,3}{4!}\lambda\,(G_F(0))^2\int dt = -\frac{i}{32\,m_0^2}\lambda T\,, \tag{3.2.8}$$

which diverges as $T\to\infty$. We used here that $G_F(0)=\frac{1}{2m_0}$ and that there are three inequivalent choice of pairs.

The next term in the sum, which has two vertices, is itself made up of three types of terms. The first type has all Wick contractions between ϕ's in the same vertex. The next type has two of the ϕ fields within each vertex contracted together and the other two fields contracted with the leftover fields in the other vertex. Finally, the third type has all four fields in a vertex Wick contracted with a field in the other vertex.

The first type of term is $V_{2,1}\equiv\frac{1}{2}(V_1)^2$. The second type is

$$\begin{aligned} V_{2,2} &\equiv -\frac{1}{2}\left(\frac{6}{4!}\right)^2 2\lambda^2\,(G_F(0))^2\int dt_1dt_2\,(G_F(t_1-t_2))^2 \\ &= -\frac{1}{2}\left(\frac{6}{4!}\right)^2 2\lambda^2\frac{1}{(2m_0)^4}\int dTd\tau e^{-2im_0|\tau|} = -\left(\frac{1}{4(2m_0)^2}\right)^2\lambda^2\frac{2\,T}{2\,i\,m_0} = \frac{i}{256\,m_0^5}\lambda^2 T\,. \end{aligned} \tag{3.2.9}$$

Here, there is a factor of 6 for each vertex for the number of ways of choosing two fields to Wick contract. There is an additional factor of 2 for the different ways of pairing the leftover fields from the first vertex with the leftover fields from the second. The third type of term is

$$\begin{aligned} V_{2,3} &\equiv -\frac{1}{2}\left(\frac{1}{4!}\right)^2 4!\,\lambda^2\int dt_1dt_2\,(G_F(t_1-t_2))^4 \\ &= -\frac{1}{48}\lambda^2\frac{1}{(2m_0)^4}\int dt_1dt_2e^{-4im_0|t_1-t_2|} = -\frac{1}{48}\lambda^2\frac{1}{(2m_0)^4}\frac{2\,T}{4\,i\,m_0} = \frac{i}{6\cdot 256\,m_0^5}\lambda^2 T\,. \end{aligned} \tag{3.2.10}$$

We can then write

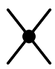
$$\mathcal{Z}(0) = \exp(V_1+V_{2,2}+V_{2,3}) = \exp\left(-i\left[\frac{\lambda}{32\,m_0^2}-\frac{7\,\lambda^2}{1536\,m_0^5}\right]T\right)+\mathrm{O}(\lambda^3)\,. \tag{3.2.11}$$

The term inside the square brackets is the correction to the ground-state energy to second order in perturbation theory. This should be expected, since $\mathcal{Z}(0) = \langle 0, T/2|0, -T/2\rangle = \langle 0|e^{-iHT}|0\rangle$ up to an overall constant. The constant was chosen to cancel out the part from the unperturbed Hamiltonian, thus we are left with $\mathcal{Z}(0) = e^{-i\Delta E\, T}$, where ΔE is the correction to the ground-state energy.

Let's now repeat what we have just done using some pictures, called Feynman diagrams. A Feynman diagram is made from a set of building blocks. Each building block has an associated Feynman rule and the Feynman diagrams have a set of rules that one follows called Feynman rules. Each Wick contraction joins two fields, resulting in a propagator. We will use as a symbol of this propagator a straight line with a dot and the relevant time coordinate at each end:

t_1 $t_2 = G_F(t_1 - t_2)\,, \qquad (3.2.12)$

where $G_F(t_1-t_2)$ is the rule associated with this symbol. For a vertex the symbol and associated rule are

 $= -i\,\lambda \displaystyle\int dt \qquad (3.2.13)$

The factor of $\frac{1}{4!}$ does not appear in (3.2.13) because we are assuming that there are four other fields that will attach to the vertex and there are 4! ways for choosing how these four fields pair up with the four fields that make up the vertex.

The Feynman diagram for V_1 is then

$V_1 =$  $, \qquad (3.2.14)$

with two "ears" attached to a vertex. This is an example of a "vacuum bubble" diagram. Using the rules as stated so far it seems it should be equal to

$$V_1 \stackrel{?}{=} -i\,\lambda \frac{T}{(2m_0)^2}\,. \qquad (3.2.15)$$

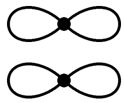
But this is 8 times larger than what was computed in (3.2.8) because we overcounted. The Feynman rule for the vertex assumed that there were 24 ways to make pairs, but as we argued earlier there are only 3.

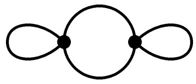
However, we can fix this by taking into account symmetry factors. To see how this works, let us take four objects and order them. There are of course 4! inequivalent ways of doing this. Let us further say that the first object is paired with the second and the third with the fourth. However, exchanging 1 with 2 gives the same pairing, so if we are only counting distinct pairs we should divide by a factor of 2. Likewise, we should divide by a factor of 2 for the exchange of 3 with 4. Finally, the pairing is invariant under the exchange of (1,2) with (3,4), leading us to divide by a further factor of 2. The complete symmetry factor is then $S = 2 \times 2 \times 2 = 8$. Hence, for every diagram we should include an additional rule:

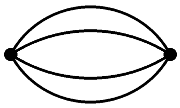
Additional Feynman rule: *Divide by the symmetry factor.*

In the case of V_1, we then find (3.2.8). In general, for a propagator that starts and ends at the same vertex there is a symmetry factor of 2. Furthermore, if there are n propagators running between the same two vertices then there is a factor of $n!$. There is also a symmetry factor of 2 for invariance under a reflection and n for an invariance under a rotation of of $2\pi/n$. There is also a factor of $n!$ if a diagram is unchanged under a permutation of n identical subdiagrams.

For the terms that are second order terms in λ, the three different vacuum bubble diagrams are

$V_{2,1} =$ 

$V_{2,2} =$ 

$V_{2,3} =$  . (3.2.16)

The symmetry factor for $V_{2,1}$ is 2 multiplied by the symmetry factors already present in the two V_1 terms. The symmetry factor for $V_{2,2}$ is $2 \times 2 \times 2 \times 2 = 16$, where there is a factor of 2 for each "ear", another factor of 2 for the two propagators running between the two vertices and a final factor of 2 for the reflection symmetry (which is the same as rotation by $2\pi/n$ so there is no additional factor of 2 for this). Finally, the symmetry factor for $V_{2,3}$ is $4! \times 2 = 48$, where the factor of 4! comes from the symmetry of the 4 propagators and the factor of 2 for the reflection symmetry.

The diagram for $V_{2,1}$ has an obvious difference from the $V_{2,2}$ and $V_{2,3}$ diagrams. The first diagram is disconnected, meaning that there are two or more parts which are not connected to each other by at least one propagator. The latter two diagrams are called connected. It follows from (3.2.11) that

$$\log \mathcal{Z}(0) = V_1 + V_{2,2} + V_{2,3} + \mathrm{O}(\lambda^3)\,. \tag{3.2.17}$$

Notice that the terms in the sum are connected diagrams. In fact this carries out to all orders in λ, namely

$$\log \mathcal{Z}(0) = \sum_{j=1}^{\infty} \mathcal{V}_j\,. \tag{3.2.18}$$

where $\mathcal{V}_j$ is the j^{th} connected vacuum bubble under some ordering scheme. Eq. (3.2.18) is straightforward to verify. Suppose we consider one of the diagrams in $\mathcal{Z}(0)$, which has N_1 copies of $\mathcal{V}_1$, N_2 copies of $\mathcal{V}_2$, *etc.*(we are assuming that only a finite number of the N_j are nonzero). Let us label this diagram as $\widehat{V}_{N_1,N_2\ldots}$. According to the Feynman rules, we have that

$$\widehat{V}_{N_1,N_2\dots} = \prod_{j=1}^{\infty} \frac{1}{N_j!} \left(\mathcal{V}_j\right)^{N_j} . \tag{3.2.19}$$

The different disconnected diagrams are distinguished by the N_j. Therefore, the sum over all diagrams is

$$\mathcal{Z}(0) = \sum_{N_1,N_2\dots=0}^{\infty} \widehat{V}_{N_1,N_2\dots} = \prod_{j=1}^{\infty}\sum_{N_j=0}^{\infty} \frac{1}{N_j!} \left(\mathcal{V}_j\right)^{N_j} = \prod_{j=1}^{\infty} \exp(\mathcal{V}_j) = \exp\left(\sum_{j=1}^{\infty}\mathcal{V}_j\right) \tag{3.2.20}$$

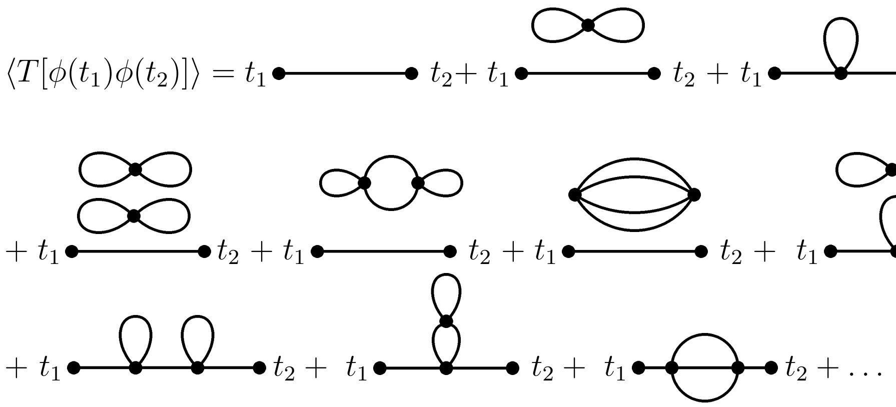
thus (3.2.18) holds.

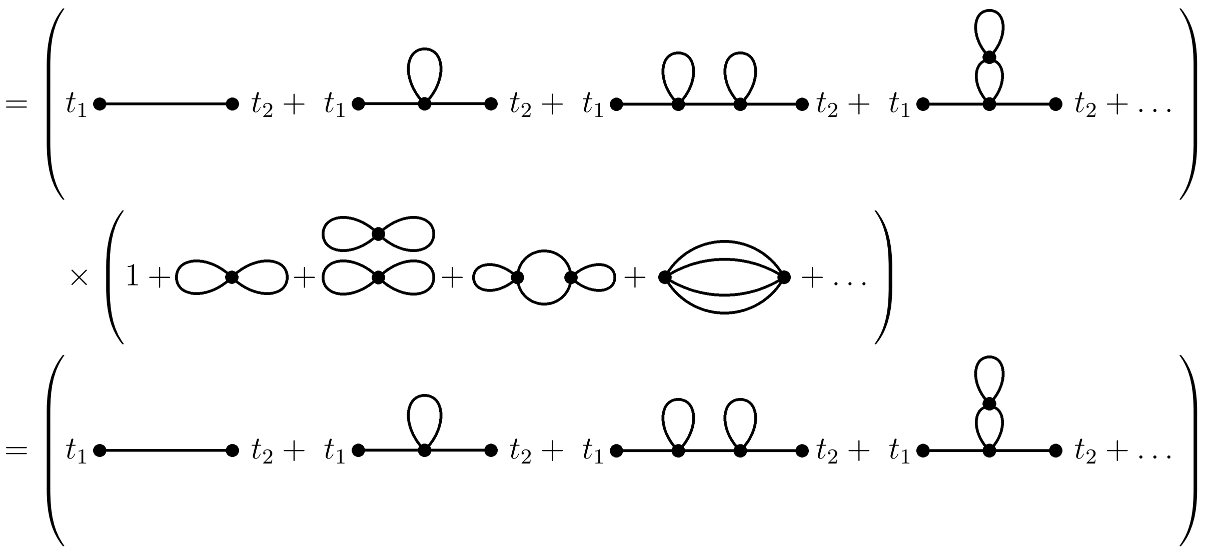
3.2.2 2-point correlators for the perturbed harmonic oscillator

Let us now consider the 2-point correlator for the anharmonic oscillator. We will compute this by explicitly inserting two fields into the path integral, hence the correlator is

$$\begin{aligned}
&\langle T[\phi(t_1)\phi(t_2)]\rangle = \\
&\int \mathcal{D}\phi\, \phi(t_1)\phi(t_2) \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i}{4!}\lambda\right)^n \prod_{j=1}^{n} \int dt'_j \phi^4(t'_j) \exp\left(i\int dt\, \left(\tfrac{1}{2}\,\dot\phi^2(t) - \tfrac{1}{2}\,m_0^2\phi^2(t)\right)\right) \\
&= \langle T[\phi(t_1)\phi(t_2)]\rangle_{\text{free}} - \frac{i}{4!}\,\lambda \int dt' \langle T[\phi(t_1)\phi(t_2)\phi^4(t')]\rangle_{\text{free}} + \\
&\qquad + \frac{1}{2}\frac{(-i)^2}{(4!)^2}\,\lambda^2 \int dt'_1 dt'_2 \langle T[\phi(t_1)\phi(t_2)\phi^4(t'_1)\phi^4(t'_2)]\rangle_{\text{free}} + \dots .
\end{aligned} \tag{3.2.21}$$

The sum over diagrams would then look like

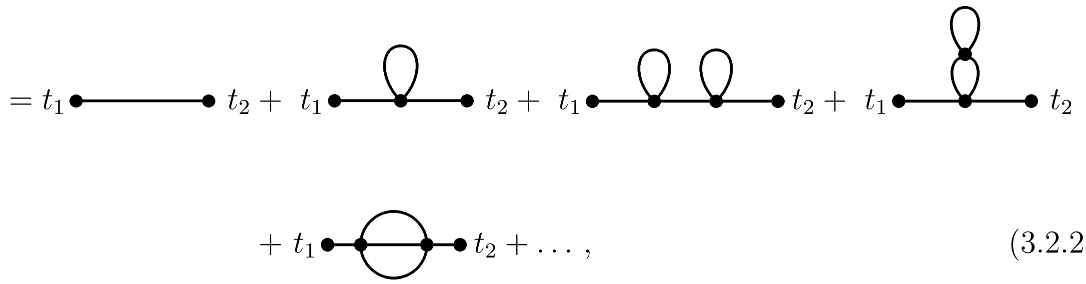




(3.2.22)

Therefore, the normalized propagator is

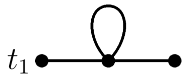
$$\frac{\langle T[\phi(t_1)\phi(t_2)]\rangle}{\langle 0|0\rangle}$$



(3.2.23)

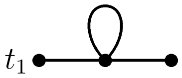
where all vacuum bubbles have divided out. From now on, we will assume that the n-point functions are the normalized ones.

Let us now start computing the terms in (3.2.23). The first term is the usual free propagator, $G_F(t_1-t_2)$. The second term, according to the Feynman rules is



$$= \frac{1}{2}(-i\lambda)\int dt\,\frac{1}{(2m_0)^3}e^{-im_0|t_1-t|-im_0|t_2-t|}\,. \tag{3.2.24}$$

where the factor of $\frac{1}{2}$ comes from the symmetry factor of having a propagator's two ends attached to the same vertex. Note that there is no symmetry factor for reflection because the ends of the diagram are attached to fields at different times. The integral is straightforward and yields

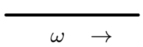


$$= \frac{-i\lambda}{8m_0^2}\left(|t_1-t_2|-\frac{i}{m_0}\right)G_F(t_1-t_2)\,. \tag{3.2.25}$$

We could keep going on, but at this point it is much more instructive, and easier, to switch over to frequency space. We will first need to rederive the Feynman rules. We have previously seen that

$$\int dt_1 dt_2\, e^{i\omega_1 t_1+i\omega_2 t_2}G_F(t_1-t_2) = 2\pi\delta(\omega_1+\omega_2)\frac{i}{\omega_1^2-m_0^2+i\,\epsilon} = 2\pi\delta(\omega_1+\omega_2)\,\widetilde{G}(\omega_1)\,. \tag{3.2.26}$$

We will use the following symbol for the propagator in frequency space



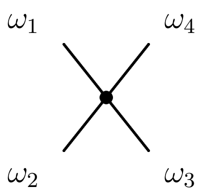
$$= \frac{i}{\omega^2-m_0^2+i\,\epsilon}\,, \tag{3.2.27}$$

where the arrow shows the direction that ω is running. The δ-function in (3.2.26) ensures that the energy ω entering the propagator on the left equals the energy exiting on the right.

For the vertex, we Fourier transform to find

$$\begin{aligned}\frac{-i\lambda}{4!}\int dt\,\phi^4(t) &= \frac{-i\lambda}{4!}\int dt\int\left(\prod_{j=1}^4\frac{d\omega_j}{2\pi}\right)e^{i(\omega_1+\omega_2+\omega_3+\omega_4)t}\,\widetilde{\phi}(\omega_1)\widetilde{\phi}(\omega_2)\widetilde{\phi}(\omega_3)\widetilde{\phi}(\omega_4)\\ &= \frac{-i\lambda}{4!}\int\left(\prod_{j=1}^4\frac{d\omega_j}{2\pi}\right)2\pi\delta(\omega_1+\omega_2+\omega_3+\omega_4)\,\widetilde{\phi}(\omega_1)\widetilde{\phi}(\omega_2)\widetilde{\phi}(\omega_3)\widetilde{\phi}(\omega_4)\,.\end{aligned} \tag{3.2.28}$$

The δ-function ensures energy conservation at the vertex. There is no change in the combinatiorics, so the rule we use is then

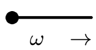


$$= (-i\,\lambda)\,2\pi\delta(\omega_1+\omega_2+\omega_3+\omega_4)$$

where the labels show the energy going into the vertex. We will keep the external fields in coordinate space, so using

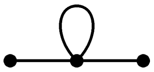
$$\phi(t) = \int \frac{d\omega}{2\pi}\, e^{-i\omega t}\, \widetilde{\phi}(\omega) \tag{3.2.29}$$

we include the additional rule

t  $= e^{-i\omega t}\,.$ (3.2.30)

We then integrate over all energies that are left undetermined by the δ-functions and divide by an symmetry factors to arrive at the result.

Going back to the diagram in (3.2.24) and applying the frequency space Feynman rules we find

t_1  t_2

$$\begin{aligned} &= \int \frac{d\omega}{2\pi}\, e^{-i\omega(t_1-t_2)} \frac{i}{\omega^2 - m_0^2 + i\epsilon} \left[\frac{-i\lambda}{2} \int \frac{d\omega'}{2\pi}\, \frac{i}{\omega'^2 - m_0^2 + i\epsilon}\right] \frac{i}{\omega^2 - m_0^2 + i\epsilon} \\ &= \int \frac{d\omega}{2\pi}\, e^{-i\omega(t_1-t_2)} \frac{i}{\omega^2 - m_0^2 + i\epsilon} \left[\frac{-i\lambda}{2}\, \frac{(2\pi i)i}{(2\pi)(-2m_0)}\right] \frac{i}{\omega^2 - m_0^2 + i\epsilon} \\ &= \int \frac{d\omega}{2\pi}\, e^{-i\omega(t_1-t_2)} \frac{i}{\omega^2 - m_0^2 + i\epsilon} \left[\frac{-i\lambda}{4m_0}\right] \frac{i}{\omega^2 - m_0^2 + i\epsilon}\,, \end{aligned} \tag{3.2.31}$$

where the integral over ω' was done by closing the contour in the upper half plane, which circles the pole at $\omega' = -m_0 + i\epsilon$ and which has residue $-\frac{1}{2m_0}$.

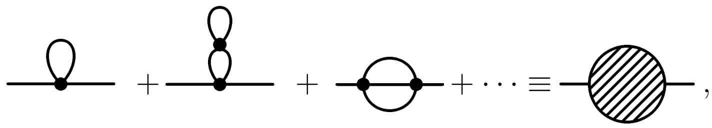
The diagram in (3.2.31) is an example of a one-loop diagram. Because of this loop there is one-more undetermined frequency than in t_1 t_2. A loop is a path of propagators that starts and stops at the same vertex, perhaps going through other vertices along the loop. Suppose we have a set of frequencies in our diagram that satisfy the conservation rules at every vertex. If we then add some frequency into the loop, the frequency conservation still holds since if a loop enters a vertex it must also exit it. This can be applied to every independent loop (some loops are equivalent to the sum of two or more loops.). Hence, for every loop there is an undetermined frequency that must be integrated over.

There are infinitely many Feynman diagrams for the two-point function. However, we can sum up a large class of these diagrams, leaving a subclass of diagrams to compute, albeit still an infinite number. This is particularly easy to do in frequency space. Let us define a class of diagrams called "one particle irreducible" (1PI). A 1PI diagram is any connected diagram that stays connected after any internal propagator in the diagram is cut. An internal propagator is one that is not connected to the external fields in the correlator (the $\phi(t_1)$ and $\phi(t_2)$ in the case of the 2-point correlator.) Thus, the first, second, fourth and fifth diagrams in the sum in (3.2.23) are 1PI, while the third diagram is not.

We next consider the *truncated* 1PI diagrams, where all external propagators are lopped off. This is accomplished by dividing by a propagator for every endpoint. A truncated diagram is distinguished from a nontruncated one in that the truncated diagram has no dots on its ends. Note that by truncating the propagator we end up with $(\text{———})^{-1} = -i(\omega^2 - m_0^2 + i\epsilon)$, which is the inverse propagator since we divided by two propagators, one for each end-point.

If we now define the bare *self energy* $\Sigma_0(\omega)$ in terms of the sum over all truncated 1PI diagrams for the two-point function which include at least one vertex,

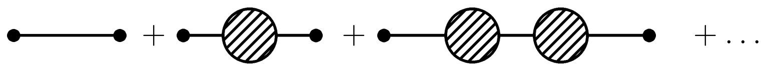
$$-i\,\Sigma_0(\omega) =$$



$$(3.2.32)$$

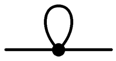
then the complete nontruncated propagator, which we call $G_0^{(2)}(\omega)$, is[1]

$$G_0^{(2)}(\omega) =$$



$$= \frac{i}{\omega^2 - m_0^2 + i\epsilon} \times \sum_{n=0}^{\infty} \left(\frac{\Sigma_0(\omega)}{\omega^2 - m_0^2 + i\epsilon} \right)^n = \frac{i}{\omega^2 - m_0^2 - \Sigma_0(\omega) + i\epsilon}\,. \tag{3.2.33}$$

Looking back at (3.2.31), we see that the one-loop contribution to $-i\Sigma_0(\omega)$ is



$$= \frac{-i\lambda}{2} \int \frac{d\omega'}{2\pi} \frac{i}{\omega'^2 - m_0^2 + i\epsilon} = \frac{-i\lambda}{4m_0}\,. \tag{3.2.34}$$

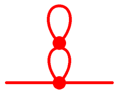
Thus, we see that the one-loop contribution shifts the position of the poles in the propagator, with the new poles at $\omega^2 = m_0^2 + \frac{\lambda}{4m_0} - i\epsilon$. Thinking of the harmonic oscillator as a $0+1$ dimensional scalar field theory with the mass-shell condition $\omega^2 = m^2$, this shift in the pole positions corresponds to a shift in the particle mass to $m = m_0 + \frac{\lambda}{8m_0^2}$ to leading order in λ. m_0, the mass that appears in the Lagrangian is called the bare mass. The new mass m is called the *physical mass* and would be the measured mass in an experiment.

Let us now show how this shift in the mass is related to one of the most basic results taught in a first semester course in quantum mechanics. In $0+1$ dimensions the single particle state is the first excited state of the harmonic oscillator. Hence the mass should equal the energy difference between the first excited state and the ground-state, and the mass-shift is the perturbative correction to this difference. We already showed that the first order correction to the ground-state energy is $\Delta E_0 = \frac{\lambda}{32m_0^2}$. For the first excited state the correction is $\Delta E_1 = \frac{5\lambda}{32m_0^2}$. Hence, the shift in the particle's mass is $\Delta E_1 - \Delta E_0 = \frac{\lambda}{8m_0^2}$.

The next diagram in (3.2.32) gives a second order correction to the pole position.

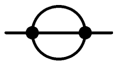
[1]The superscript (2) is for a 2-point function. The the subscript 0 signifies the "bare" propagator where no countertems are included. The concept of a counterterm will be more fully explained later.

Instead of doing this diagram (we save it for an exercise), Here is how to do the exercise:



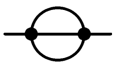
$$
\begin{aligned}
&= \frac{(-i\lambda)^2}{2\cdot 2}\int\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\left(\frac{i}{\omega_1^2-m_0^2+i\,\epsilon}\right)^2\frac{i}{\omega_2^2-m_0^2+i\,\epsilon}\\
&= \frac{(-i\lambda)^2}{2\cdot 2}\frac{1}{2m_0}\int\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\left(\frac{i}{\omega_1^2-m_0^2+i\,\epsilon}\right)^2\\
&= \frac{(-i\lambda)^2}{2\cdot 2}\frac{1}{2m_0}\frac{-2i}{8m_0^3}=\frac{i\,\lambda^2}{32\,m_0^4}\,,
\end{aligned}
\tag{3.2.35}
$$

let us move on to the third one, called the *sunset* diagram. Here we find that

$\omega\rightarrow$ 

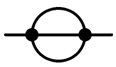
$$
=\frac{1}{6}(-i\lambda)^2\int\frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\frac{i}{\omega_1^2-m_0^2+i\epsilon}\frac{i}{\omega_2^2-m_0^2+i\epsilon}\frac{i}{(\omega-\omega_1-\omega_2)^2-m_0^2+i\epsilon}\,,
\tag{3.2.36}
$$

where the factor of $\frac{1}{6}$ is the symmetry factor. Notice that there are two integrations since we have a 2-loop graph. There are three propagators in the loops, where the energy through the third is determined from the other two by energy conservation in the vertices. Doing the integral over ω_1 first, we find after closing off the contour in the lower half-plane that the contour encircles two poles at $\omega_1=m_0-i\epsilon$ and $\omega-\omega_2+m_0-i\epsilon$ and the resulting integral is

$\omega\rightarrow$ 

$$
=\frac{1}{6}(-i\lambda)^2\frac{2}{2m_0}\int\frac{d\omega_2}{2\pi}\frac{i}{(\omega_2-\omega)^2-(2m_0)^2+i\epsilon}\frac{i}{\omega_2^2-m_0^2+i\epsilon}\,,
\tag{3.2.37}
$$

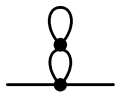
Closing off the ω_2 contour in the lower half plane we encircle two poles at $\omega_2=m_0-i\epsilon$ and $\omega_2=\omega-2m_0-i\epsilon$. We then find

$\omega\rightarrow$ 

$$
=-\frac{\lambda^2}{8m_0^2}\frac{i}{\omega^2-(3m_0)^2+i\epsilon}\,.
\tag{3.2.38}
$$

Combining all contributions up to $O(\lambda^2)$, for $\omega^2\approx m_0^2$, we have that

$$
\omega^2-m_0^2-\Sigma_0(\omega)\approx\omega^2-m_0^2-\frac{\lambda}{4m_0^2}-\Delta m_{db}^2+\frac{\lambda^2}{64m_0^4}+\frac{\lambda^2}{512m_0^6}(\omega^2-m_0^2)\,,
\tag{3.2.39}
$$

where Δm_{db}^2 is the contribution from  , which is of order λ^2. Hence, up to second order in λ we have that

$$
\frac{i}{\omega^2-m_0^2-\Sigma_0(\omega)}\approx\frac{i\,Z}{\omega^2-m_0^2-\Delta m^2}\,,
\tag{3.2.40}
$$

where

$$Z = 1 + \delta_Z = 1 - \frac{\lambda^2}{512 m_0^6} + \mathrm{O}(\lambda^3)\,. \tag{3.2.41}$$

and where $\delta_Z = \frac{\partial}{\partial\omega^2}\Sigma_0(\omega)|_{\omega^2=m_0^2}$.

The factor of Z is called the "(on-shell) wave-function renormalization" for ϕ. To see where it comes from, let us go back to the free theory case. Here we have that

$$\phi(0)|0\rangle_{\mathrm{free}} = \frac{1}{\sqrt{2m_0}}\, a^\dagger|0\rangle_{\mathrm{free}}\,. \tag{3.2.42}$$

If we define the free n-particle state as

$$|n\rangle_{\mathrm{free}} \equiv \frac{\sqrt{2nm_0}}{\sqrt{n!}}\,(a^\dagger)^n|0\rangle_{\mathrm{free}}\,, \tag{3.2.43}$$

then $|\langle n|\phi(0)|0\rangle_{\mathrm{free}}|^2 = \delta_{n1}$, that is $\phi(0)$ creates a one-particle state. In the interacting theory we assume that there are still n-particle states, normalized to be $\langle n|m\rangle = 2m_n\delta_{mn}$, but the total mass no longer satisfies $m_n = nm_0$, while the inner products have the more general form $|\langle n|\phi(0)|0\rangle|^2 = Z_n$. For either the free or the interacting case we can write the propagator as

$$\begin{aligned}\langle T[\phi(t)\phi(0)]\rangle &= \sum_n \langle 0|e^{iHt}\phi(0)e^{-iHt}|n\rangle\frac{1}{2m_n}\langle n|\phi(0)|0\rangle \qquad t>0\,,\\ &= \sum_n \langle 0|\phi(0)|n\rangle\frac{1}{2m_n}\langle n|e^{iHt}\phi(0)e^{-iHt}|0\rangle \qquad t<0\end{aligned} \tag{3.2.44}$$

where we have inserted a complete set of states. In the free case we have that $m_n = nm_0$ and $|\langle 0|\phi(0)|n\rangle^2 = \delta_{n1}$ so we get the usual Feynman propagator. In the interacting case, the frequency transform of the propagator is

$$\int dt\, e^{i\omega t}\sum_n \frac{Z_n}{2m_n}e^{-im_n|t|} = \sum_n \frac{i\,Z_n}{\omega^2 - m_n^2 + i\epsilon}\,. \tag{3.2.45}$$

The Z we have found is Z_1 in this expression, and we now see that it is related to the probability for $\phi(0)$ to create a one particle state out of the vacuum. One can choose to get rid of the Z factor by defining a renormalized field $\phi_R(t) = Z^{-1/2}\phi(t)$, in which case the residue at the one particle pole is the same as in the free case. This is the strategy we will use in $3+1$ dimensions.

According to our logic there should be other poles, which we now demonstrate. The poles will occur for the values of ω where $\omega^2 - m_0^2 - \Sigma_0(\omega) = 0$. If $\omega^2 \approx (3m_0)^2$, then the propagator is approximately

$$\frac{i}{\omega^2 - m_0^2 - \Sigma_0(\omega)} \approx \frac{i(\omega^2-(3m_0)^2)}{8m_0^2(\omega^2-(3m_0)^2) - \lambda^2/(8m_0^2)}\,, \tag{3.2.46}$$

which has a pole[2] at $\omega^2 = (3m_0)^2 + \frac{\lambda^2}{64m_0^4}$ and a residue $Z_3 = \frac{\lambda^2}{512m_0^6}$. The probability P_3 to create the three particle state out of the vacuum with $\phi(0)$ is then

$$P_3 = \frac{Z_3/(2m_3)}{\sum Z_n/(2m_n)} \approx \frac{Z_3}{3} = \frac{\lambda^2}{1536\, m_0^6}\,, \tag{3.2.47}$$

where we used that $Z_1 \approx 1$ and that $m_n \approx n m_0$. Note that it is crucial that the residue of the pole be positive in order to have a positive probability.

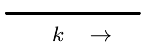
3.3 2-point correlators for the interacting scalar field theory

3.3.1 Similarities and differences with the anharmonic oscillator

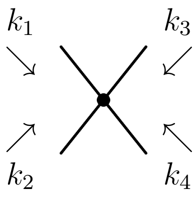
In the $3+1$ dimensional field theory much of the procedure is the same as the $0+1$ field theory. In particular, the correlator is

$$\begin{aligned}
&\langle T[\phi(x)\phi(y)]\rangle = \\
&\int \mathcal{D}\phi\,\phi(x)\phi(y)\sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{i}{4!}\lambda\right)^n\prod_{j=1}^{n}\int d^4x'_j\phi^4(t'_j)\exp\left(i\int d^4x'\ \left(\tfrac{1}{2}\,\partial_\mu\phi\partial^\mu\phi(x') - \tfrac{1}{2}\,m^2\phi^2(x')\right)\right)\\
&= \langle T[\phi(x)\phi(y)]\rangle_{\text{free}} - \frac{i}{4!}\,\lambda\int d^4x'\langle T[\phi(x)\phi(y)\phi^4(x')]\rangle_{\text{free}} + \\
&\qquad + \frac{1}{2}\frac{(-i)^2}{(4!)^2}\,\lambda^2\int d^4x'_1 d^4x'_2\langle T[\phi(x)\phi(y)\phi^4(x'_1)\phi^4(x'_2)]\rangle_{\text{free}} + \dots\\
&= \langle T[\phi(x)\phi(y)e^{i\int d^4x'\mathcal{L}_I(x')}]\rangle_{\text{free}}\,.
\end{aligned} \tag{3.3.1}$$

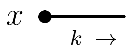
We can still use Wick's theorem, so that diagramatically $\langle T[\phi(x)\phi(y)]\rangle$ is given by (3.2.22) and the normalized propagator is given by (3.2.23), with the t_1 and t_2 in the diagrams replaced by x and y respectively. The Feynman rules are also very similar with



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



$$= (-i\,\lambda)\,(2\pi)^4\delta^4(k_1 + k_2 + k_3 + k_4) \tag{3.3.2}$$



$$= e^{-ik\cdot x}\,,$$

and with the same rules about symmetry factors. We also have that every loop comes with an integration over the 4-momentum through the loop, with measure factor $\frac{d^4k}{(2\pi)^4}$.

[2] The pole position is not precisely at this point to this order in λ because we have dropped some terms which are unimportant for computing the residue, which is our main concern here.

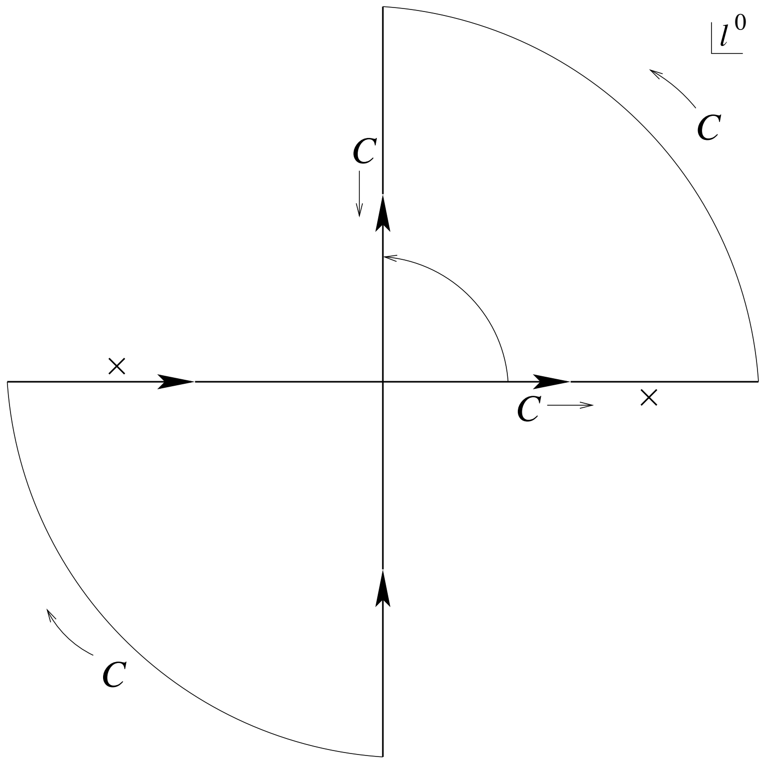
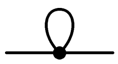


Figure 3.1: Wick rotation of the integral over ℓ^0 along the real axis to the imaginary axis. The closed contour C circles no poles and hence gives 0.

Furthermore, the bare self-energy[3] $\Sigma_0(k^2)$ satisfies the same diagrammatic equation as in (3.2.32).

But not everything is a straightforward generalization of the simple harmonic oscillator because of the infinite degrees of freedom. Let us first consider the order λ contribution to $\Sigma_0(k^2)$ coming from the single bubble diagram. Now we have that

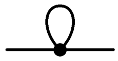
 $$= \frac{-i\lambda}{2} \int \frac{d^4\ell}{(2\pi)^4} \frac{i}{\ell^2 - m^2 + i\epsilon}\,, \tag{3.3.3}$$

where ℓ^μ is the 4-momentum running through the loop. To do this integral, we first examine the integral over ℓ^0. Here, the integral is well defined for fixed $\vec{\ell}$, with a pole below the real axis for positive ℓ^0 and one above the real axis for negative ℓ^0. Since the integrand falls off as $1/(\ell^0)^2$ for large ℓ^0, we can do what is called a Wick (or Euclidean) rotation. Namely, we note that we can rotate the integration axis counterclockwise without changing the value of the integral as long as no poles are crossed (see figure 1). The point is that the closed contour C encircles no poles so its integral is 0. C contains an integral along the real axis, integrations at ∞ which give no contribution since the integrand falls off fast enough, and another contribution from the rotated axis. But this last integration is running backward, so from this we see that the integration running

[3]By Lorentz invariance the self-energy and propagators can only depend on k^2, so we will write their arguments as such.

forward gives the same result as that from the real axis. However, if the rotation had crossed a pole, then our integral over C would be nonzero.

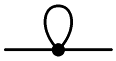
Hence, after the Wick rotation, we can replace ℓ^0 with $i\,\ell^0$ and the integral becomes



$$= \frac{-i\lambda}{2}\int \frac{d^4\ell}{(2\pi)^4}\,\frac{1}{\ell^2+m^2}\,. \tag{3.3.4}$$

where now the integral is over 4-dimensional flat Euclidean space and which has a length squared $\ell^2=\ell_0^2+\ell_1^2+\ell_2^2+\ell_3^2$. The $i\,\epsilon$ has been dropped because the integral no longer passes near a pole[4].

We can convert the integral in (3.3.4) to an integral over spherical coordinates, where here the sphere is the 3-sphere S^3. The integrand has no angular dependence, so it becomes



$$= \frac{-i\lambda}{2}\,\frac{\Omega_3}{(2\pi)^4}\int_0^\infty \frac{\ell^3\,d\ell}{\ell^2+m^2}\,. \tag{3.3.5}$$

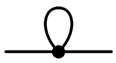
where Ω_3 is the volume of the unit 3-sphere. The volume of an n-dimensional unit sphere can be found using the following trick:

$$\int d^{n+1}x\,e^{-x^2}=\pi^{(n+1)/2}=\Omega_n\int_0^\infty x^n dx\,e^{-x^2}=\frac{\Omega_n}{2}\int_0^\infty z^{(n-1)/2}dze^{-z}=\frac{1}{2}\Omega_n\Gamma((n+1)/2)\,. \tag{3.3.6}$$

Thus,

$$\Omega_n=\frac{2\,\pi^{(n+1)/2}}{\Gamma((n+1)/2)}\,, \tag{3.3.7}$$

and so $\Omega_3=\frac{2\,\pi^2}{\Gamma(2)}=2\,\pi^2$. The integral is now



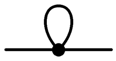
$$= \frac{-i\lambda}{16\pi^2}\int_0^\infty d\ell\,\frac{\ell^3}{\ell^2+m^2}\,. \tag{3.3.8}$$

This integral is obviously divergent. To remove the divergence we have to *regularize* (also called *regulate*) the integral.

The divergence arises because any scalar field with a particular momentum $\widetilde{\phi}(k)$ can couple to any other field with an arbitrarily high momentum. One way to regularize the integral is to assume that there is an upper value for the momentum of these other fields. Perhaps it is because the scalar field theory needs to be replaced with another theory when the momenta get arbitrarily large. Maybe it is because the theory is really on a lattice with very small spacing, which limits the momentum (but also breaks Lorentz invariance and locality). Whatever. For our purposes, we will assume that the integral over ℓ stops at some high value Λ. Presumably the value of Λ is much bigger than k, the

[4] Some books, (e.g. Peskin) put a subscript on the momentum, ℓ_E to signify it is Euclidean. However, whether it is Euclidean or not is usually evident from the sign in front of m^2, so we will neglect this little nicety.

momentum running through the propagator, as well as m. Such a limit on the integral is called an ultraviolet (UV) cut-off, since it cuts off the integral at high momentum. Inserting the cut-off we then get



$$= \frac{-i\lambda}{16\pi^2}\int_0^\Lambda d\ell\,\frac{\ell^3}{\ell^2+m^2} = \frac{-i\lambda}{32\pi^2}\left(\Lambda^2 - m^2\log\frac{\Lambda^2+m^2}{m^2}\right) \approx \frac{-i\lambda}{32\pi^2}\left(\Lambda^2 - m^2\log\frac{\Lambda^2}{m^2}\right), \qquad (3.3.9)$$

where in the last line we have assumed that $\Lambda >> m$. Hence, the correction to m^2 in the propagator is

$$\Delta m^2 = \frac{\lambda}{32\pi^2}\left(\Lambda^2 - m^2\log\frac{\Lambda^2}{m^2}\right), \qquad (3.3.10)$$

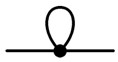
so even if $\lambda << 1$ there can be a big positive shift[5] in m^2.

Even though it looks like we can have a big correction, let us assume that the correction is small so that perturbation theory is valid. It is common practice to express all physical quantities like scattering amplitudes etc., in terms of the physical masses of the particles, that is the positions of the poles, and not in terms of the "bare masses", the mass terms that appear in the Lagrangian. To this end, we will assume that m^2 is the position of the pole. This means that we have to include terms in the Lagrangian to cancel off the Δm^2 terms from the self-energy. Such terms are called *counterterms*. In fact, we will also include counterterms to keep the residue of the poles unity. To this end, we include another type of Feynman diagram, called a counterterm diagram[6]

$$= i\left(\delta_Z(k^2-m^2) - \delta_{m^2}\right) \qquad (3.3.11)$$

so that

$$-i\Sigma(k^2) =$$



$$+$$

$$+\ldots\,. \qquad (3.3.12)$$

We have dropped the 0 subscript on $\Sigma(k^2)$ to indicate this is the physical self-energy. The wave-function counterterm is $\delta_Z = \frac{\partial}{\partial k^2}\Sigma_0(k^2)\big|_{k^2=m^2}$, and as you can see from (3.3.11) and (3.3.12), cancels off the correction to the (k^2-m^2) term in $\Sigma_0(k^2)$. Given $\Sigma(k^2)$, the physical propagator is

$$G^{(2)}(k^2) = \frac{i}{k^2 - m^2 - \Sigma(k^2)}\,. \qquad (3.3.13)$$

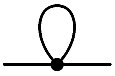
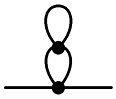
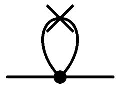
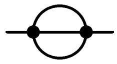
To one-loop order the other counterterm is given by

$$\delta_{m^2} = -\frac{\lambda}{32\pi^2}\left(\Lambda^2 - m^2\log\frac{\Lambda^2}{m^2}\right), \qquad \delta_Z = 0\,. \qquad (3.3.14)$$

[5]It turns out, so far anyway, there are no observed fundamental scalars in nature (the Higgs particle which is presently being searched for at the LHC is believed to be a fundamental scalar). By fundamental we mean a scalar particle not composed of other particles. One possible reason for a lack of fundamental scalars is that the quantum corrections coming from their self interactions push their masses to a very high value, making them unobservable in present day accelerators.

[6]Note that Peskin defines the counter terms slightly differently.

At higher loops the counterterm diagram will incorporate all corrections so that the physical pole stays at m^2 and with the proper residue. The counterterms also need to be included in all sub diagrams, so for example to two loop order the diagrams that contribute to the self-energy are

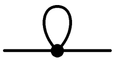
$-i\,\Sigma(k^2) =$  $+$  $+$  $+$  $+$ **?** $+$  $+\dots,$

(3.3.15)

where the counterterm in the second diagram includes terms up to order λ^2, while the counterterm in the fourth diagram only needs terms to order λ. The question mark is one more counterterm diagram, but more about that one in the final section of this chapter.

3.3.2 Dimensional regularization

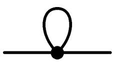
It turns out there is a better way to regularize divergences, which will be particular effective when we consider gauge theories. This is called dimensional regularization. In the case of $0+1$ scalar field theory we found that the one-loop correction is finite, while for $3+1$ it is divergent. Let us now suppose that the number of space-time dimensions D can be varied. In fact, let us suppose that it can be varied continuously. This may seem crazy, because what does it mean to have a noninteger dimension? But really, we are only assuming a continuous dimension when it comes to doing the loop integrals. The D dimensional generalization of the Euclidean one-loop diagram is then



$$= \frac{-i\,\lambda\,\mu^{4-D}}{2}\int\frac{d^D\ell}{(2\pi)^D}\,\frac{1}{\ell^2+m^2}\,. \tag{3.3.16}$$

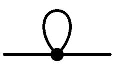
We have replaced λ with $\lambda\,\mu^{4-D}$ where μ has dimensions of mass so that λ stays dimensionless in any dimension. The parameter μ is unphysical and at the end of the day should drop out of all physical quantities.

We can do the integral in (3.3.16) using the following trick:



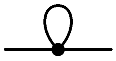
$$\begin{aligned} &= \frac{-i\,\lambda\,\mu^{4-D}}{2}\int_0^\infty d\rho\int\frac{d^D\ell}{(2\pi)^D}\,e^{-\rho(\ell^2+m^2)} = \frac{-i\,\lambda\,\mu^{4-D}}{2(4\pi)^{D/2}}\int_0^\infty d\rho\,\rho^{-D/2}e^{-\rho\, m^2}\\ &= \frac{-i\lambda\;\mu^{4-D}\,m^{D-2}}{2(4\pi)^{D/2}}\int_0^\infty dz\,z^{-D/2}e^{-z}\,. \end{aligned} \tag{3.3.17}$$

The integral is clearly finite if $D<2$. It is logarithmically divergent if $D=2$ and power law divergent for $D>2$. However, the integral is the Γ-function $\Gamma(1-D/2)$ if $D<2$ and this can be analytically continued to values of $D>2$. Hence, we will set



$$= \frac{-i\lambda\,\mu^{4-D}\,m^{D-2}\,\Gamma(1-D/2)}{2(4\pi)^{D/2}}\,. \tag{3.3.18}$$

The Γ-function has poles at 0 and the negative integers and so our analytically continued result is divergent if D is a positive even integer. Therefore the 3+1 result is still divergent and needs to be regularized. The regularization procedure we will use is dimensional regularization where we assume that $D = 4 - 2\varepsilon$ with[7] $\varepsilon << 1$. In this case we see that



$$= \frac{i\lambda m^2}{2(4\pi)^2}\left(\frac{1}{\varepsilon} + 1 - \gamma_E + \log(4\pi) - \log\frac{m^2}{\mu^2}\right) + \mathrm{O}(\varepsilon)\,, \qquad (3.3.19)$$

where $\gamma_E \approx 0.5772$ is the Euler-Mascheroni constant. In this expansion we used that

$$\mu^{2\varepsilon}\, m^{2-2\varepsilon} = m^2\, e^{\varepsilon\,\log\mu^2 - \varepsilon\,\log m^2} = m^2\left(1 - \varepsilon\log\frac{m^2}{\mu^2}\right) + \mathrm{O}(\varepsilon^2)\,. \qquad (3.3.20)$$

To see what the divergent ε^{-1} means, let us consider a different integral, namely

$$\int \frac{d^4\ell}{(2\pi)^4}\,\frac{1}{(\ell^2+m^2)^2}\,, \qquad (3.3.21)$$

which is still divergent, but less so than the one in (3.3.4). If we do the integral using the cutoff Λ then we get

$$\int \frac{d^4\ell}{(2\pi)^4}\,\frac{1}{(\ell^2+m^2)^2} = \frac{1}{16\pi^2}\left(\log\frac{\Lambda^2+m^2}{m^2} + 1 - \frac{m^2}{\Lambda^2+m^2}\right) \approx \frac{1}{16\pi^2}\log\frac{\Lambda^2}{m^2}\,, \qquad (3.3.22)$$

Using dimensional regularization with the scale factor μ^{4-D} included to keep the dimension of the integral fixed, we find

$$\begin{aligned}\mu^{4-D}\int \frac{d^D\ell}{(2\pi)^D}\,\frac{1}{(\ell^2+m^2)^2} &= \mu^{4-D}\int_0^\infty \rho\, d\rho \int \frac{d^D\ell}{(2\pi)^D}\, e^{-\rho(\ell^2+m^2)} = \frac{\mu^{4-D}m^{D-4}}{(4\pi)^{D/2}}\Gamma(2-D/2)\\ &= \frac{1}{16\pi^2}\left(\frac{1}{\varepsilon} - \gamma_E + \log(4\pi) + \log\frac{\mu^2}{m^2}\right) + \mathrm{O}(\varepsilon)\,. \qquad (3.3.23)\end{aligned}$$

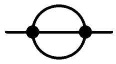
Comparing terms we see that $\varepsilon^{-1} \approx \log\frac{\Lambda^2}{\mu^2}$. Using this in (3.3.19), we discover that dimensional regularization gives us the second term in (3.3.9) but helpfully gets rid of the more divergent term.

The two different regularizations give us different answers for the one-loop mass-shift, which begs the question: which is the correct procedure? The answer is that they both are correct. This is because for each result we must put in a counterterm to cancel the shift so that the pole is at the physical mass. Different mass-shift, different counterterm, same physical pole. However, the regularization procedure we choose should not do untoward harm to our theory. Scalar field theory is rather resilient so either procedure is fine. Later on in the course when we discuss QED we will see that the regularization procedure must preserve gauge invariance in order to avoid unphysical states. Dimensional regularization does an admirable job in this regard.

[7]We will use the script epsilon symbol ε when varying dimensions as opposed to the ordinary epsilon ϵ which appears in the $i\,\epsilon$ of the propagator.

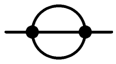
3.3.3 The sunset diagram and the interpretation of its cuts

We next turn to the two-loop self-energy coming from the sunset diagram

$k \rightarrow$  $$= \frac{(-i\lambda)^2}{6}\int \frac{d^4\ell_1}{(2\pi)^4}\frac{d^4\ell_2}{(2\pi)^4}\frac{i}{\ell_1^2-m^2+i\,\epsilon}\,\frac{i}{\ell_2^2-m^2+i\,\epsilon}\,\frac{i}{(k-\ell_1-\ell_2)^2-m^2+i\,\epsilon}\,. \tag{3.3.24}$$

The more loops a diagram has the more complicated the calculation, and unfortunately, two loops is already horrific. In any case let's do it anyway.

One can Wick rotate the integrations over ℓ_1^0 and ℓ_2^0 if we also assume that k^2 is not too big (we will see what too big means later). Afterward, we can analytically continue to bigger values. We will replace k with k_E where $k_E^2 = -k^2$ in the integration, and substitute back at the end of the calculation. After Wick rotating we then have[8]

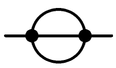
$k \rightarrow$  $$= \frac{i\,\lambda^2}{6}\int \frac{d^4\ell_1}{(2\pi)^4}\frac{d^4\ell_2}{(2\pi)^4}\frac{1}{\ell_1^2+m^2}\,\frac{1}{\ell_2^2+m^2}\,\frac{1}{(k_E-\ell_1-\ell_2)^2+m^2}\,. \tag{3.3.25}$$

Simple power counting shows that this diagram is UV divergent since the measure factors have dimension 8 while the integrand has dimension -6 and thus the integral blows up as ℓ_1 and ℓ_2 get large.

To do the integrals in (3.3.25) we make use of the following identity:

$$\begin{aligned}
\frac{1}{C_1C_2\ldots C_n} &= \int_0^\infty d\rho_1\,d\rho_2\ldots d\rho_n\,e^{-\sum C_i\rho_i}\\
&= \int_0^\infty \rho^{n-1}d\rho\int_0^1 dx_1\,dx_2\ldots dx_n\,\delta\left(1-\sum x_i\right)\,e^{-\rho\sum C_ix_i} \qquad (3.3.26)\\
&= \int_0^1 dx_1\,dx_2\ldots dx_n\,\delta\left(1-\sum x_i\right)\frac{(n-1)!}{(C_1x_1+C_2x_2+\ldots C_nx_n)^n}\,,
\end{aligned}$$

where along the way we substituted $\rho=\sum\rho_i$ and $\rho_i = x_i\,\rho$. The x_i are called *Feynman parameters*. Using dimensional regularization, we can then write the integral in (3.3.25) as

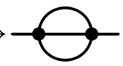
$k \rightarrow$  $$\begin{aligned}= \frac{i\,\lambda^2\mu^{8-2D}}{6}\int\frac{d^D\ell_1}{(2\pi)^D}\frac{d^D\ell_2}{(2\pi)^D}\int_0^\infty\rho^2\,d\rho\int_0^1 dx_1\,dx_2\,dx_3\,\delta(1-x_1-x_2-x_3)\\
\times\,\exp\left(-\rho\left(x_1\ell_1^2+x_2\ell_2^2+x_3(k_E-\ell_1-\ell_2)^2+m^2\right)\right)\,.\end{aligned} \tag{3.3.27}$$

Completing the square for ℓ_1,

$$x_1\ell_1^2+x_3(k_E-\ell_1-\ell_2)^2 = (x_1+x_3)\left(\ell_1+\frac{x_3}{x_1+x_3}(\ell_2-k_E)\right)^2-\frac{x_3^2}{x_1+x_3}(\ell_2-k_E)^2\,, \tag{3.3.28}$$

[8]We have dropped the $i\epsilon$ terms. At the end of this section we will need them back, but this can be achieved by replacing m^2 with $m^2-i\,\epsilon$.

and then doing the Gaussian integral over ℓ_1 gives

$k \rightarrow$ 

$$= \frac{i\,\lambda^2\mu^{8-2D}}{6(4\pi)^{D/2}}\int\frac{d^D\ell_2}{(2\pi)^D}\int_0^\infty \rho^{2-D/2}\,d\rho\int_0^1\frac{dx_1\,dx_2\,dx_3}{(x_1+x_3)^{D/2}}\,\delta(1-x_1-x_2-x_3)$$
$$\times\,\exp\left(-\rho\left(x_2\ell_2^2+\left(x_3-\frac{(x_3)^2}{x_1+x_3}\right)(k_E-\ell_2)^2+m^2\right)\right)\,. \tag{3.3.29}$$

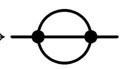
Completing the square for ℓ_2,

$$x_2\ell_2^2+\left(x_3-\frac{(x_3)^2}{x_1+x_3}\right)(k_E-\ell_2)^2 \;=\; \frac{x_1x_2+x_2x_3+x_3x_1}{x_1+x_3}\left(\ell_2-\frac{x_1x_3}{x_1x_2+x_2x_3+x_3x_1}k_E\right)^2+X\,k_E^2\,, \tag{3.3.30}$$

with X defined as

$$X\equiv\frac{x_1x_2x_3}{x_1x_2+x_2x_3+x_3x_1}\,, \tag{3.3.31}$$

and then doing the Gaussian integral over ℓ_2 followed by the integral over ρ leads to

$k \rightarrow$ 

$$= \frac{i\,\lambda^2\mu^{8-2D}}{6(4\pi)^{D}}\int_0^\infty \rho^{2-D}\,d\rho\int_0^1\frac{dx_1\,dx_2\,dx_3}{(x_1x_2+x_2x_3+x_3x_1)^{D/2}}\,\delta(1-x_1-x_2-x_3)\,\exp\left(-\rho\left(X\,k_E^2+m^2\right)\right)$$
$$= \frac{i\,\lambda^2\mu^{8-2D}}{6(4\pi)^{D}}\,\Gamma(3-D)\int_0^1\frac{dx_1\,dx_2\,dx_3}{(x_1x_2+x_2x_3+x_3x_1)^{D/2}}\,\delta(1-x_1-x_2-x_3)\,\left(m^2-X\,k^2\right)^{D-3}\,, \tag{3.3.32}$$

where in the last step we replaced k_E^2 with $-k^2$. Setting $D=4-2\varepsilon$, we see that there is a singularity from the Γ-function, $\Gamma(3-D)\approx-\frac{1}{2\varepsilon}$. Let us now use the expansion

$$\left(m^2-X\,k^2\right)^{1-2\varepsilon}\approx m^{-4\varepsilon}\left[1-2\,\varepsilon\,\log\left(1-X\,\frac{k^2}{m^2}\right)\right]\left(m^2-X\,k^2\right) \tag{3.3.33}$$

If we take the first term in the square brackets we find that there is a singularity from the integrals over the Feynman parameters if $m^2\neq0$. This arises at the three corners of the integration region where two of the Feynman parameters approach 0. In this case $X\to0$ so the k^2 term does not contribute to this singularity. The Feynman parameter integrals are then approximately

$$\int_0^1\frac{dx_1\,dx_2\,dx_3}{(x_1x_2+x_2x_3+x_3x_1)^{2-\varepsilon}}\,\delta(1-x_1-x_2-x_3)\approx3\int\frac{dx_1\,dx_2}{(x_1+x_2)^{2-\varepsilon}}\approx\frac{3}{\varepsilon}\,, \tag{3.3.34}$$

which leads to a Δm^2 term with a double pole in ε,

$$\Delta m^2=\frac{\lambda^2}{4(4\pi)^4\,\varepsilon^2}m^2+\mathrm{O}(\varepsilon^{-1})=\frac{\lambda^2}{4(4\pi)^4}m^2\left(\log\frac{\Lambda^2}{\mu^2}\right)^2+\mathrm{O}\left(\log\frac{\Lambda^2}{\mu^2}\right)\,. \tag{3.3.35}$$

The leading singular term is then canceled with the counterterm

$$\delta_{m^2} = -\frac{\lambda^2}{4(4\pi)^4} m^2 \left(\log\frac{\Lambda^2}{\mu^2}\right)^2 . \tag{3.3.36}$$

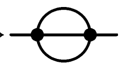
The Feynman integral for the k^2 term is finite, leaving an overall single pole in ε. If we take the first term in the square brackets in (3.3.33) we find that the k^2 term is (the actual evaluation is left as an exercise)

$$\begin{aligned} &-\frac{i\,\lambda^2\mu^{4\varepsilon}m^{-4\varepsilon}\Gamma[-1+2\varepsilon]}{12(4\pi)^{4-2\varepsilon}}(1+\varepsilon\, C)\,k^2 = \\ &\qquad = \frac{i\,\lambda^2}{12(4\pi)^4}\left(\frac{1}{2\varepsilon}+1-\gamma_E+\log(4\pi)-\frac{C}{2}+\log\frac{\mu^2}{m^2}\right)k^2 + \mathrm{O}(\varepsilon)\,, \end{aligned} \tag{3.3.37}$$

where C is the finite integral

$$C = \int_0^1 \frac{dx_1\,dx_2\,dx_3\,(x_1x_2x_3)}{(x_1x_2+x_2x_3+x_3x_1)^3}\,\log(x_1x_2+x_2x_3+x_3x_1)\,\delta(1-x_1-x_2-x_3)\,. \tag{3.3.38}$$

If we now take the second term in the square brackets in (3.3.33) we have the additional finite term

$k \rightarrow$  $\Big|_{\mathrm{finite}}$

$$= \frac{i\,\lambda^2}{6(4\pi)^4}\,m^2 f\left(\frac{k^2}{m^2}\right)\,, \tag{3.3.39}$$

where

$$f(z) \;=\; \int_0^1 \frac{dx_1dx_2dx_3\,\delta(1-x_1-x_2-x_3)}{(x_1x_2+x_2x_3+x_3x_1)^2}\,(1-X\,z)\log\left(1-X\,z\right)\,.$$

Putting together the terms from (3.3.37) and (3.3.39), we find that the wave-function counterterm δ_Z is

$$\begin{aligned} \delta_Z &= \frac{\partial}{\partial k^2}\Sigma_0(k^2)\Big|_{k^2=m^2} \\ &= -\frac{\lambda^2}{12(4\pi)^4}\left(\frac{1}{2\varepsilon}+1-\gamma_E+\log(4\pi)-\frac{C}{2}+\log\frac{\mu^2}{m^2}+2f'(1)\right)\,, \end{aligned} \tag{3.3.40}$$

where $f'(z) = \frac{d}{dz}f(z)$.

After adding all counterterms, the physical propagator becomes

$$G^{(2)}(k^2) = \frac{i}{k^2-m^2-\Sigma(k^2)+i\,\epsilon}\,, \tag{3.3.41}$$

where the physical self-energy $\Sigma(k^2)$ is

$$\Sigma(k^2) = \frac{\lambda^2}{6(4\pi)^4}\left(m^2f(1)+(k^2-m^2)f'(1)-m^2f\left(\frac{k^2}{m^2}\right)\right)\,. \tag{3.3.42}$$

This is the physical propagator to two loop order and is independent of μ, which we previously claimed was unphysical. Notice that $\Sigma(m^2) = \frac{\partial}{\partial k^2}\Sigma(k^2)\Big|_{k^2=m^2} = 0$ so that the physical pole is at m^2 with residue 1.

We will not explicitly evaluate $f(z)$ in $\Sigma(k^2)$, but it has a very interesting property that tells us something important about the physics. This function is real for $z < 9$, but develops an imaginary part when $z > 9$. This is because the argument of the log is negative for some of the integration region if $z > 9$. Assuming that we choose the branch of the log so that it has no imaginary part for $z < 9$, then if we assume that z has a small positive (negative) imaginary part, then the imaginary part of the log is $-i\pi$ ($+i\pi$) when the real part of the argument is less than zero. The maximum value of X in the integration region is $1/9$ which is reached when $x_1 = x_2 = x_3 = 1/3$. Hence the argument of the log will have a negative real part for some of the integration region if $\mathrm{Re}(z) > 9$. Therefore, assuming that z is real we find that

$$\mathrm{Im}(f(z\pm i\epsilon)) = \pm\pi\int_0^1 \frac{dx_1dx_2dx_3\,\delta(1-x_1-x_2-x_3)}{(x_1x_2+x_2x_3+x_3x_1)^2}\,(X\,z-1)\,\theta(X\,z-1)\,. \tag{3.3.43}$$

Thus, $f(z)$ is real for z real and $z \le 9$, but has a branch cut along the real axis for $z > 9$ where the imaginary part is discontinuous. Which side of the cut z is on is determined by the $i\,\epsilon$ term that we had previously dropped after the Wick rotation, but which can be restored by replacing m^2 with $m^2 - i\,\epsilon$. Under this replacement $z \to z + i\,\epsilon$ for $k^2 > 0$, so z is on the topside of the branch cut.

To understand this branch cut, let us write the propagator as

$$G^{(2)}(k^2) \approx \frac{i}{k^2-m^2+i\,\epsilon} + \frac{i\,\Sigma(k^2)}{(k^2-m^2)^2}\,, \tag{3.3.44}$$

where we dropped the $i\epsilon$ in the second term because it does not have a pole at $k^2 = m^2$, since $\Sigma(m^2) = \frac{\partial}{\partial k^2}\Sigma(k^2)|_{k^2=m^2} = 0$. We can then write

$$\frac{\Sigma(k^2)}{(k^2-m^2)^2} = m^{-2}F(z)\,, \tag{3.3.45}$$

where

$$F(z) = \frac{\lambda^2}{6(4\pi)^4}\,\frac{f(1)+(z-1)f'(1)-f(z)}{(z-1)^2}\,. \tag{3.3.46}$$

$F(z)$ is analytic everywhere in z except along the branch cut and behaves as $F(z) \sim \frac{\log(z)}{z}$ as $z \to \infty$. Therefore, by Cauchy's theorem we have for general complex z that

$$F(z) = -\frac{1}{2\pi i}\oint_C \frac{dz'}{z-z'}F(z')\,, \tag{3.3.47}$$

where C is the contour shown in figure 2. This formula is valid so long as z is not on the cut. Since $z^{-1}\,F(z) \sim z^{-2}\log(z)$ as $z \to \infty$, there is no contribution to the contour

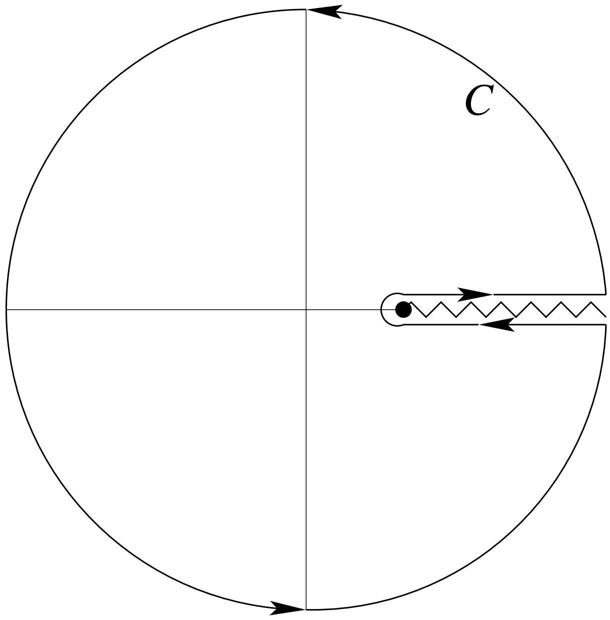


Figure 3.2: Contour C for integration in (3.3.47). The jagged line is the branch cut.

integral from the the circle at infinity and so everything comes from the integral around the branch cut. Hence we have

$$F(z) = -\frac{1}{2\pi i}\int_9^\infty \frac{dz'}{z-z'}\left(F(z'+i\epsilon)-F(z'-i\epsilon)\right) = \int_9^\infty \frac{\chi(z')\,dz'}{z-z'}\,, \tag{3.3.48}$$

where

$$\chi(z) = \frac{\lambda^2}{6(4\pi)^4}\frac{1}{(z-1)^2}\int_0^1 \frac{dx_1dx_2dx_3\,\delta(1-x_1-x_2-x_3)}{(x_1x_2+x_2x_3+x_3x_1)^2}\,(X\,z-1)\,\theta(X\,z-1)\,. \tag{3.3.49}$$

$\chi(z)=0$ if $z\leq 9$ and is strictly positive for $z>9$, falling off as $\chi(z)\sim z^{-1}$ for large z. We previously argued that z has a small positive imaginary piece, hence we finally reach that

$$\frac{\Sigma(k^2)}{(k^2-m^2)^2} = \int_9^\infty \frac{\chi(z')\,dz'}{k^2-z'm^2+i\epsilon}\,. \tag{3.3.50}$$

Let us now show that (3.3.50) is the form we expect. In the case of the anharmonic oscillator we saw that there was a pole at $\omega^2=(3m_0)^2$ because $\phi(0)$ had some probability of producing a three particle state. In $0+1$ dimensions there is only one three-particle state, but in $3+1$ dimensions there are infinitely many. The 4-momentum of the three particles will satisfy $k^2=(k_1+k_2+k_3)^2\equiv M^2$, which has a minimum allowed value of $M^2=(3m)^2$ when $\vec{k}_1=\vec{k}_2=\vec{k}_3=0$. Above this value there is a continuum of states. Hence we expect that $G^{(2)}(k^2)$ should include an integration over a continuous distribution of poles. Since the poles are determined by M^2 this can be expressed through a "spectral density function", $\rho(M^2) \geq 0$, such that

$$G^{(2)}(k^2) = \frac{i}{k^2 - m^2 + i\,\epsilon} + \int_{M^2=(3m)^2}^{\infty} \frac{dM^2}{2\pi} \rho(M^2) \frac{i}{k^2 - M^2 + i\,\epsilon} \tag{3.3.51}$$

where the 2π is a convention. We can see that the second term in $G^{(2)}(k^2)$ has precisely the form in (3.3.50) if we make the identification

$$\rho(M^2) = \frac{2\pi}{m^2}\,\chi\left(\frac{M^2}{m^2}\right)\,. \tag{3.3.52}$$

3.4 4-point correlators and the effective coupling

We now consider 4-point correlators, up to the one-loop level. It is convenient at this point to write the path integral as $\mathcal{Z}(J) = e^{iW(J)}$. The normalized 2-point correlator is then given by

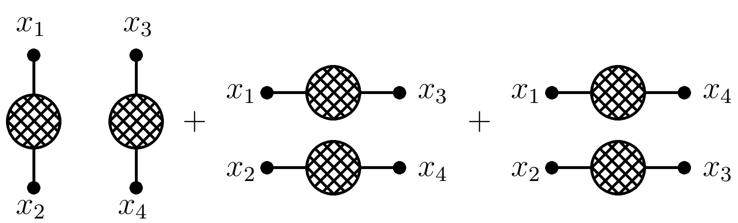
$$\frac{\langle T[\phi(x_1)\phi(x_2)]\rangle}{\langle 0|0\rangle} = \mathcal{Z}^{-1}(0)\frac{-i\delta}{\delta J(x_1)}\frac{-i\delta}{\delta J(x_2)}\mathcal{Z}(J)\Big|_{J=0} = \frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}W(J)\Big|_{J=0}\,, \tag{3.4.1}$$

where we used that $\frac{\delta}{\delta J(x)}W(J)\Big|_{J=0} = 0$. The latter condition is the statement that the one-point functions are zero, $\langle\phi(x)\rangle = 0$, which is the case for any interacting theory which is invariant under $\phi(x) \to -\phi(x)$.

The 4-point correlator in this language is

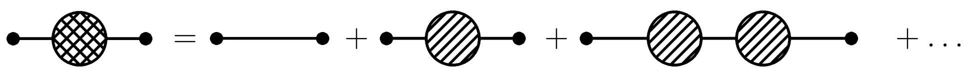
$$\begin{aligned}\frac{\langle T[\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)]\rangle}{\langle 0|0\rangle} \quad = \quad & \frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}W(J)\Big|_{J=0}\frac{\delta}{\delta J(x_3)}\frac{\delta}{\delta J(x_4)}W(J)\Big|_{J=0} \\ & +\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_3)}W(J)\Big|_{J=0}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_4)}W(J)\Big|_{J=0} \\ & +\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_4)}W(J)\Big|_{J=0}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_3)}W(J)\Big|_{J=0} \\ & +\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_3)}\frac{\delta}{\delta J(x_4)}W(J)\Big|_{J=0}\,. \end{aligned} \tag{3.4.2}$$

Comparing with (3.4.1) we see that the first three terms in (3.4.2) are products of 2-point functions. These then are the disconnected diagrams



(3.4.3)

where

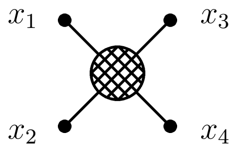


(3.4.4)

Since (3.4.3) are all possible disconnected diagrams (without vacuum bubbles which have been divided out when normalizing the propagator), the fourth term in (3.4.2) must be all *connected* diagrams. This argument is easily extendible to all higher point correlators, hence, $W(J)$ is the generator of all connected diagrams[9].

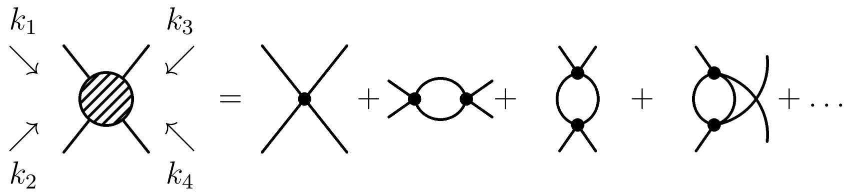
$$\frac{\langle T[\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)]\rangle_{\text{connected}}}{\langle 0|0\rangle} = \frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_3)}\frac{\delta}{\delta J(x_4)}W(J)\Big|_{J=0}$$

$$=$$



(3.4.5)

The truncated diagram in momentum space where we lop off the propagators connected to the end points is



(3.4.6)

The first term in (3.4.6) is the single vertex and is equal to the vertex Feynman rule in (3.3.2). Before evaluating the next three one-loop graphs it is convenient to introduce some new terminology. The Mandelstam variables are defined as

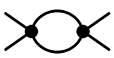
$$\begin{aligned} s &\equiv (k_1+k_2)^2 = (k_3+k_4)^2 \\ t &\equiv (k_1+k_3)^2 = (k_2+k_4)^2 \\ u &\equiv (k_1+k_4)^2 = (k_2+k_3)^2\,. \end{aligned} \tag{3.4.7}$$

If the incoming lines are on-shell, *i.e* $k_i^2 = m^2$ for all i, then it is a straightforward exercise to show that

$$s+t+u = 4\,m^2\,. \tag{3.4.8}$$

In the first loop graph we have $k_1^\mu + k_2^\mu$ entering the vertex, so for this reason this is called the s-channel loop graph. The successive graphs are the t- and u-channel graphs respectively. By Lorentz invariance the s-channel graph can only depend on s, the t-channel only on t and the u-channel only on u. By symmetry, we also see that after computing the s-channel graph we can immediately find the t- and u-channel contributions by substituting t and u for s respectively.

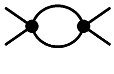
Using the Feynman rules we find that



$$\begin{aligned} &= (2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\frac{1}{2}(-i\lambda)^2\int\frac{d^4\ell}{(2\pi)^4}\frac{i}{\ell^2-m^2+i\epsilon}\frac{i}{(k_1+k_2-\ell)^2-m^2+i\epsilon} \\ &= i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\frac{1}{2}\lambda^2\int\frac{d^4\ell}{(2\pi)^4}\frac{1}{\ell^2+m^2}\frac{1}{(k_{1E}+k_{2E}-\ell)^2+m^2} \end{aligned} \tag{3.4.9}$$

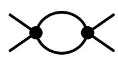
[9] A combinatorial argument for this can be found in Srednicki.

where the subscript E again refers to the Euclidean values. We again dropped the $i\epsilon$ which we can bring back in by shifting m^2 by $-i\epsilon$. We next use dimensional regularization and Feynman parameters to get



$$
\begin{aligned}
&= i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\frac{1}{2}\lambda^2\mu^{4-D}\int_0^\infty \rho\, d\rho\int_0^1 dx\int\frac{d^D\ell}{(2\pi)^D}\\
&\qquad\times\exp\left(-\rho\left(x\ell^2+(1-x)(k_{1E}+k_{2E}-\ell)^2+m^2\right)\right)\\
&= i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\frac{\lambda^2\,\mu^{4-D}}{2(4\pi)^{D/2}}\Gamma(2-D/2)\int_0^1\frac{dx}{\left(m^2-x(1-x)s\right)^{2-D/2}}\,,
\end{aligned}
\tag{3.4.10}
$$

where in doing the gaussian integral over ℓ we completed the square and set $(k_{1E}+k_{2E})^2=-s$. If we now let $D=4-2\,\varepsilon$, then up to finite terms we get



$$
= i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\frac{\lambda^2}{32\,\pi^2}\left(\frac{1}{\epsilon}-\gamma_E+\log(4\pi)+\log\frac{\mu^2}{m^2}+Q\left(\frac{s}{m^2}\right)\right)\,,
\tag{3.4.11}
$$

where

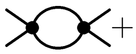
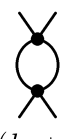
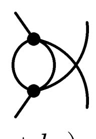
$$
Q(z)=-\int_0^1 dx\,\log\left(1-x(1-x)z\right)=2\left(1-\sqrt{\frac{4}{z}-1}\,\arctan\left(\frac{1}{\sqrt{\frac{4}{z}-1}}\right)\right).
\tag{3.4.12}
$$

If $z>4$ then the argument of the log is negative for the integration region $\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{4}{z}}<x<\frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{4}{z}}$. It then follows that

$$
\mathrm{Im}(Q(z\pm i\epsilon))=\pm\pi\sqrt{1-\frac{4}{z}}\,.
\tag{3.4.13}
$$

Hence $Q(z)$ has a branch cut for $z>4$. We will explain the consequences of the branch cut later in the course when we discuss particle scattering and the optical theorem.

Including all three channels we find that the one-loop contribution to the connected 4-point function is

 +  + 

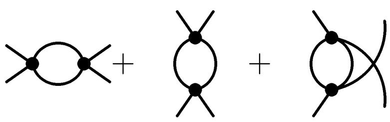
$$
\begin{aligned}
&= i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\\
&\times\frac{3\,\lambda^2}{32\,\pi^2}\left(\frac{1}{\epsilon}-\gamma_E+\log(4\pi)+\log\frac{\mu^2}{m^2}+\frac{1}{3}\left(Q\left(\frac{s}{m^2}\right)+Q\left(\frac{t}{m^2}\right)+Q\left(\frac{u}{m^2}\right)\right)\right)\,.
\end{aligned}
\tag{3.4.14}
$$

We need to add an infinite counterterm to cancel off the divergence, such that remaining piece is the physical coupling. However, there is a new wrinkle that we did not encounter with δ_Z and δ_{m^2}. In these cases we have a pole at a physical mass and we adjust the counterterms to put us at that pole with the correct residue. We can do the same here, where we add the counterterm to give us the physical coupling. The thing is, we have to decide at which values of s, t and u we will fix the coupling, because ultimately the counterterm is local which prevents it from having more than polynomial dependence on the momenta.

To make this argument more transparent, let us suppose that the values of s, t and u are so large that we can ignore the mass. In this case the arguments of the Q-functions have a modulus which is much greater than 1. Hence, the functions can be approximated as

$$Q(z) \approx -\int_0^1 dx\, \log\left(-x(1-x)z\right) = -\log(-z) + 2\,, \tag{3.4.15}$$

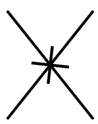
and the diagrams sum to



$$= i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\frac{3\,\lambda^2}{32\,\pi^2}\left(\frac{1}{\epsilon}-\gamma_E+\log(4\pi)+2-\frac{1}{3}\log\left(-\frac{stu}{\mu^6}\right)\right) \tag{3.4.16}$$

To cancel off the divergence we need to add a counterterm δ_λ to the coupling that appears in the Lagrangian,

$$\delta_\lambda = $$

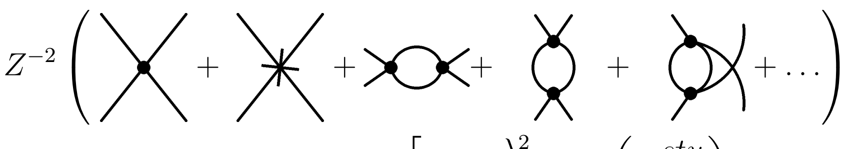


$$= -i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\frac{3\,\lambda^2}{32\,\pi^2}\left(\frac{1}{\epsilon}-\gamma_E+\log(4\pi)+2\right) \tag{3.4.17}$$

where we have also chosen to cancel off the annoying γ_E and $\log(4\pi)$ terms, as well as the 2 for good measure. Hence the "bare coupling", λ_0, the coupling that appears in the Lagrangian, is

$$\lambda_0 = \lambda + \frac{3\,\lambda^2}{32\,\pi^2}\left(\frac{1}{\epsilon}-\gamma_E+\log(4\pi)+2\right) + \mathrm{O}(\lambda^3)\,. \tag{3.4.18}$$

The physical truncated and connected 4-point diagram is then



$$= -i\,(2\pi)^4\delta^4(k_1+k_2+k_3+k_4)\left[\lambda + \frac{\lambda^2}{32\,\pi^2}\log\left(-\frac{stu}{\mu^6}\right) + \mathrm{O}(\lambda^3)\right]\,. \tag{3.4.19}$$

where we have multiplied by $(Z^{-1/2})^4$ so that we have the truncated diagram for the *renormalized* fields $\phi_R(x) = Z^{-1/2}\phi_0(x)$, where $\phi_0(x)$ are the bare fields that appear in the Lagrangian. However, we have already seen that Z gets no corrections until two-loop order[10], so at the one-loop order we have $Z = 1$.

We would think that the real part inside the square brackets of equation (3.4.19) should give the physical coupling at the one-loop order. But there seems to be a problem because of the μ appearing in a supposedly physical result. If we try to absorb the $\log \mu^2$ term into the bare coupling we would need to introduce another mass scale so that we could make a log with a dimensionless argument. We cannot include the $\log stu$ because this would lead to a nonlocal counterterm. There is only one possibility; the physical coupling is scale dependent. Here is how it would work: Our procedure should be independent of μ, so suppose we choose $\mu = M$ where $(stu)^2 \equiv (M)^{12}$. Furthermore, we let $\lambda = \bar{\lambda}$ be the physical coupling when $M = \mu$. Hence for M near μ we have that

$$\lambda(M^2) \approx \bar{\lambda} + \frac{3\bar{\lambda}^2}{32\pi^2} \log \frac{M^2}{\mu^2} \,. \tag{3.4.20}$$

If we now take a derivative with respect to $\log M^2$, we find that

$$\frac{\partial \lambda}{\partial \log M^2} \approx \frac{3\bar{\lambda}^2}{32\pi^2} \approx \frac{3\lambda^2}{32\pi^2} \,. \tag{3.4.21}$$

We now claim that actually to this order in λ the true equation is

$$\beta(\lambda) \equiv \frac{\partial \lambda}{\partial \log M^2} = \frac{3\lambda^2}{32\pi^2} \,, \tag{3.4.22}$$

where $\beta(\lambda)$ is called the β-function for λ. The righthand side depends on λ and not $\bar{\lambda}$ because after changing M by a little bit, if the physical result is independent of μ, then we are free to choose it at the new value of M with the new coupling. We then do the procedure all over again. If the right hand side had depended on $\bar{\lambda}$ then the β-function, and hence the physical coupling, would have depended on μ through a choice of a $\bar{\lambda}$. Let me emphasize that the physics is independent of μ because choosing a different μ compensates by choosing a different value for the coupling.

Since the sign of the β-function in (3.4.22) is positive, the coupling is increasing as M increases. This means that scalar field theory becomes more strongly coupled as we go to higher energies. This change in the coupling as the scale changes is called the running of the coupling constant. We can actually solve for the equation in (3.4.22), where we find

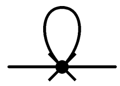
$$\lambda(M^2) = -\frac{32\pi^2}{3 \log \dfrac{M^2}{M_0^2}} \,, \tag{3.4.23}$$

where M_0^2 is an integration constant. This gives us a new scale in the theory which is physical and is defined through the coupling. For example we can define the intrinsic scale as the value of M where the coupling is 1. This result diverges at $M = M_0$ and so perturbation theory would have broken down before this value, so (3.4.23) is not particulary trustworthy near this value. But it is probably reliable if $\lambda < 1$. This result is also only valid for $M >> m$. Once M gets near m the running slows down.

[10]Wave-function renormalization will need to be taken into account at the one-loop order when we consider QED.

3.5 A few comments

1. The missing diagram in (3.3.15) contains the counterterm for the vertex

   ? =  (3.5.1)

2. You might be worried about doing perturbation theory with infinite terms at each order of perturbation theory. However, the philosophy you should take is that each order in λ includes the loop graphs *and* the counterterms, so combined together everything is finite and perturbation theory is well behaved, at least if the couplings are small enough.

3. It turns out that δ_Z, δ_m^2 and δ_λ are the only counterterms you need to cancel divergences in ϕ^4 theory. This is a hallmark of a renormalizable field theory. Nonrenormalizable theories require an infinite number of counterterms. Non renormalizable theories have couplings whose dimension is negative (*e.g.* $\lambda\phi^6$ in 3+1 dimensions).

4. Throughout these notes we have been employing a particular regularization "scheme" called the on-shell scheme, where the counterterms are chosen so that the poles of the propagators are at the physical values of the masses. There are other schemes that are often used. One is the minimal subtraction (MS) scheme, where one only adds counterterms to cancel off the ε^{-1} term. Another scheme is the modified minimal subtraction ($\overline{\text{MS}}$) where one also cancels off the factors of γ_E and $\log(4\pi)$. In both of these schemes there is no dependence on μ^2 in the counterterms, and hence in the bare parameters. The resulting m_R^2 must then depend on μ^2 and is referred to as the *renormalized* mass squared and is distinct from the physical value where the pole sits.

Chapter 4

Free Dirac fields

4.1 The Dirac field

So far we have discussed the scalar field, whose quantization leads to spinless particles. But many of the elementary particles in nature are spin 1/2 particles, including electrons, muons, taus, neutrinos and quarks so we need to find the corresponding local fields for these particles.

4.1.1 The Dirac equation

I am sure that many of you are familiar with the story of how Dirac derived his famous equation. After Schrödinger introduced his eponymous equation, many started looking for a relativistic generalization. The obvious thing to try was making it second order in the time derivatives to match the second order spatial derivatives. So in the case of a free particle it would have the form

$$-\left(\frac{\partial^2}{\partial^2 t} - \vec{\nabla}^2\right)\Psi(\vec{x},t) = m^2\,\Psi(\vec{x},t)\,. \tag{4.1.1}$$

The problems with this equation, at least as a generalization of the Schrödinger equation are well known. In particular there are arbitrarily low negative energy solutions. Even worse, there is no well behaved probability current, since the natural choice for the probability density $\Psi^*\partial^0\Psi$ is not positive definite. In any case, we recognize the equation in (1.3.1) as the Klein-Gordon equation and we now know that we should treat it not as the Schrödinger equation of relativistic quantum mechanics, but as a classical field equation whose field we then quantize[1].

The root of these problems for the Klein-Gordon equation is the second order time derivative. Dirac's big insight was not to make the time derivatives second order, but to make the spatial derivatives first order. However, in order for this to work $\phi(x)$ had to

[1] For this reason the quantization of the scalar field $\phi(x)$ is often called second quantization, the first quantization being the Klein-Gordon equation itself which was initially treated as a relativistic Shrödinger equation.

have a set of indices and a certain set of matrices needed to be introduced. In particular Dirac proposed the equation

$$(i\partial\!\!\!/ - m)\psi(x) = 0\,, \tag{4.1.2}$$

where $\partial\!\!\!/ \equiv \gamma^\mu\partial_\mu$ and the γ^μ are a set of matrices for each space-time component. If we act on this equation with $-i\partial\!\!\!/ - m$ we should get a relativistic second order equation

$$(\partial^2 + m^2)\,\psi(x) = 0 \tag{4.1.3}$$

which can be true only if

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\,\eta^{\mu\nu}\,. \tag{4.1.4}$$

The relations in (4.1.4) are called a Clifford algebra and Dirac found that the matrices must be 4×4 (at least) in order to have a faithful representation of (4.1.4). One choice of basis for the matrices is the Weyl basis (which Peskin uses), with

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\,, \tag{4.1.5}$$

where I is the 2×2 identity matrix and σ^i, $i = 1, 2, 3$ are the three Pauli matrices. Since γ^μ are 4×4, $\psi(x)$ has four components. We will use front-alphabet Greek letters (α, β, *etc.*) for the components, $\psi_\alpha(x)$, and refer to them as Dirac indices.

4.1.2 The Lorentz algebra and its spinor representations

In this section we study how Lorentz transformations act on $\psi(x)$. Lorentz transformations are the space-time generalizations of spatial rotations, the latter generated by the different components of the angular momentum operator. Let's review why this is so. Suppose we have a wave-function that is a function of the spherical coordinates, $\Psi(r, \theta, \phi)$. A rotation in the $x - y$ plane leads to the shift $\phi \to \phi + \Delta\phi$. We can express the wave-function with this shift as the Taylor expansion

$$\begin{aligned}\Psi(r,\theta,\phi+\Delta\phi) &= \Psi(r,\theta,\phi) + \Delta\phi\,\partial_\phi\Psi(r,\theta,\phi) + \tfrac{1}{2}(\Delta\phi)^2\partial_\phi^2\Psi(r,\theta,\phi) + \ldots \\ &= e^{\Delta\phi\,\partial_\phi}\Psi(r,\theta,\phi) = e^{i\Delta\phi\,L_z}\Psi(r,\theta,\phi)\,.\end{aligned} \tag{4.1.6}$$

Hence L_z generates rotations in the $x - y$ plane. The other generators of angular momentum rotate angles in the other planes. To this end, let us define $J^{ij} = \varepsilon^{ijk}L_k$. In terms of the coordinates and its derivatives this becomes

$$J^{ij} = x^i p_j - x^j p_i = -i(x^i\partial_j - x^j\partial_i)\,. \tag{4.1.7}$$

We can then generalize this to the full set of Lorentz generators as

$$J^{\mu\nu} = i\,\left(x^\mu\partial^\nu - x^\nu\partial^\mu\right)\,. \tag{4.1.8}$$

(The extra sign in J^{ij} compared to $J^{\mu\nu}$ comes from lowering the spatial index on the derivative.)

Let us now show how $J^{\mu\nu}$ generates a Lorentz transformation. Under a transformation, the coordiate x^μ transforms as

$$x^\mu \to x^{\mu'} = \Lambda^{\mu'}{}_\nu\, x^\nu\,. \tag{4.1.9}$$

Suppose the Lorentz transformation is a boost in the x^1 direction and let us further assume that the boost is very small, hence we have

$$x^0 \to x^0 - \epsilon\, x^1\,, \qquad x^1 \to x^1 - \epsilon\, x^0\,, \qquad x^{2,3} \to x^{2,3}\,, \tag{4.1.10}$$

where $\epsilon << 1$. It's not hard to show that this transformation can be written as

$$x^\mu \to x^\mu - i\epsilon\,[J^{01}, x^\mu] \approx (1 - i\epsilon\, J^{01})x^\mu(1 + i\epsilon\, J^{01})\,. \tag{4.1.11}$$

If we now consider a finite version of this transformation, we can think of this as N infinitesimal transformations such that $N\epsilon = \eta$. In this case we find

$$x^\mu \to (1 - i\epsilon\, J^{01})^N x^\mu (1 + i\epsilon\, J^{01})^N = e^{-i\eta J^{01}} x^\mu e^{+i\eta J^{01}}\,, \tag{4.1.12}$$

For a more general Lorentz transformation, the infinitesimal version has

$$x^\mu \to x^\mu - \frac{i}{2}\epsilon_{\nu\lambda}[J^{\nu\lambda}, x^\mu] = x^\mu + \epsilon_\nu{}^\mu x^\nu\,, \tag{4.1.13}$$

which exponentiates to

$$x^\mu \to x^{\mu'} = U^{-1}(\Lambda)\, x^\mu\, U(\Lambda)\,, \tag{4.1.14}$$

where $U(\Lambda) = \exp(+\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu})$ and the $\omega_{\mu\nu}$ are a set of real parameters characterizing the transformation (*e.g.* rotation angles).

The Lorentz generators satisfy the algebra

$$[J^{\mu\lambda}, J^{\nu\sigma}] = i\left(\eta^{\lambda\nu}J^{\mu\sigma} - \eta^{\mu\nu}J^{\lambda\sigma} - \eta^{\lambda\sigma}J^{\mu\nu} + \eta^{\mu\sigma}J^{\lambda\nu}\right)\,, \tag{4.1.15}$$

which can be readily verified using the definitions in (4.1.8). In 3+1 dimensions this algebra is called $SO(1,3)$ but it also generalizes to any space-time dimension. The explicit generators in (4.1.8) are a representation of this algebra for acting on coordinates. It turns out we can also construct a representation with the γ^μ matrices, just as we can construct an algebra of the rotations with the Pauli matrices, σ^i. In particular we let

$$S^{\mu\nu} = \frac{i}{4}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)\,. \tag{4.1.16}$$

Using the Clifford algebra in (4.1.4) as well as the identity

$$[AB, CD] = A\{B, C\}D - \{A, C\}BD + CA\{B, D\} - C\{A, D\}B\,, \tag{4.1.17}$$

it is straightforward to show that $S^{\mu\nu}$ satisfies the Lorentz algebra in (4.1.15). One can further show that

$$\frac{i}{2}\epsilon_{\nu\lambda}[S^{\nu\lambda}, \gamma^\mu] = -\epsilon_\nu{}^\mu\gamma^\nu\,, \tag{4.1.18}$$

and so γ^μ transforms the same way as x^μ under Lorentz transformations. Therefore

$$\Lambda^{\mu'}{}_\nu\gamma^\nu = U_\gamma^{-1}(\Lambda)\gamma^\mu U_\gamma(\Lambda) \tag{4.1.19}$$

where $U_\gamma(\Lambda) = \exp(+\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu})$.

We now demand that the Dirac equation is covariant, meaning that if it is satisfied in one Lorentz frame it is satisfied in all frames. Given a set of γ^μ in frame $\mathbf{S}$ one might think that the Dirac equation in $\mathbf{S}'$ is

$$(i\gamma^{\mu'}\partial_{\mu'} - m)\psi'(x') = 0\,, \tag{4.1.20}$$

where $\gamma^{\mu'} = \Lambda^{\mu'}{}_\nu\gamma^\nu$. But this means the observer in $\mathbf{S}'$ uses a different set of γ-matrices than the observer in $\mathbf{S}$. While this is possible to do, it is more natural and less cumbersome that every observer use the same set of γ-matrices. So in reality the equation used by the observer in $\mathbf{S}'$ is

$$U_\gamma(\Lambda)(i\gamma^{\mu'}\partial_{\mu'} - m)U_\gamma^{-1}(\Lambda)\psi'(x') = 0\,, \tag{4.1.21}$$

which transforms $\gamma^{\mu'}$ back to γ^μ. Since $i\gamma^{\mu'}\partial_{\mu'} = i\gamma^\mu\partial_\mu$, we must have that $\psi(x)$ transforms as

$$\psi(x) \to \psi'(x') = U_\gamma(\Lambda)\psi(x) = \exp(+\tfrac{i}{2}\omega_{\mu\nu}S^{\mu\nu})\psi(x)\,, \tag{4.1.22}$$

in order to satisfy the Dirac equation in both frames.

Writing the $S^{\mu\nu}$ in the Weyl basis, we find that

$$S^{0i} = -\frac{i}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix} \qquad S^{ij} = \frac{1}{2}\,\varepsilon_{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \tag{4.1.23}$$

The form of S^{ij} shows that $\psi(x)$ is spin 1/2 under rotations, so we will refer to it as a Dirac spinor.

We then construct the two independent sets of generators

$$\begin{aligned} S_L^i &= \tfrac{1}{4}\,\varepsilon_{ijk}\,S^{jk} + \tfrac{i}{2}\,S^{0i} = \begin{pmatrix} \frac{1}{2}\sigma^i & 0 \\ 0 & 0 \end{pmatrix} \\ S_R^i &= \tfrac{1}{4}\,\varepsilon_{ijk}\,S^{jk} - \tfrac{i}{2}\,S^{0i} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2}\sigma^i \end{pmatrix}. \end{aligned} \tag{4.1.24}$$

Hence, all S_L^i commute with all S_R^i and each set of generators comprises an $SU(2)$ Lie algebra,

$$[S_L^i, S_L^j] = i\,\varepsilon_{ijk}\,S_L^k \qquad [S_R^i, S_R^j] = i\,\varepsilon_{ijk}\,S_R^k\,. \tag{4.1.25}$$

This algebra is the same found for three dimensional angular momentum. Hence, the Lorentz algebra decomposes as $SO(1,3) \simeq SU(2)_L \times SU(2)_R$, where the subscripts label the two $SU(2)$'s. Using the form of the generators in (4.1.24) we can write $\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$, where each of ψ_L and ψ_R have two components. ψ_L is a doublet under $SU(2)_L$ but a singlet (invariant) under $SU(2)_R$ while ψ_R is the opposite. We call the separate ψ_L and ψ_R parts of ψ Weyl fermions.

4.1.3 Solutions to the Dirac equation

To find the solutions to the Dirac equation we first Fourier transform $\psi(x)$,

$$\widetilde{\psi}(k) = \int d^4x\, e^{ik\cdot x}\, \psi(x)\,. \tag{4.1.26}$$

The Dirac equation then becomes

$$(\not{k} - m)\widetilde{\psi}(k) = 0\,. \tag{4.1.27}$$

Acting with $\not{k} + m$ on (4.1.27), we find that $k^2 = m^2$. In terms of the Weyl components we can express (4.1.27) as

$$\begin{aligned} (k^0 - \vec{k}\cdot\vec{\sigma})\widetilde{\psi}_R(k) - m\widetilde{\psi}_L(k) &= 0 \\ (k^0 + \vec{k}\cdot\vec{\sigma})\widetilde{\psi}_L(k) - m\widetilde{\psi}_R(k) &= 0\,. \end{aligned} \tag{4.1.28}$$

If $m = 0$ then we find that $\widetilde{\psi}_L$ and $\widetilde{\psi}_R$ decouple. In this case, when $k^0 > 0$ the spin direction aligns along $\vec{k}$ for $\widetilde{\psi}_R(k)$ and anti-aligns along $\vec{k}$ for $\widetilde{\psi}_L(k)$. It is in this sense that $\widetilde{\psi}_R(k)$ and $\widetilde{\psi}_L(k)$ are respectively right- and left-handed. We call the component of spin along $\vec{k}$ the *helicity*, h, so $\widetilde{\psi}_R(\vec{k})$ has helicity $h = 1/2$, while $\widetilde{\psi}_L(\vec{k})$ has helicity $h = -1/2$. Note that for $k^0 < 0$ it is the opposite.

Now let us take the other extreme and assume that $m \neq 0$ and we are in a frame where $\vec{k} = 0$. In this case we either have $k^0 = m$ and $\widetilde{\psi}_L = \widetilde{\psi}_R$, or $k_0 = -m$ and $\widetilde{\psi}_L = -\widetilde{\psi}_R$. In each of these cases there are two independent components, which correspond to the spin: S^{ij} acting on $\widetilde{\psi}(k)$ preserves the identification between $\widetilde{\psi}_L(k)$ and $\widetilde{\psi}_R(k)$ and clearly has eigenvalues $\pm\frac{1}{2}$.

To find the solution for general $\vec{k}$ we boost the solution with $\vec{k} = 0$. At this point it is convenient to define a parameter η called the "rapidity", which for the positive energy modes satisfies

$$k^0 = m\cosh\eta \qquad |\vec{k}| = m\sinh\eta\,. \tag{4.1.29}$$

This clearly satisfies the condition $(k^0)^2 - \vec{k}^{\,2} = m^2$. The boost that takes us to the new frame is then[2]

$$U_\gamma(\Lambda) = e^{-i\eta k^i S^{0i}/|\vec{k}|} = \cosh\frac{\eta}{2}\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \sinh\frac{\eta}{2}\begin{pmatrix} \vec{n}\cdot\vec{\sigma} & 0 \\ 0 & -\vec{n}\cdot\vec{\sigma} \end{pmatrix}, \tag{4.1.30}$$

where $\vec{n} = \dfrac{\vec{k}}{|\vec{k}|}$, the unit vector along $\vec{k}$. Choosing a basis such that the spin operator is diagonalized along $\vec{k}$, we find that the two independent components for $\widetilde{\psi}(k)$ are proportional to the Dirac spinors

$$u^1(k) \equiv \sqrt{m}\begin{pmatrix} e^{-\eta/2} \\ 0 \\ e^{\eta/2} \\ 0 \end{pmatrix}, \qquad u^2(k) \equiv \sqrt{m}\begin{pmatrix} 0 \\ e^{\eta/2} \\ 0 \\ e^{-\eta/2} \end{pmatrix}, \tag{4.1.31}$$

[2] Note that we boost in the opposite direction of $\vec{k}$.

where the factor of $\sqrt{m}$ is for later convenience. For a generic basis of spinors ξ^s, $s = 1, 2$ such that $\xi^{s\dagger}\xi^r = \delta^{sr}$, the independent normalized solutions for the positive frequency modes, $u^s(\vec{k})$, are

$$u^s(\vec{k}) = \left(\begin{array}{c} \left(\frac{1-\vec{n}\cdot\vec{\sigma}}{2}\sqrt{k^0 + |\vec{k}|} + \frac{1+\vec{n}\cdot\vec{\sigma}}{2}\sqrt{k^0 - |\vec{k}|}\right)\xi^s \\ \left(\frac{1+\vec{n}\cdot\vec{\sigma}}{2}\sqrt{k^0 + |\vec{k}|} + \frac{1-\vec{n}\cdot\vec{\sigma}}{2}\sqrt{k^0 - |\vec{k}|}\right)\xi^s \end{array} \right) \tag{4.1.32}$$

The combinations $\frac{1+\vec{n}\cdot\vec{\sigma}}{2}$ and $\frac{1-\vec{n}\cdot\vec{\sigma}}{2}$ are examples of projectors. A projector is any operator that squares to itself and a generic spinor ξ^s can be written as a sum over the projectors, $\xi^s = \frac{1+\vec{n}\cdot\vec{\sigma}}{2}\xi^s + \frac{1-\vec{n}\cdot\vec{\sigma}}{2}\xi^s$, which expresses ξ^s as a sum of eigenstates of $\vec{n}\cdot\vec{\sigma}$. If we define $\sigma^\mu \equiv (I, \sigma^i)$ and $\overline{\sigma}^\mu \equiv (I, -\sigma^i)$, then $u^s(\vec{k})$ can be written as

$$u^s(\vec{k}) = \left(\begin{array}{c} \sqrt{k\cdot\sigma}\,\xi^s \\ \sqrt{k\cdot\overline{\sigma}}\,\xi^s \end{array} \right), \tag{4.1.33}$$

where $k\cdot\sigma = k^\mu\sigma_\mu$. We understand this expression to be a sum over the eigenstates of $\vec{k}\cdot\vec{\sigma}$ and $\vec{k}\cdot\vec{\overline{\sigma}}$, with each term in the sum multiplied by the square root of the eigenvalue.

Under a Lorentz transformation $u^s(\vec{k})$ transforms as $u^s(\vec{k}) \to e^{+\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}u^s(\vec{k})$. But the inner product $u^{s\dagger}(\vec{k})u^r(\vec{k})$ is *not* Lorentz invariant . The γ-matrices satisfy $(\gamma^0)^\dagger = \gamma^0$, $(\gamma^i)^\dagger = -\gamma^i$. Therefore, $(S^{ij})^\dagger = S^{ij}$, while $(S^{0i})^\dagger = -S^{0i}$. It then follows that $(e^{+\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}})^\dagger \neq e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}$ if it is not a pure rotation. However, we also note that $\gamma^0 S^{ij} = S^{ij}\gamma^0$, while $\gamma^0 S^{0i} = -S^{0i}\gamma^0$. Hence, $(e^{+\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}})^\dagger\gamma^0 = \gamma^0 e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}$. Therefore, under a Lorentz transformation,

$$\overline{u}^s(\vec{k}) \equiv u^{s\dagger}(\vec{k})\gamma^0 \to \overline{u}^s(\vec{k})e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}, \tag{4.1.34}$$

and so $\overline{u}^s(\vec{k})u^r(\vec{k})$ is Lorentz invariant[3]. In particular the invariant is

$$\overline{u}^s(\vec{k})u^r(\vec{k}) = 2\sqrt{k^2}\,\delta^{sr} = 2m\,\delta^{sr}. \tag{4.1.35}$$

Here we used that $\sigma_i\sigma_j = -\sigma_j\sigma_i$ if $i \neq j$ from which it is one short step to show that $k\cdot\sigma k\cdot\overline{\sigma} = k^2$. One can also show the useful relation

$$\overline{u}^s(\vec{k})\gamma^\mu u^r(\vec{k}) = 2k^\mu\delta^{rs}. \tag{4.1.36}$$

which can be derived with an explicit computation, or with a clever trick[4].

[3] The reader might be confused about having Lorentz invariance even though explicit spin indices s and r are present in the invariant. The important point is that $s = 1, 2$ are labels for explicit spinors. For example, let us suppose that we have two observers A and B in inertial frames and that observer B's inertial frame is a 90 degree rotation in the $x - y$ plane from observer A's inertial frame. Now suppose we choose a basis of spinors such that for observer A they are $+$ and $-$ in the x direction. Observer B would say that the basis spinors are $+$ and $-$ in the y direction. The two spinors have not changed, only the coordinates used to describe them have changed.

[4] The trick will be useful for doing 3.2 in Peskin. (4.1.36) is a corrolary of the result you will derive there.)

We can derive another useful relation by summing over the two spin components. Since we can choose any spin basis we prefer, we can choose the one used in (4.1.31) where the spins are diagonalized along the spatial component of the momentum. Hence we find for Dirac components α and β, $\alpha,\beta = 1\ldots 4$,

$$\sum_{s=1,2} u^s_\alpha(\vec{k})\overline{u}^s_\beta(\vec{k}) = m\begin{pmatrix} 1 & 0 & e^{-\eta} & 0 \\ 0 & 1 & 0 & e^{\eta} \\ e^{\eta} & 0 & 1 & 0 \\ 0 & e^{-\eta} & 0 & 1 \end{pmatrix}_{\alpha\beta} = \begin{pmatrix} m & k\cdot\sigma \\ k\cdot\overline{\sigma} & m \end{pmatrix}_{\alpha\beta} = (\not{k}+m)_{\alpha\beta}\,. \tag{4.1.37}$$

For the negative frequency modes, the same boost that takes us to $(k^0,\vec{k})$ for the positive modes takes us to $(-k^0,-\vec{k})$ for the negative modes. Defining a new Dirac spinor $v^s(\vec{k})$, we have

$$v^s(\vec{k}) = \begin{pmatrix} +\sqrt{k\cdot\sigma}\,\xi^s \\ -\sqrt{k\cdot\overline{\sigma}}\,\xi^s \end{pmatrix}, \tag{4.1.38}$$

where the inner product is

$$\overline{v}^s(\vec{k})v^r(\vec{k}) = -2\sqrt{k^2} = -2m\delta^{sr}\,. \tag{4.1.39}$$

We also have

$$\overline{v}^s(\vec{k})\gamma^\mu v^r(\vec{k}) \;=\; 2k^\mu\delta^{rs}\,, \tag{4.1.40}$$

as well as

$$\sum_{s=1,2} v^s(\vec{k})_\alpha\overline{v}^s_\beta(\vec{k}) \;=\; m\begin{pmatrix} -1 & 0 & e^{-\eta} & 0 \\ 0 & -1 & 0 & e^{\eta} \\ e^{\eta} & 0 & -1 & 0 \\ 0 & e^{-\eta} & 0 & -1 \end{pmatrix}_{\alpha\beta} = \begin{pmatrix} -m & k\cdot\sigma \\ k\cdot\overline{\sigma} & -m \end{pmatrix}_{\alpha\beta} = (\not{k}-m)_{\alpha\beta}\,. \tag{4.1.41}$$

Note that while $u^s(\vec{k})$ satisfies (4.1.27), the equation for motion for $v^s(\vec{k})$ is

$$(\not{k}+m)v^s(\vec{k}) = 0\,. \tag{4.1.42}$$

Using this we can find the null inner products

$$\begin{aligned} \overline{u}^s(\vec{k})v^r(\vec{k}) &= 0 \\ \overline{u}^s(-\vec{k})\gamma^0 v^r(\vec{k}) &= 0\,. \end{aligned} \tag{4.1.43}$$

We can show these by inserting $\not{k}+m$ a the Dirac spinors,

$$0 \;=\; \overline{u}^s(\vec{k})(\not{k}+m)v^r(\vec{k}) = \Big((\not{k}+m)u(\vec{k})\Big)^\dagger\gamma^0 v(\vec{k}) = 2m\,\overline{u}^s(\vec{k})v^r(\vec{k}) \tag{4.1.44}$$

$$0 \;=\; \overline{u}^s(-\vec{k})\gamma^0(\not{k}+m)v^r(\vec{k}) = \Big((k^0\gamma^0+\vec{k}\cdot\vec{\gamma}+m)u(-\vec{k})\Big)^\dagger v(\vec{k}) = 2m\,\overline{u}^s(-\vec{k})\gamma^0 v^r(\vec{k})\,.$$

Here we used that $\gamma^0\gamma^i = -\gamma^i\gamma^0$ as well as $\gamma^{0\dagger} = \gamma^0$ and $\gamma^{i\dagger} = -\gamma^i$. Note that in general, $\overline{u}^s(-\vec{k})\gamma^\mu v^r(\vec{k}) \neq 0$.

A generic classical solution for $\psi(x)$ is then

$$\psi(x) = \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3 2k^0}\left(e^{-ik\cdot x}\, A^s_{\vec{k}}\, u^s(\vec{k}) + e^{+ik\cdot x}\, B^{s*}_{\vec{k}}\, v^s(\vec{k})\right), \tag{4.1.45}$$

where $k^0 = +\sqrt{\vec{k}\cdot\vec{k}}$ and $B^{s*}_{\vec{k}} \neq A^{s*}_{\vec{k}}$. Since $\psi(x)$ is not real we also have

$$\overline{\psi}(x) = \psi^\dagger(x)\gamma^0 = \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3 2k^0}\left(e^{+ik\cdot x}\, A^{s*}_{\vec{k}}\, \overline{u}^s(\vec{k}) + e^{-ik\cdot x}\, B^s_{\vec{k}}\, \overline{v}^s(\vec{k})\right), \tag{4.1.46}$$

4.2 Quantization of the Dirac field

Let us now take the same approach as we did for the scalar field and treat the Dirac spinor $\psi(x)$ as an operator field and not as a wave-function. To quantize this field, we again replace $A^s_{\vec{k}}$, $B^s_{\vec{k}}$ and their conjugates with creation and annihilation operators $a^s{}_{\vec{k}}$, $b^s{}_{\vec{k}}$, $a^{s\dagger}{}_{\vec{k}}$ and $b^{s\dagger}{}_{\vec{k}}$,

$$\begin{aligned}
\psi(x) &= \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3 2k^0}\left(e^{-ik\cdot x}\, a^s{}_{\vec{k}}\, u^s(\vec{k}) + e^{+ik\cdot x}\, b^{s\dagger}{}_{\vec{k}}\, v^s(\vec{k})\right), \\
\overline{\psi}(x) &= \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3 2k^0}\left(e^{+ik\cdot x}\, a^{s\dagger}{}_{\vec{k}}\, \overline{u}^s(\vec{k}) + e^{-ik\cdot x}\, b^s{}_{\vec{k}}\, \overline{v}^s(\vec{k})\right).
\end{aligned} \tag{4.2.1}$$

We also assume that we have the commutation relations

$$[a^s{}_{\vec{k}}, a^{r\dagger}{}_{\vec{k}'}] = (2\pi)^3(2k^0)\delta^3(\vec{k}-\vec{k}')\delta^{sr}, \qquad [b^s_{\vec{k}}, b^{r\dagger}{}_{\vec{k}'}] = (2\pi)^3(2k^0)\delta^3(\vec{k}-\vec{k}')\delta^{sr}, \tag{4.2.2}$$

with all other commutators zero[5]. We also assume that there is a vacuum state where $a_{\vec{k}}|0\rangle = b_{\vec{k}}|0\rangle = 0$, so that there are no negative energy states.

With the commutation relations in (4.2.2) it is clear that $[\psi_\alpha(x), \psi_\beta(y)] = 0$, where $\alpha, \beta = 1\ldots4$. But consider $[\psi_\alpha(x), \psi^\dagger_\beta(y)]$ and let us further assume that $(x-y)^2 < 0$. To make life simpler, we choose an inertial frame where $x^0 = y^0$. Then the commutator of $\psi_\alpha(x)$ with $\psi^\dagger_\beta(y)$ is

$$\begin{aligned}
&[\psi_\alpha(x^0\vec{x}), \psi^\dagger_\beta(x^0,\vec{y})] \\
&\quad= \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3 2k^0}\left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})}u^s_\alpha(\vec{k})(\overline{u}^s(\vec{k})\gamma^0)_\beta - e^{-i\vec{k}\cdot(\vec{x}-\vec{y})}v^s_\alpha(\vec{k})(\overline{v}^s(\vec{k})\gamma^0)_\beta\right) \\
&\quad= \int \frac{d^3k}{(2\pi)^3 2k^0}\left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})}((\not{k}+m)\gamma^0)_{\alpha\beta} - e^{-i\vec{k}\cdot(\vec{x}-\vec{y})}((\not{k}-m)\gamma^0)_{\alpha\beta}\right)
\end{aligned} \tag{4.2.3}$$

[5] Since we are using $2k^0$ instead of $\sqrt{2k^0}$ in the measure factors in (4.1.45), the commutators will come with an extra factor of $2k^0$. Peskin's convention is to carry around the factors of $\sqrt{2k^0}$.

which in general is *not* zero, hence violating causality. This is not good. However, a closer inspection of the last line in (4.2.3) reveals that if the second term had come with a + sign instead of a − sign, then we would have found the result

$$\begin{aligned}
\overset{-\to+}{=} &\int\frac{d^3k}{(2\pi)^3 2k^0}\left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})}((\not{k}+m)\gamma^0)_{\alpha\beta}+e^{-i\vec{k}\cdot(\vec{x}-\vec{y})}((\not{k}-m)\gamma^0)_{\alpha\beta}\right)\\
= &\int\frac{d^3k}{(2\pi)^3 2k^0}\left(e^{i\vec{k}\cdot(\vec{x}-\vec{y})}(2k^0\gamma^0\gamma^0)_{\alpha\beta}\right)=\delta^3(\vec{x}-\vec{y})\delta_{\alpha\beta}\,, \qquad (4.2.4)
\end{aligned}$$

where we used that the terms in the integrand proportional to $\vec{k}$ and m are odd under $\vec{k}\to-\vec{k}$, while the k^0 term is even. The result in (4.2.4) is exactly the type of term we want!

So now we ask the question how does one change the − sign in (4.2.3) to a + sign? The key is noticing that we got the minus sign from a commutator, which is automatically antisymmetric in its entries, $[b^\dagger{}_{\vec{k}}, b_{\vec{k}'}] = -[b_{\vec{k}'}, b^\dagger{}_{\vec{k}}]$. To get a + sign, we should use an analog of a commutator which is automatically *symmetric* in its indices[6]. To this end we define the *anticommutator*,

$$\{A,B\}\equiv AB+BA\,. \qquad (4.2.5)$$

We then replace the commutation relations in (4.2.2) with the anticommutation relations

$$\begin{aligned}
&\{a^s{}_{\vec{k}}\,,a^{r\dagger}{}_{\vec{k}'}\}=(2\pi)^3(2k^0)\delta^3(\vec{k}-\vec{k}')\delta^{sr}, \qquad \{b^s{}_{\vec{k}}\,,b^{r\dagger}{}_{\vec{k}'}\}=(2\pi)^3(2k^0)\delta^3(\vec{k}-\vec{k}')\delta^{sr},\\
&\{a^s{}_{\vec{k}}\,,a^r{}_{\vec{k}'}\}=\{b^s{}_{\vec{k}}\,,b^r{}_{\vec{k}'}\}=\{a^{s\dagger}{}_{\vec{k}}\,,a^{r\dagger}{}_{\vec{k}'}\}=\{b^{s\dagger}{}_{\vec{k}}\,,b^{r\dagger}{}_{\vec{k}'}\}=0\\
&\{a^s{}_{\vec{k}}\,,b^r{}_{\vec{k}'}\}=\{a^s{}_{\vec{k}}\,,b^{r\dagger}{}_{\vec{k}'}\}=\{a^{s\dagger}{}_{\vec{k}}\,,b^r{}_{\vec{k}'}\}=\{a^{s\dagger}{}_{\vec{k}}\,,b^{r\dagger}{}_{\vec{k}'}\}=0\,. \qquad (4.2.6)
\end{aligned}$$

Therefore, the anticommutation relations for the Dirac fields are

$$\begin{aligned}
&\{\psi_\alpha(x),\psi_\beta(y)\}=\{\psi^\dagger_\alpha(x),\psi\dagger_\beta(y)\}=0\\
&\{\psi_\alpha(x^0,\vec{x}),\psi^\dagger_\beta(x^0,\vec{y})\}=\delta^3(\vec{x}-\vec{y})\,\delta_{\alpha\beta}\,. \qquad (4.2.7)
\end{aligned}$$

We still assume that the vacuum is annihilated by $a^s{}_{\vec{k}}$ and $b^s{}_{\vec{k}}$. If we define $c^{+s\dagger}{}_{\vec{k}} = a^{s\dagger}{}_{\vec{k}}$ and $c^{-s\dagger}{}_{\vec{k}} = b^{s\dagger}{}_{\vec{k}}$, then the Fock space for the Dirac field is made up of the n-particle states which we write as

$$|\vec{k}_1,q_1,s_1;\vec{k}_2,q_2,s_2;\ldots\vec{k}_n,q_n,s_n\rangle\equiv\prod_{j=1}^n c^{q_js_j\dagger}{}_{\vec{k}_j}\,|0\rangle\,, \qquad (4.2.8)$$

where $q_j = \pm$. This Fock space is significantly different from that of the free scalar field. Firstly, the particles now have internal quantum numbers: their spin as well as the quantum number q. Second, under an exchange of momentum, q and spin, the states satisfy

$$|\ldots\vec{k}_j,q_j,s_j;\ldots\vec{k}_k,q_k,s_k;\ldots\rangle=-|\ldots\vec{k}_k,q_k,s_k;\ldots\vec{k}_j,q_j,s_j;\ldots\rangle\,, \qquad (4.2.9)$$

[6] Another way to try and solve the problem of the + sign is to make the switch $b^s{}_{\vec{k}}\leftrightarrow b^{s\dagger}{}_{\vec{k}}$. However, this just trades one disastrous result for another, because with this convention $b^{s\dagger}{}_{\vec{k}}$ would create negative energy states and the total energy would not be bounded below.

hence, the particles are identical fermions. If we have identical fermions then no two particles can be in the same state. This immediately follows from the anti-commutation relations. For example, consider

$$|\vec{k},+,s;\vec{k},+,s\rangle = a^{s\dagger}{}_{\vec{k}}a^{s\dagger}{}_{\vec{k}}|0\rangle\,. \tag{4.2.10}$$

But by the anticommutation relations we have that $\{a^{s\dagger}{}_{\vec{k}}, a^{s\dagger}{}_{\vec{k}}\} = 2a^{s\dagger}{}_{\vec{k}}a^{s\dagger}{}_{\vec{k}} = 0$, hence no such state exists.

To summarize, requiring Lorentz invariance forces our spin 1/2 particles to be fermions. We have previously seen that the the particles from scalar fields are bosons. The two results are part of the *spin statistics theorem*, which states that integer spin particles are bosons and half-integer spin particles are fermions.

4.3 The Dirac propagator

As in the case of the real scalar field, our main interest are the time-ordered correlators for the Dirac fields. Clearly,

$$\langle\psi_\alpha(x)\psi_\beta(y)\rangle = \langle\overline{\psi}_\alpha(x)\overline{\psi}_\beta(y)\rangle = 0\,, \tag{4.3.1}$$

so the only relevant two point function is

$$\langle T[\psi_\alpha(x)\overline{\psi}_\beta(y)]\rangle\,. \tag{4.3.2}$$

Because we are dealing with fermionic fields, the time ordering is defined as

$$\begin{aligned}\langle T[\psi_\alpha(x)\overline{\psi}_\beta(y)]\rangle &= \langle\psi_\alpha(x)\overline{\psi}_\beta(y)\rangle \qquad x^0 > y^0\,,\\ &= -\langle\overline{\psi}_\beta(y)\psi_\alpha(x)\rangle \qquad x^0 < y^0\,,\end{aligned} \tag{4.3.3}$$

Using (4.2.1) and the anticommutation relations in (4.2.6) we find

$$\begin{aligned}S_F(x-y)_{\alpha\beta} &\equiv \langle T[\psi_\alpha(x)\overline{\psi}_\beta(y)]\rangle\\ &= \int\frac{d^3k}{(2\pi)^3 2k^0}\left(\theta(x^0-y^0)\,e^{-ik\cdot(x-y)}(\not{k}+m)_{\alpha\beta} - \theta(y^0-x^0)\,e^{+ik\cdot(x-y)}(\not{k}-m)_{\alpha\beta}\right)\,.\end{aligned} \tag{4.3.4}$$

We can drop the α and β indices, understanding that the correlator is a 4×4 matrix. If we now act on this correlator with $i\not{\partial}-m$ we find

$$\begin{aligned}&(i\not{\partial}-m)S_F(x-y) =\\ &\int\frac{d^3k}{(2\pi)^3 2k^0}\Big[\theta(x^0-y^0)\,e^{-ik\cdot(x-y)}(\not{k}-m)(\not{k}+m) - \theta(y^0-x^0)\,e^{+ik\cdot(x-y)}(\not{k}+m)(\not{k}-m)\\ &\qquad +i\,\partial_{x^0}\theta(x^0-y^0)\,e^{-ik\cdot(x-y)}\gamma^0(\not{k}+m) - i\,\partial_{x^0}\theta(y^0-x^0)\,e^{+ik\cdot(x-y)}\gamma^0(\not{k}-m)\Big]\\ &\qquad = i\,\delta^4(x-y)\,,\end{aligned} \tag{4.3.5}$$

where we used that $(\not{k}-m)(\not{k}+m) = k^2 - m^2 = 0$, $\partial_{x^0}\theta(x^0-y^0) = \delta(x^0 - y^0)$, $\gamma^0\gamma^0 = 1$ and substituted $-\vec{k}$ for $\vec{k}$ in the second term of the second line. Therefore, $-i\, S_F(x-y)$ is a Green's function for $i\not{\partial} - m$.

As in the case of the scalar field, the expression for the correlator in momentum space is much more compact. Here we find

$$\begin{aligned}
\widetilde{S}_F(k) &= \int d^4x\, e^{ik\cdot x} \int \frac{d^3q}{(2\pi)^3 2q^0} \left(\theta(x^0)\, e^{-iq\cdot x}(\not{q}+m) - \theta(-x^0)\, e^{+iq\cdot x}(\not{q}-m)\right) \\
&= \frac{1}{2\omega(\vec{k})}\left(\frac{i}{k^0-\omega(\vec{k})+i\,\epsilon}((\omega(\vec{k})-k^0)\gamma^0+\not{k}+m) + \frac{i}{k^0+\omega(\vec{k})-i\,\epsilon}((\omega(\vec{k})+k^0)\gamma^0-\not{k}-m)\right) \\
&= \frac{i(\not{k}+m)}{k^2-m^2+i\,\epsilon} = \frac{i}{\not{k}-m+i\,\epsilon}\,,
\end{aligned} \tag{4.3.6}$$

where we replaced $\omega(\vec{k})$ by $\omega(\vec{k}) - i\,\epsilon$ to make the integrals well defined. In the last step we used that

$$k^2 - m^2 + i\,\epsilon = (\not{k} - m + i\,\epsilon)(\not{k} + m - i\,\epsilon)\,. \tag{4.3.7}$$

We interpret the propagator $\langle T[\psi(x)\overline{\psi}(y)]\rangle_{\alpha\beta}$ to mean that for $x^0 > y^0$, $\overline{\psi}(y)$ creates a fermion with $q = -1$ out of the vacuum. The particle then propagates to x^μ. Instead of specifying a spin we have a Dirac index β for the fermion at y^μ and another Dirac index α for the fermion at x^μ. If $x^0 < y^0$ then $\psi(x)$ creates an *anti-fermion* out of the vacuum with $q = +1$ which then propagates to y^μ. Note that if $x^\mu - y^\mu$ is space-like then one observer could see a fermion going forward in time from y^μ to x^μ, while another observer would see an antifermion going forward in time from x^μ to y^μ. Thus, by Lorentz invariance we see that for every fermion there must be a corresponding anti-fermion.

By inserting a complete set of single particle states, we can write the single particle fermion state $|y^\mu, \beta\rangle$ in terms of spins and momenta as

$$\begin{aligned}
|y^\mu, \beta'\rangle\gamma^0_{\beta'\beta} &= \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3(2k^0)}\, |\vec{k}, s\rangle\langle\vec{k}, s|y^\mu, \beta'\rangle\gamma^0_{\beta'\beta} \\
&= \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3(2k^0)}\, |\vec{k}, s\rangle\, e^{ik\cdot x}\, \overline{u}^s(\vec{k})_\beta \\
&= \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3(2k^0)}\, e^{ik\cdot x}\, \overline{u}^s(\vec{k})_\beta\, a^{s\dagger}{}_{\vec{k}}\, |0\rangle = \overline{\psi}_\beta(y)\, |0\rangle\,.
\end{aligned} \tag{4.3.8}$$

It might seem that there are four independent components for α even though there are only two spins. However, only two of the components are independent because the Dirac equation $(\not{k} - m)u^s(\vec{k}) = 0$ relates the components to each other.

4.4 The Dirac field Lagrangian

The Lagrangian $\mathcal{L}$ should have the form

$$\mathcal{L} = \Pi_\alpha(x)\partial_0\psi_\alpha(x) - \mathcal{H}(x) \tag{4.4.1}$$

where $\mathcal{H}(x)$ is the Lagrangian. Looking back at (4.2.7) we observe that $i\psi^\dagger_\alpha(x)$ has the correct anticommutation relations with $\psi_\alpha(x)$ to be identified with $\Pi_\alpha(x)$. Hence, the first term in (4.4.1) can be written as $i\psi^\dagger(x)\partial_0\psi(x) = i\overline{\psi}(x)\gamma^0\partial_0\psi(x)$. Alternatively, we could have called $\psi^\dagger_\alpha(x)$ the coordinate and $i\psi_\alpha(x)$ the canonical momentum, in which case we would have $i\psi(x)\partial_0\psi^\dagger(x)$. In the action we can integrate by parts, so this last term is equivalent to $-i\partial_0\psi(x)\psi^\dagger(x)$. Furthermore, the classical Dirac fields anticommute with each other, hence the last term is the same as our original expression. Let me stress that the anticommutativity is crucial for consistency.

The Hamiltonian can be written as

$$\begin{aligned} H &= \sum_{s=1,2}\int\frac{d^3k}{(2\pi)^3(2k^0)}\,\frac{k^0}{2}\left(a^{s\dagger}{}_{\vec{k}}a^s{}_{\vec{k}} - a^s{}_{\vec{k}}a^{s\dagger}{}_{\vec{k}} + b^{s\dagger}{}_{\vec{k}}b^s{}_{\vec{k}} - b^s{}_{\vec{k}}b^{s\dagger}{}_{\vec{k}}\right) \\ &= \sum_{s=1,2}\int\frac{d^3k}{(2\pi)^3(2k^0)}\,k^0\left(a^{s\dagger}{}_{\vec{k}}a^s{}_{\vec{k}} - b^s{}_{\vec{k}}b^{s\dagger}{}_{\vec{k}}\right) \end{aligned} \tag{4.4.2}$$

We then observe that

$$\begin{aligned} k^0\,a^{s\dagger}{}_{\vec{k}}\,a^s{}_{\vec{k}} &= \sum_{r=1,2}\int d^3x\int\frac{d^3q}{(2\pi)^3 2q^0}2k^0e^{ix\cdot(q-k)}\frac{1}{2}\,a^{r\dagger}{}_{\vec{q}}\,\overline{u}^r(\vec{q})\gamma^0u^s(\vec{k})\,a^s{}_{\vec{k}} \\ &= \sum_{r=1,2}\int d^3x\int\frac{d^3q}{(2\pi)^3 2q^0}e^{ix\cdot(q-k)}\,a^{r\dagger}{}_{\vec{q}}\,\overline{u}^r(\vec{q})(\vec{k}\cdot\vec{\gamma}+m)u^s(\vec{k})\,a^s{}_{\vec{k}} \\ &= \sum_{r=1,2}\int d^3x\int\frac{d^3q}{(2\pi)^3 2q^0}e^{ix\cdot q}\,a^{r\dagger}{}_{\vec{q}}\,\overline{u}^r(\vec{q})(-i\vec{\nabla}\cdot\vec{\gamma}+m)e^{-ix\cdot k}\,a^s{}_{\vec{k}}u^s(\vec{k}) \\ -k^0\,b^{s\dagger}{}_{\vec{k}}\,b^s{}_{\vec{k}} &= \sum_{r=1,2}\int d^3x\int\frac{d^3q}{(2\pi)^3 2q^0}2k^0e^{-ix\cdot(q-k)}\frac{1}{2}\,b^{r\dagger}{}_{\vec{q}}\,\overline{v}^r(\vec{q})(-\gamma^0)v^s(\vec{k})\,b^s{}_{\vec{k}} \qquad (4.4.3) \\ &= \sum_{r=1,2}\int d^3x\int\frac{d^3q}{(2\pi)^3 2q^0}e^{-ix\cdot(q-k)}\,b^{r\dagger}{}_{\vec{q}}\,\overline{v}^r(\vec{q})(-\vec{k}\cdot\vec{\gamma}+m)v^s(\vec{k})\,b^s{}_{\vec{k}} \\ &= \sum_{r=1,2}\int d^3x\int\frac{d^3q}{(2\pi)^3 2q^0}e^{-ix\cdot q}\,b^{r\dagger}{}_{\vec{q}}\,\overline{v}^r(\vec{q})(-i\vec{\nabla}\cdot\vec{\gamma}+m)e^{ix\cdot k}\,b^s{}_{\vec{k}}v^s(\vec{k})\,. \end{aligned}$$

We also use

$$\begin{aligned} \overline{u}^r(-\vec{k})(-\vec{k}\cdot\vec{\gamma}+m)v^s(\vec{k}) = -k^0\overline{u}^r(-\vec{k})\gamma^0v^s(\vec{k}) &= 0 \\ \overline{v}^r(-\vec{k})(\vec{k}\cdot\vec{\gamma}+m)u^s(\vec{k}) = k^0\overline{v}^r(-\vec{k})\gamma^0u^s(\vec{k}) &= 0\,, \end{aligned} \tag{4.4.4}$$

to finally write the Hamiltonian as

$$\begin{aligned} H = &\int d^3x\left(\sum_{r=1,2}\int\frac{d^3q}{(2\pi)^3 2q^0}\left(e^{ix\cdot q}\,a^{r\dagger}{}_{\vec{q}}\,\overline{u}^r(\vec{q}) + e^{-ix\cdot q}\,b^r{}_{\vec{q}}\,\overline{v}^r(\vec{q})\right)\right) \\ &\qquad\times(-i\vec{\nabla}\cdot\vec{\gamma}+m)\left(\sum_{s=1,2}\int\frac{d^3k}{(2\pi)^3(2k^0)}\left(e^{-ix\cdot k}\,a^s{}_{\vec{k}}\,u^s(\vec{k}) + e^{ix\cdot k}\,b^{s\dagger}{}_{\vec{k}}\,v^s(\vec{k})\right)\right) \\ = &\int d^3x\,\overline{\psi}(x)(-i\vec{\nabla}\cdot\vec{\gamma}+m)\psi(x) = \int d^3x\,\mathcal{H}(x)\,. \end{aligned} \tag{4.4.5}$$

The Lagrangian is then

$$\mathcal{L}(x) = i\overline{\psi}(x)\gamma^0\partial^0\psi(x) - \mathcal{H}(x) = \overline{\psi}(x)(i\partial\!\!\!/ - m)\psi(x)\,. \tag{4.4.6}$$

Since $\mathcal{L}$ is dimension 4 in 3+1 dimensions, $\psi(x)$ is dimension 3/2.

The equations of motion should follow from $\mathcal{L}(x)$. Treating $\psi(x)$ and $\overline{\psi}(x)$ as independent we have

$$-\frac{\partial\mathcal{L}}{\partial\overline{\psi}} = (-i\partial\!\!\!/ + m)\psi(x) = 0\,, \tag{4.4.7}$$

and

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} - \frac{\partial\mathcal{L}}{\partial\psi} = -i\,\partial_\mu\overline{\psi}\gamma^\mu - m\overline{\psi} = \overline{\psi}(-i\overleftarrow{\partial\!\!\!/} - m) = 0\,, \tag{4.4.8}$$

where the backward pointing arrow means the derivative is acting toward the left. In deriving this last result we used that $\frac{\partial}{\partial\psi}\overline{\psi} = -\overline{\psi}\frac{\partial}{\partial\psi}$. We can also derive the last equation of motion by integrating by parts in the action $S = \int d^4\mathcal{L}(x)$, which leads to

$$\mathcal{L}(x) = \overline{\psi}(x)(-i\overleftarrow{\partial\!\!\!/} - m)\psi(x)\,. \tag{4.4.9}$$

From here we just set $\frac{\partial\mathcal{L}}{\partial\psi} = 0$.

By treating $\psi(x)$ as a field the Dirac equation becomes a classical equation of motion. But it is an equation of motion for a fermion field and classically two fermion fields satisfy $\psi^1(x)\psi^2(y) = -\psi^2(y)\psi^1(x)$.[7] Hence $\psi(x)$ cannot be an observable since observables are real numbers. However, fermion bilinears, that is two fermi fields together, can be observables (*e.g.* $\overline{\psi}(x)\gamma^\mu\psi(x)$ is an observable).

4.5 Symmetries

The Dirac Lagrangian (4.4.6) has some interesting symmetries which we now discuss.

4.5.1 Lorentz symmetries and particle spins

The fermions and anti-fermions have spin labels, s, but we have not yet checked how the labels correspond to the actual spin of the particles. In this subsection we address this question.

To find the spin that goes along with the label s, we need to construct the angular momentum currents and from this the charges. To find the currents we invoke Noether's theorem. The change to $\psi(x)$ under Lorentz transformations follows from (4.1.22),

$$\psi'(x) = e^{+\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}\psi(\Lambda^{-1}x) \tag{4.5.1}$$

[7]The anti-commutation of the fields does not go away in a classical limit. If we insert explicit factors of $\hbar$ into the anti-commutation relations in (4.2.7), then the righthand side has an $\hbar$ factor. Taking the classical limit sets the righthand side to zero leaving completely anti-commuting fields.

where the $\Lambda^{-1}x$ factor comes from replacing $x^{\mu'}$ with x^μ in $\psi'(x')$. Therefore, the infinitesimal transformation to $\psi(x)$ is

$$\delta\psi(x) = +\frac{i}{2}\epsilon_{\mu\nu}\left(S^{\mu\nu} + i\left(x^\mu\partial^\nu - x^\nu\partial^\mu\right)\right)\psi(x)\,, \tag{4.5.2}$$

and so the Noether current is

$$\begin{aligned} j^{\lambda\mu\nu} &= \frac{\partial\mathcal{L}}{\partial\partial_\lambda\psi(x)}(+i)\left(S^{\mu\nu} + i\left(x^\mu\partial^\nu - x^\nu\partial^\mu\right)\right)\psi(x) + \left(\eta^{\nu\lambda}x^\mu - \eta^{\mu\lambda}x^\nu\right)\mathcal{L}(x) \\ &= -i\,\overline{\psi}\left(+i\,\gamma^\lambda S^{\mu\nu} - x^\mu\left(\gamma^\lambda\partial^\nu + \eta^{\lambda\nu}\partial\!\!\!/\right) + x^\nu\left(\gamma^\lambda\partial^\mu + \eta^{\lambda\mu}\partial\!\!\!/\right)\right)\psi\,. \end{aligned} \tag{4.5.3}$$

Therefore the charges are

$$Q^{\mu\nu} = \int d^3 j^{0\mu\nu} = \int d^3x\,\psi^\dagger(x)\,S^{\mu\nu}\psi(x) + \dots \equiv Q^{\mu\nu}_\Sigma + \dots\,, \tag{4.5.4}$$

where the dots represent contributions from the derivatives. This part of the charge is due to the orbital angular momentum. Note that there is only one charge per $\mu\nu$ component so spin and orbital angular momentum will not be conserved separately.

If we now substitute in the expressions for $\psi(x)$ and $\psi^\dagger(x)$ from (4.2.1), set $x^0 = 0$ and integrate over d^3x, we get

$$Q^{\mu\nu} = \sum_{r,s=1,2}\int\frac{d^3k}{(2\pi)^3(2k^0)^2}\left(u^{r\dagger}(\vec{k})a^{r\dagger}_{\vec{k}} + v^{r\dagger}(-\vec{k})b^r_{-\vec{k}}\right)S^{\mu\nu}\left(u^s(\vec{k})a^s_{\vec{k}} + v^s(-\vec{k})b^{s\dagger}_{-\vec{k}}\right) + \dots \tag{4.5.5}$$

We now use the fact that $Q^{\mu\nu}|0\rangle = 0$ by the Lorentz invariance of the vacuum. Therefore, acting with the Lorentz charges on a single fermion state $a^{s\dagger}_{\vec{k}}|0\rangle$ gives

$$Q^{\mu\nu}a^{s\dagger}_{\vec{k}}|0\rangle = [Q^{\mu\nu}, a^{s\dagger}_{\vec{k}}]|0\rangle\,. \tag{4.5.6}$$

In order to simplify things, we consider only Lorentz charges corresponding to rotations in the plane orthogonal to $\vec{k}$. In this case

$$[Q^{ij}, a^{s\dagger}_{\vec{k}}]|0\rangle = [Q^{ij}_\Sigma, a^{s\dagger}_{\vec{k}}]|0\rangle = \frac{1}{2k^0}\sum_{r=1,2}u^{r\dagger}(\vec{k})\,S^{ij}u^s(\vec{k})a^{r\dagger}_{\vec{k}}|0\rangle\,. \tag{4.5.7}$$

Using that $S^{ij} = \frac{1}{2}\varepsilon_{ijk}\begin{pmatrix}\sigma^k & 0\\ 0 & \sigma^k\end{pmatrix}$ and that $k^iS^{ij} = 0$, we then find

$$Q^{ij}a^{s\dagger}_{\vec{k}}|0\rangle = \tfrac{1}{2}\varepsilon_{ijk}\sum_{r=1,2}\sigma^{k,sr}a^{r\dagger}_{\vec{k}}|0\rangle\,, \tag{4.5.8}$$

where σ^k is the spin component along $\vec{k}$. If we choose this spin to be diagonal, then the two values of s correspond to spin $\pm\frac{1}{2}$.

We next consider a single anti-fermion, where we find

$$\begin{aligned}
Q^{\mu\nu} b^{s\dagger}_{\vec{k}}|0\rangle &= [Q^{ij}_{\Sigma}, b^{s\dagger}_{\vec{k}}]|0\rangle = -\frac{1}{2k^0}\sum_{r=1,2} v^{s\dagger}(\vec{k})\, S^{ij} v^{r}(\vec{k}) b^{r\dagger}_{\vec{k}}|0\rangle \\
&= -\tfrac{1}{2}\varepsilon_{ijk}\sum_{r=1,2}\sigma^{k,sr} b^{r\dagger}_{\vec{k}}|0\rangle\,. \qquad (4.5.9)
\end{aligned}$$

Hence, we again see that in the diagonal basis the two values of s correspond to spin $\pm\frac{1}{2}$, but the assignment of the spins to s is the opposite of the fermion's assignment.

This has an interesting consequence for Weyl fermions. For a right-handed Weyl fermion, the spin is aligned along $\vec{k}$. The argument in the previous paragraph shows that for the anti-fermion, the spin is anti-aligned along $\vec{k}$. Hence, the anti-fermion is actually left-handed. For a left-handed Weyl fermion the situation is the opposite.

4.5.2 Internal continuous symmetries

The Lagrangian is invariant under the transformation $\psi(x) \to e^{i\theta}\psi(x)$, since $\overline{\psi}(x)$ will then transform as

$$\overline{\psi}(x) = \psi^\dagger(x)\gamma^0 \to \psi^\dagger(x)e^{-i\theta}\gamma^0 = \overline{\psi}(x)e^{-i\theta}\,. \qquad (4.5.10)$$

The infinitesimal transformation is

$$\psi(x) \to \psi(x) + i\,\epsilon\,\psi(x)\,, \qquad \overline{\psi}(x) \to \overline{\psi}(x) - i\,\epsilon\,\overline{\psi}(x)\,. \qquad (4.5.11)$$

Hence, using Noether's theorem we find for the current (often called the vector current)

$$j^\mu_V(x) = \frac{\partial\mathcal{L}(x)}{\partial\partial_\mu\psi} i\psi(x) = -i\overline{\psi}(x)\gamma^\mu(i\,\psi(x)) = \overline{\psi}(x)\gamma^\mu\psi(x) \qquad (4.5.12)$$

where the subscript V stands for vector. We can then quickly show that the charge is

$$Q = \int d^3 j^0_V = \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3 2k^0}\left(a^{s\dagger}{}_{\vec{k}} a^{s}{}_{\vec{k}} + b^{s}{}_{\vec{k}} b^{s\dagger}{}_{\vec{k}}\right)\,. \qquad (4.5.13)$$

If we then consider the commutator of Q with the creation and annihilation operators we get

$$[Q, a^{s\dagger}{}_{\vec{k}}] = +a^{s\dagger}{}_{\vec{k}}\,, \qquad [Q, a^{s}{}_{\vec{k}}] = -a^{s}{}_{\vec{k}}\,, \qquad [Q, b^{s\dagger}{}_{\vec{k}}] = -b^{s\dagger}{}_{\vec{k}}\,, \qquad [Q, b^{s}{}_{\vec{k}}] = +b^{s}{}_{\vec{k}}\,, \qquad (4.5.14)$$

where in deriving these commutators we used that

$$[a^{r\dagger}{}_{\vec{q}}\, a^{r}{}_{\vec{q}}, a^{s\dagger}{}_{\vec{k}}] = a^{r\dagger}{}_{\vec{q}}\{a^{r}{}_{\vec{q}}, a^{s\dagger}{}_{\vec{k}}\} - \{a^{r\dagger}{}_{\vec{q}}, a^{s\dagger}{}_{\vec{k}}\}a^{r}{}_{\vec{q}}\,, \qquad (4.5.15)$$

etc.

Hence, $a^{s\dagger}{}_{\vec{k}}$ creates a particle with charge $q = +1$ while $b^{s\dagger}{}_{\vec{k}}$ creates a particle with charge $q = -1$. These are the same values for q discussed in section 2. When we discuss

QED we will see that these correspond to the electric charges of the positron and the electron. We can also derive the commutator of Q directly with the fields,

$$[Q,\psi(x)] = -\psi(x) \qquad [Q,\overline{\psi}(x)] = +\overline{\psi}(x)\,, \tag{4.5.16}$$

which follows from the commutators in (4.5.14), the definition of j^0 in (4.5.12) and the anticommutation relations in (4.2.7).

If $m=0$ then there is another continuous symmetry present called a "chiral" symmetry. To see how to implement this symmetry we define a new γ-matrix

$$\gamma^5 \equiv i\,\gamma^0\gamma^1\gamma^2\gamma^3\,. \tag{4.5.17}$$

If we square it we find

$$(\gamma^5)^2 = -(\gamma^0\gamma^1\gamma^2\gamma^3)(\gamma^0\gamma^1\gamma^2\gamma^3) = -(-1)^2(\gamma^0\gamma^0\gamma^1\gamma^1\gamma^2\gamma^2\gamma^3\gamma^3) = +1\,. \tag{4.5.18}$$

It is also straightforward to show that $\{\gamma^5,\gamma^\mu\}=0$, since γ^μ commutes with one of the γ-matrices within γ^5 and anti-commutes with the other three. Because $(\gamma^5)^2=1$, we can construct the projectors $\frac{1}{2}(1\pm\gamma^5)$,

$$\left(\tfrac{1}{2}(1\pm\gamma^5)\right)^2 = \tfrac{1}{2}(1\pm\gamma^5)\,. \tag{4.5.19}$$

To see what the projectors project onto, it is convenient to go to the Weyl basis, where we find

$$\gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & +I \end{pmatrix}, \tag{4.5.20}$$

thus we find

$$\begin{aligned} \tfrac{1}{2}(1-\gamma^5)\psi &= \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \psi_L \\ 0 \end{pmatrix} \\ \tfrac{1}{2}(1+\gamma^5)\psi &= \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} 0 \\ \psi_R \end{pmatrix}. \end{aligned} \tag{4.5.21}$$

The projectors project onto the left or the right handed Weyl spinors.

Let us now use γ^5 to generate a symmetry. Since γ^5 anti-commutes with every γ-matrix, it commutes with the product of any two. Therefore, under the transformation $\psi\to e^{i\theta\gamma^5}\psi$ we have that

$$\begin{aligned} \overline{\psi} = \psi^\dagger\gamma^0 \quad &\to \quad \psi^\dagger e^{-i\theta\gamma^5}\gamma^0 = \psi^\dagger\gamma^0 e^{+i\theta\gamma^5} = \overline{\psi}\,e^{+i\theta\gamma^5} \\ \overline{\psi}\gamma^\mu = \psi^\dagger\gamma^0\gamma^\mu \quad &\to \quad \psi^\dagger e^{-i\theta\gamma^5}\gamma^0\gamma^\mu = \psi^\dagger\gamma^0\gamma^\mu e^{-i\theta\gamma^5} = \overline{\psi}\gamma^\mu e^{-i\theta\gamma^5}\,. \end{aligned} \tag{4.5.22}$$

Thus, $i\,\overline{\psi}\partial\!\!\!/\psi$ is invariant under this transformation but $\overline{\psi}\psi$ is not. Hence this is only a symmetry if $m=0$. To find the corresponding current, we note that the infinitesimal transformation is $\psi\to\psi+i\epsilon\gamma^5\psi$, hence the current is

$$j^\mu_A(x) = \frac{\partial\mathcal{L}(x)}{\partial\partial_\mu\psi}\,i\,\gamma^5\psi(x) = -i\overline{\psi}(x)\gamma^\mu(i\,\gamma^5\psi(x)) = \overline{\psi}(x)\gamma^\mu\gamma^5\psi(x)\,. \tag{4.5.23}$$

This is called the axial vector current or sometimes just axial current.

Since the generator of Lorentz transformations $S^{\mu\nu}$ is a product of two γ-matrices, we have $[\gamma^5, S^{\mu\nu}] = 0$. Hence, $i\,\overline{\psi}\partial\!\!\!/\gamma^5\psi$ and $\overline{\psi}\gamma^5\psi$ are Lorentz invariant. Since these are Lorentz invariant, the corresponding combinations with the projectors inserted are also Lorentz invariant,

$$\begin{aligned} i\,\overline{\psi}\partial\!\!\!/\,\tfrac{1}{2}(1-\gamma^5)\psi &= i\,\overline{\psi}_L\partial\!\!\!/\psi_L \\ i\,\overline{\psi}\partial\!\!\!/\,\tfrac{1}{2}(1+\gamma^5)\psi &= i\,\overline{\psi}_R\partial\!\!\!/\psi_R \\ \overline{\psi}\,\tfrac{1}{2}(1-\gamma^5)\psi &= \overline{\psi}\,\tfrac{1}{2}(1+\gamma^5)\psi = 0\,. \end{aligned} \tag{4.5.24}$$

Hence, if $m = 0$ it is possible to have a Lorentz invariant Lagrangian with only one Weyl fermion and thus half the number of fermion degrees of freedom. This should not come as a total shock, since we have previously seen that when $m = 0$ the left- and right-handed components decouple in the Dirac equation. We can give another argument why it is possible to have only left-handed fermions if $m = 0$ and conversely not possible if $m \neq 0$. Suppose we have a massive particle whose spin is pointing backward along its line of flight. Hence this is left-handed. Since the particle is massive, there exists a boost that brings us to the particle's rest frame. Thus, if we boost beyond this we are in an inertial frame where the particle is moving in the opposite direction from the original frame. But the spin points in the same direction, so now the particle is right-handed. If $m = 0$ there exists no boost that brings us to the particle's rest frame and hence no boost that goes beyond the rest frame such that the particle moves in the opposite direction. Thus, a left-handed particle is left-handed for any observer and so in this sense is a Lorentz invariant. A Lagrangian with only the left-handed or right-handed fermions is called a "chiral" Lagrangian. The handedness of the fermions is also called their chirality.

4.5.3 Discrete symmetries

The Dirac Lagrangian has three important discrete symmetries. The first of these symmetries is called parity (P) and transforms the space-time coordinates to $(x^0, \vec{x}) \to (x^0, -\vec{x})$. Parity satisfies $P^2 = 1$ and so $P^{-1} = P$. Note that parity is not a Lorentz transformation. To see this, recall that a Lorentz transformation matrix $\Lambda^{\mu'}{}_\nu$ has $\det\Lambda = 1$, while the parity transformation has $\det P = -1$. The Lorentz algebra generates the $SO(1,3)$ Lorentz group, where the "SO" stands for "special orthogonal". It is special because the determinant of all transformations is 1. This group can be enlarged to $O(1,3)$ by including parity and all other operations that include parity with a Lorentz transformation such that the determinant is -1.

Under parity the derivatives undergo a unitary transformation

$$\partial^\mu \to P\partial^\mu P: \qquad P\partial^0 P = \partial^0\,, \qquad P\partial^i P = -\partial^i\,. \tag{4.5.25}$$

Therefore, in order for $i\,\overline{\psi}\partial\!\!\!/\psi$ to be invariant we require that the transformation for $\psi(x)$ is

$$\psi(x) \to P\psi(x)P = \gamma^0\psi(x)\,. \tag{4.5.26}$$

Using the projectors we see that

$$\tfrac{1}{2}(1+\gamma^5)\psi \to \gamma^0\tfrac{1}{2}(1+\gamma^5)\psi = \tfrac{1}{2}(1-\gamma^5)\gamma^0\psi\,, \tag{4.5.27}$$

and so under parity we have that $\psi_L \leftrightarrow \psi_R$. Hence, the chiral Lagrangian $i\,\overline{\psi}_L\partial\!\!\!/\psi_L$ is *not* invariant under parity. A little thought should convince you that indeed a parity transformation flips left-handed to right. Under parity, the spatial component of the momentum changes sign. However, angular momentum does not change sign ($\vec{r}\times\vec{p}\to(-\vec{r})\times(-\vec{p})$) and so the spin points in the same direction. Hence the chirality of the fermions flips.

The vector current and axial current have different transformations under parity. For the vector current we have

$$\begin{aligned} j_V^0 &= \overline{\psi}\,\gamma^0\psi \to \overline{\psi}\,\gamma^0\gamma^0\gamma^0\psi = \overline{\psi}\,\gamma^0\psi \\ j_V^i &= \overline{\psi}\,\gamma^i\psi \to \overline{\psi}\,\gamma^0\gamma^i\gamma^0\psi = -\overline{\psi}\,\gamma^i\psi\,, \end{aligned} \tag{4.5.28}$$

while for the axial vector current it is the opposite,

$$\begin{aligned} j_A^0 &= \overline{\psi}\,\gamma^0\gamma^5\psi \to \overline{\psi}\,\gamma^0\gamma^0\gamma^5\gamma^0\psi = -\overline{\psi}\,\gamma^0\gamma^5\psi \\ j_A^i &= \overline{\psi}\,\gamma^i\gamma^5\psi \to \overline{\psi}\,\gamma^0\gamma^i\gamma^5\gamma^0\psi = \overline{\psi}\,\gamma^i\gamma^5\psi\,, \end{aligned} \tag{4.5.29}$$

The next discrete symmetry is C for charge conjugation. Similarly to P, $C^2=1$, $C^{-1}=C$. In the case of a complex scalar field, charge conjugation takes $\phi\to C\phi C=\phi^*$. For the fermion field we also expect it to contain a complex conjugation. However, the fermion action is not invariant under $\psi\to\psi^*$. For the mass term we would have

$$\overline{\psi}\,\psi \to \psi^T\,\gamma^0\,\psi^* = -\psi^\dagger\,\gamma^{0T}\psi = -\overline{\psi}\,\psi\,, \tag{4.5.30}$$

where we used $\gamma^{0T}=\gamma^0$ and the anti-commutativity of the Dirac fields, while for the derivative term we would get

$$i\,\overline{\psi}\partial\!\!\!/\,\psi \to i\,\psi^T\,\gamma^0\partial\!\!\!/\,\psi^* = -i\,\psi^T\,\gamma^0\,\overleftarrow{\partial\!\!\!/}\,\psi^* = i\,\psi^\dagger\partial\!\!\!/^T\gamma^{0T}\,\psi\,. \tag{4.5.31}$$

Now $\gamma^{\mu\,T}\gamma^0=+\gamma^0\gamma^\mu$ for all μ except $\mu=2$ where we have $\gamma^{2\,T}\gamma^0=-\gamma^0\gamma^2$. Hence, we can make both the mass term and the derivative term invariant if the transformation for $\psi(x)$ is

$$\psi\to C\psi(x)C = i\gamma^2\,\psi^*(x)\,. \tag{4.5.32}$$

The combination $i\,\gamma^2$ is real and symmetric and satisfies $(i\,\gamma^2)^2=1$. Moreover, it anticommutes with γ^0 and $\gamma^0\gamma^2$ but commutes with $\gamma^0\gamma^1$ and $\gamma^0\gamma^3$, so it changes the signs of the terms that needed their signs changed and leaves the other signs alone.

When we act with C on the currents we get

$$\begin{aligned} j_V^\mu &\to \psi^T(i\,\gamma^2)\gamma^0\gamma^\mu(i\gamma^2)\psi^* = -\psi^\dagger(i\,\gamma^2)\gamma^{\mu\,T}\gamma^0(i\gamma^2)\psi = -j_V^\mu \\ j_A^\mu &\to \psi^T(i\,\gamma^2)\gamma^0\gamma^\mu\gamma^5(i\gamma^2)\psi^* = -\psi^\dagger(i\,\gamma^2)\gamma^5\gamma^{\mu\,T}\gamma^0(i\gamma^2)\psi = j_A^\mu\,. \end{aligned} \tag{4.5.33}$$

where we used that $[\gamma^0\gamma^\mu, \gamma^5] = 0$, $\gamma^{5\,T} = \gamma^5$ as well as the anti-commutivity of the fields. Therefore, we find that $Q \to -Q$ under charge conjugation and thus changes all values of q to $-q$.

Under a Lorentz transformation, it follows from (4.1.22) that

$$i\,\gamma^2\psi^* \to i\,\gamma^2 e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu *}}\,\psi^* = e^{+\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}\,(i\,\gamma^2)\psi^*\,. \tag{4.5.34}$$

Hence, $(i\,\gamma^2)\psi^*$ transforms the same way under Lorentz transformations as ψ. This means that charge conjugation does not affect the spins of particles.

The chiral Lagrangians are not invariant under charge conjugation. Given the above properties of γ^5, it is straightforward to show that

$$i\,\overline{\psi}\not{\partial}\,\gamma^5\,\psi \to -i\,\overline{\psi}\not{\partial}\,\gamma^5\,\psi \tag{4.5.35}$$

under charge conjugation, and so

$$i\,\overline{\psi}_L\not{\partial}\,\psi_L \to i\;\overline{\psi}_R\not{\partial}\,\psi_R\,. \tag{4.5.36}$$

However, since P transforms the chiral Lagrangians in the same way, they are invariant under the combination CP. The chiral theories only have one spin for each fermion and anti-fermion. If the fermion is left-handed, the anti-fermion is right-handed as we argued at the end of §5.1. Hence, under charge conjugation the left-handed fermion becomes a left-handed anti-fermion and the right-handed anti-fermion becomes a right-handed fermion. If we now combine this with P then left-handed exchanges with right-handed and we are back to our original configuration.

The last discrete symmetry we discuss is time reversal (T). This symmetry is different than P and C because the transformation is antilinear. Unlike an ordinary linear operator $\mathcal{O}$ that acts on a Hilbert space as $\mathcal{O}\,c|\chi\rangle = c|\mathcal{O}\chi\rangle$ where c is any complex number, time reversal acts as $Tc|\chi\rangle = c^*T|\chi\rangle$. The reason that it must act this way is because of how the state develops in time. Assuming that the Hamiltonian H is invariant under time reversal, *i.e.* $[T, H] = 0$, then T acting on a state evolving in time is

$$T|\chi, t\rangle = |T\chi, -t\rangle = T\,e^{-iHt}|\chi\rangle = e^{iHt}|T\chi\rangle\,. \tag{4.5.37}$$

We can figure out the action of T on the Dirac field $\psi(t,\vec{x})$ by considering how it should act on the vector current $\overline{\psi}(t,\vec{x})\gamma^\mu\psi(t,\vec{x})$. The 0 component is a charge density and this should be invariant under $t \to -t$, while the spatial components are charge currents, and if we run time backward the direction of the current reverses. Hence we expect

$$\begin{aligned} T\psi(t,\vec{x})T &= B\psi(-t,\vec{x}) \\ T\overline{\psi}(t,\vec{x})\gamma^\mu\psi(t,\vec{x})T &= T\overline{\psi}(t,\vec{x})T\gamma^{\mu *}T\psi(t,\vec{x})T \\ &= \overline{\psi}(-t,\vec{x})\gamma^0 B^\dagger\gamma^0\gamma^{\mu *}B\psi(-t,\vec{x})\,, \end{aligned} \tag{4.5.38}$$

and so

$$\gamma^0 B^\dagger\gamma^0\gamma^{0*}B = \gamma^0 \qquad \gamma^0 B^\dagger\gamma^0\gamma^{i*}B = -\gamma^i\,. \tag{4.5.39}$$

Since $\gamma^{0*}=\gamma^0$, the first relation tells us that B is unitary ($B^\dagger=B^{-1}$). Since $\gamma^{1*}=\gamma^1$, $\gamma^{2*}=-\gamma^2$ and $\gamma^{3*}=\gamma^3$, we find that $B=e^{i\phi}\gamma^1\gamma^3$ where $e^{i\phi}$ is an undetermined phase. You might think that by doing two time reversals we could determine the phase, since this should be the same as doing nothing. However, this is not the case. In fact we can see that

$$T(T(\psi(t,\vec{x})T)T=T(e^{i\phi}\gamma^1\gamma^3\psi(-t,\vec{x}))T=e^{i\phi}e^{-i\phi}\gamma^1\gamma^3\gamma^1\gamma^3\phi(t,\vec{x})=-\psi(t,\vec{x})\,. \tag{4.5.40}$$

Not only can we not determine the phase, there is an extra sign that we cannot get rid of! However, all *physical* operators come with an even number of Dirac fields, so the sign will not show up in these.

The Dirac kinetic term is invariant under T, since $T(i\,\partial_0)T=(-i)(-\partial_0)=i\,\partial_0$, $T(i\,\partial_i)T=(-i)(\partial_i)=-i\,\partial_i$, so combined with our result for the current we have that $i\overline{\psi}\partial\!\!\!/\,\psi$ is invariant. Inserting γ^5 in the current to make the axial current does not change the time reversal properties since γ^5 is real and commutes with $\gamma^1\gamma^3$. Therefore the chiral Lagrangians $i\,\overline{\psi}_{L,R}\partial\!\!\!/\,\psi_{L,R}$ are also invariant under T. The same is true for the mass term $m\overline{\psi}\psi$. It seems as if everything is T invariant. However, there is one Lorentz invariant fermion bilinear term that is not time reversal invariant, $i\,\overline{\psi}\gamma^5\psi$. This term is called a *pseudoscalar* because it is invariant under continuous Lorentz transformations but changes sign under parity. The factor of i is necessary for this to be an Hermitian operator. Therefore, under time reversal we have

$$T(i\,\overline{\psi}\gamma^5\psi)T=-iT\overline{\psi}T\gamma^5T\psi T=-i\,\overline{\psi}\gamma^5\psi \tag{4.5.41}$$

where we used $[\gamma^5,B]=0$. In fact, we can also see that this operator changes sign under the combination CP,

$$CP(i\,\overline{\psi}\gamma^5\psi)PC=C(-i\,\overline{\psi}\gamma^5\psi)C=-i\,\psi^T(i\,\gamma^2)\gamma^0\gamma^5(i\,\gamma^2)\psi^*=-i\,\psi^T\gamma^0\gamma^5\psi^*=-i\,\overline{\psi}\gamma^5\psi\,. \tag{4.5.42}$$

For a long time it was thought that there was no CP violation in nature. But in the early 1960's CP violation was discovered in rare kaon decays. However, the combination CPT is believed to be a good symmetry as it is not possible to write down a Lorentz invariant Hermitian Lagrangian which violates the symmetry. Hence, CP violation implies T violation. Since T reverses the direction of the spin and the momentum, while P only reverses the direction of the momentum, under CPT, a massless fermion with helicity $h=+1/2$ transforms into a massless antifermion with helicity $h=-1/2$. CPT invariance says if you have one you must have the other.

4.6 Fermionic path integrals

We can find path integrals for fermion fields, although the antisymmetry of the fields even at the classical level requires some new machinery in order to deal with this problem.

4.6.1 Grassmann variables

A Grassmann variable θ anticommutes with all other Grassmann variables, so if we have two Grassmann variables θ_1 and θ_2, then $\theta_1\theta_2 = -\theta_2\theta_1$. Any function of a Grassmann variable $f(\theta)$ satisfies

$$f(\theta) = f(0) + \theta\, f'(0) = f_0 + \theta f_1\,, \tag{4.6.1}$$

since $\theta^2 = -\theta^2 = 0$.

We further define the *grading* of a variable, $\chi(x)$ where $\chi(x) = 0$ if x is a c-number and $\chi(x) = 1$ if x is Grassman. The grading of a product is

$$\chi(xy) = \chi(x) + \chi(y) \quad \mod 2\,. \tag{4.6.2}$$

We also have that $xy = (-1)^{\chi(x)\chi(y)}yx$. We will only need to consider functions $f(\theta)$ of a definite grading, so we will assume that either f_0 in (4.6.1) is a c-number and f_1 is Grassmann, or vice versa.

We can also construct derivatives with respect to Grassmann variables, which have the following properties:

$$\frac{\partial}{\partial\theta}\,f(\theta) = f'(0) = f_1\,, \qquad \frac{\partial}{\partial\theta_1}\frac{\partial}{\partial\theta_2} = -\frac{\partial}{\partial\theta_2}\frac{\partial}{\partial\theta_1}\,, \qquad \frac{\partial}{\partial\theta_1}\theta_2 = -\theta_2\frac{\partial}{\partial\theta_1} \tag{4.6.3}$$

Since we will be constructing path integrals we will also need to define what we mean by an integral over Grassmann variables. The key concept we want to take from ordinary integrals is integration by parts where we can throw away the integral over the total derivative. Based on the properties of the Grassmann derivative, a Grassmann integral satisfies

$$\int d\theta f(\theta)\frac{\partial}{\partial\theta}g(\theta) = -\int d\theta(-1)^{\chi(f)}\left(\frac{\partial}{\partial\theta}f(\theta)\right)g(\theta) + \int d\theta(-1)^{\chi(f)}\frac{\partial}{\partial\theta}(f(\theta)g(\theta))\,, \tag{4.6.4}$$

Taking the derivatives in the last integral we get

$$\int d\theta(-1)^{\chi(f)}\frac{\partial}{\partial\theta}(f(\theta)g(\theta)) = \int d\theta(f_0g_1 + (-1)^{\chi(f)}f_1g_0)\,. \tag{4.6.5}$$

For this to be zero for generic coefficients we must have

$$\int d\theta\cdot 1 = 0\,. \tag{4.6.6}$$

Evaluating the remaining integrals in (4.6.4) and using (4.6.6) we have

$$\int d\theta\,\theta\, f_1g_1 \;\;=\;\; -\int d\theta(-1)^{\chi(f)}f_1\theta\, g_1 = \int d\theta\,\theta\, f_1g_1 \tag{4.6.7}$$

Hence, it is consistent to choose

$$\int d\theta\,\theta = 1\,. \tag{4.6.8}$$

Equations (4.6.6) and (4.6.8) show that integration over a Grassmann variable is the same as differentiation. If we have integrals over more than one variable then we get

$$\int d\theta_1 d\theta_2\, \theta_1\theta_2 = -\int d\theta_2 d\theta_1\, \theta_1\theta_2 = -\int d\theta_1\, \theta_1 \int d\theta_2\, \theta_2 = -1\,. \tag{4.6.9}$$

The Gaussian integral will be of particular interest for us. Let us first consider a complex Grassmann made up of two Grassmann variables, $\theta = \theta_1 + i\,\theta_2$, $\overline{\theta} = \theta_1 - i\,\theta_2$. Because of the derivative-like nature of the integrals, the relation of differentials is backwards from the ordinary case,

$$d\overline{\theta}d\theta = \det\begin{pmatrix} \frac{d\theta_1}{d\overline{\theta}} & \frac{d\theta_2}{d\overline{\theta}} \\ \frac{d\theta_1}{d\theta} & \frac{d\theta_2}{d\theta} \end{pmatrix} d\theta_1 d\theta_2 = -\frac{i}{2}\, d\theta_1 d\theta_2\,. \tag{4.6.10}$$

We thus find for the Gaussian integral with the c-number b,

$$\int d\overline{\theta}d\theta\, e^{-\overline{\theta}\,b\,\theta} = \int d\overline{\theta}d\theta\,(1-\overline{\theta}\,b\,\theta) = b = \int -\frac{i}{2}\,d\theta_1 d\theta_2\, e^{-2\,i\,b\,\theta_1\theta_2}\,, \tag{4.6.11}$$

where we used in the expansion of the exponent that $(\overline{\theta}\,b\,\theta)^2 = 0$ and hence the expansion terminates. If we have N complex Grassmann variables θ_j, $\overline{\theta}_j$, then the Gaussian integral generalizes to

$$\int \prod_{j=1}^{N} d\overline{\theta}_j d\theta_j\, e^{-\overline{\theta}_j\,B_{jk}\,\theta_k} = \int \prod_{j=1}^{N} d\overline{\theta}'_j d\theta'_j\, e^{-\sum \overline{\theta}'_j\,b_j\,\theta'_j} = \prod_{j=1}^{N} b_j = \det(B) \tag{4.6.12}$$

where we used a unitary transformation to diagonalize B_{jk}, $U_{jk}B_{k\ell}U^{-1}_{\ell m} = b_j\delta_{jm}$, $\theta'_j = U_{jk}\theta_k$ and

$$\prod_{j=1}^{N} d\overline{\theta}_j d\theta_j = \det(U^{-1})\det(U)\prod_{j=1}^{N} d\overline{\theta}'_j d\theta'_j = \prod_{j=1}^{N} d\overline{\theta}'_j d\theta'_j\,. \tag{4.6.13}$$

Observe the difference between the Grassmann Gaussian integral which gives $\det(B)$ and the ordinary Gaussian integral which is proportional to $\det(B)^{-1}$.

Also note that completing the square in the Gaussian also works nicely. For example, consider the Gaussian integral with Grassmann parameters η and $\overline{\eta}$ coupled to $\overline{\theta}$ and θ,

$$\int d\overline{\theta}d\theta\, e^{-\overline{\theta}\,b\,\theta+\overline{\eta}\,\theta+\overline{\theta}\,\eta} = \int d\overline{\theta}d\theta(1-\overline{\theta}\,b\,\theta+\overline{\eta}\,\theta\,\overline{\theta}\eta) = b+\overline{\eta}\,\eta = b\,e^{\overline{\eta}\,b^{-1}\,\eta}\,. \tag{4.6.14}$$

However, we can also express this as

$$\int d\overline{\theta}d\theta\, e^{-\overline{\theta}\,b\,\theta+\overline{\eta}\,\theta+\overline{\theta}\,\eta} = \int d\overline{\theta}d\theta\, e^{-(\overline{\theta}-\overline{\eta}\,b^{-1})\,b\,(\theta-b^{-1}\eta)}\, e^{\overline{\eta}\,b^{-1}\,\eta}\,. \tag{4.6.15}$$

Hence the result in (4.6.14) is consistent with shifting the integration variables in (4.6.15).

Finally, we can also define a δ-function, $\delta(\theta-\theta')$, where

$$\int d\theta\,\delta(\theta-\theta')\,f(\theta) = f(\theta')\,. \tag{4.6.16}$$

A quick calculation shows that (4.6.16) is satisfied if $\delta(\theta-\theta') = \theta-\theta'$.

4.6.2 The 0+1 dimensional fermionic oscillator

In $0+1$ dimensions we can have a single fermionic oscillator with commutation relations

$$\{d, d^\dagger\} = 1\,, \qquad \{d, d\} = \{d^\dagger, d^\dagger\} = 0\,. \tag{4.6.17}$$

Hence, there are only two states, $|0\rangle$ and $d^\dagger|0\rangle$. If we assume the Hamiltonian is $H = m_0\, d^\dagger d$, then the energies of the two states are 0 and m_0. Let us now find the fermionic path integral that is consistent with this system.

We write our fermionic fields as

$$\widehat{\psi}(t) \;=\; e^{-im_0 t}\, d \qquad\qquad \widehat{\overline{\psi}}(t) = e^{+im_0 t}\, d^\dagger\,, \tag{4.6.18}$$

where we put the hats over the fields to emphasize that these are quantum operators. We can write down ket states which are eigenstates of these field operators satisfying

$$\widehat{\psi}|\psi\rangle = \psi|\psi\rangle \qquad\qquad \widehat{\overline{\psi}}|\overline{\psi}\rangle = \overline{\psi}|\overline{\psi}\rangle\,, \tag{4.6.19}$$

where $\widehat{\psi} \equiv \widehat{\psi}(0)$ and the fields without the hats are ordinary Grassmann variables which anticommute with the fermionic operators. A straightforward calculation shows that

$$|\psi\rangle \equiv |0\rangle - \psi\, d^\dagger|0\rangle \qquad\qquad |\overline{\psi}\rangle \equiv \overline{\psi}\,|0\rangle - d^\dagger|0\rangle \tag{4.6.20}$$

satisfies the eigenvalue equations. The bra states are defined so that they satisfy the eigenvalue equations

$$\langle\psi|\widehat{\psi} = \langle\psi|\psi \qquad\qquad \langle\overline{\psi}|\widehat{\overline{\psi}} = \langle\overline{\psi}|\overline{\psi} \tag{4.6.21}$$

for which we find the consistent solutions

$$\langle\psi| \equiv \langle 0|\psi - \langle 0|d \qquad\qquad \langle\overline{\psi}| \equiv \langle 0| - \langle 0|d\,\overline{\psi}\,. \tag{4.6.22}$$

Notice that $|\psi\rangle$ is bosonic (assuming $|0\rangle$ is bosonic) while $\langle\psi|$ is fermionic. The situation is reversed for $|\overline{\psi}\rangle$ and $\langle\overline{\psi}|$. We also see that $\langle\overline{\psi}|$ is the adjoint of $|\psi\rangle$ while $\langle\psi|$ is the adjoint of $|\overline{\psi}\rangle$.

Given these states we can find relations similar to those in ordinary quantum mechanics,

$$\begin{aligned}
\langle\psi|\psi'\rangle &= \psi - \psi' = \delta(\psi - \psi') &\qquad \langle\overline{\psi}|\overline{\psi}'\rangle &= \overline{\psi}' - \overline{\psi} = \delta(\overline{\psi}' - \overline{\psi})\\
\langle\psi|\overline{\psi}\rangle &= 1 - \overline{\psi}\psi = e^{-\overline{\psi}\psi} &\qquad \langle\overline{\psi}|\psi\rangle &= 1 + \overline{\psi}\psi = e^{\overline{\psi}\psi}\,.
\end{aligned} \tag{4.6.23}$$

These last two relations are the analogs of $\langle x|p\rangle = e^{ipx}$, $\langle p|x\rangle = e^{-ipx}$ (remember that $i\overline{\psi}$ is the canonical momentum for ψ). Using the rules for Grassmann integration we can also write down the complete set of states,

$$\int d\psi|\psi\rangle\langle\psi| = \int d\overline{\psi}|\overline{\psi}\rangle\langle\overline{\psi}| = |0\rangle\langle 0| + d^\dagger|0\rangle\langle 0|d = \mathbb{1} \tag{4.6.24}$$

We can now ask what is the amplitude for the state $|0\rangle$ at $t=-T/2$ to evolve to the state $|0\rangle$ at time $t=+T/2$. Since the Hamiltonian is $H=m_0\widehat{\overline{\psi}}\widehat{\psi}$, we find for the amplitude

$$\langle 0|e^{-iTH}|0\rangle=\lim_{\Delta t\to 0}\langle 0|\prod_t e^{-i\Delta t H}|0\rangle\,, \tag{4.6.25}$$

Between each time step we now insert two complete sets of states,

$$\int d\psi(t)\,|\psi(t)\rangle\langle\psi(t)|\int d\overline{\psi}(t)\,|\overline{\psi}(t)\rangle\langle\overline{\psi}(t)|=\int d\overline{\psi}(t)d\psi(t)\,|\psi(t)\rangle\langle\overline{\psi}(t)|\,e^{-\overline{\psi}(t)\psi(t)}\,. \tag{4.6.26}$$

We then use that

$$\begin{aligned}\langle\overline{\psi}(t)|e^{-im_0\widehat{\overline{\psi}}\widehat{\psi}\Delta t}|\psi(t-\Delta t)\rangle &\approx \langle\overline{\psi}(t)|(1-im_0\widehat{\overline{\psi}}\widehat{\psi}\Delta t)|\psi(t-\Delta t))\rangle\\ &= \langle\overline{\psi}(t)|\psi(t-\Delta t)\rangle(1-im_0\overline{\psi}(t)\psi(t-\Delta t)\Delta t)\\ &\approx e^{\overline{\psi}(t)\psi(t-\Delta t)}e^{-im_0\overline{\psi}(t)\psi(t-\Delta t)\Delta t}\end{aligned} \tag{4.6.27}$$

Putting everything together we find the fermionic path integral

$$\begin{aligned}\mathcal{Z}=\lim_{T\to\infty}\langle 0|e^{-iTH}|0\rangle &= \lim_{\substack{T\to\infty\\ \Delta t\to 0}}\int\prod_{t=-T/2}^{T/2}\int d\overline{\psi}(t)d\psi(t)\,e^{-\overline{\psi}(-T/2)\psi(-T/2)}\\ &\qquad\times\exp\left(\sum_{t=-T/2+\Delta t}^{T/2}\left[-\overline{\psi}(t)\left(\psi(t)-\psi(t-\Delta t)\right)-im_0\overline{\psi}(t)\psi(t-\Delta t)\Delta t\right]\right)\\ &= \int\mathcal{D}\overline{\psi}\,\mathcal{D}\psi\exp\left(i\int dt\,\overline{\psi}(t)(i\,\partial_0-m_0+i\epsilon)\psi(t)\right)\\ &= \int\mathcal{D}\overline{\psi}\,\mathcal{D}\psi\,e^{i\int dt\mathcal{L}(t)}\,.\end{aligned} \tag{4.6.28}$$

where $\mathcal{D}\overline{\psi}$ and $\mathcal{D}\psi$ signifies the functional integral for the Grassmann variables $\overline{\psi}(t)$ and $\psi(t)$. The $i\,\epsilon$ appears because the path integral is a limit over a very large time.

To find the time ordered correlators we do as in the bosonic case and introduce sources. The sources are the Grassmann variables $\eta(t)$ and $\overline{\eta}(t)$ and the path integral's dependence on them is

$$\mathcal{Z}(\eta,\overline{\eta})=\int\mathcal{D}\overline{\psi}\,\mathcal{D}\psi\exp\left(i\int dt\left(\overline{\psi}(t)(i\,\partial_0-m_0+i\epsilon)\psi(t)+\overline{\eta}(t)\psi(t)+\overline{\psi}(t)\eta(t)\right)\right) \tag{4.6.29}$$

We insert a field operator[8] $\psi(t)$ inside the inner product by taking a functional derivative, $-i\,\dfrac{\delta}{\delta\overline{\eta}(t)}$, while for $\overline{\psi}(t)$ we use $i\,\dfrac{\delta}{\delta\eta(t)}$. Therefore, the time ordered correlators are

$$\begin{aligned}&\langle T[\psi(t_1)\overline{\psi}(t_2)\psi(t_3)\overline{\psi}(t_4)\dots\psi(t_{2n-1})\overline{\psi}(t_{2n})]\rangle\\ &\qquad=\prod_{j=1}^{n}\left(-i\,\frac{\delta}{\delta\overline{\eta}(t_{2j-1})}\right)\left(+i\,\frac{\delta}{\delta\eta(t_{2j})}\right)Z(\eta,\overline{\eta})\Big|_{\eta=\overline{\eta}=0}\,.\end{aligned} \tag{4.6.30}$$

[8]We drop the hats on the operators since the context should be clear by now.

As in the bosonic case the correlator automatically comes out time ordered. Furthermore, there is the appropriate sign change as in (4.3.3) because of the antisymmetry of the source derivatives.

The path integral can be evaluated by completing the square and shifting variables, where we find

$$\begin{aligned}
\mathcal{Z}(\eta,\overline{\eta}) &= \int \mathcal{D}\overline{\psi}\,\mathcal{D}\psi\, \exp\left(i\int dt\,(\overline{\psi}(t)-i\overline{\eta}S_F(t))(i\,\partial_0-m_0+i\epsilon)(\psi(t)-i\,S_F\eta(t))\right)\\
&\qquad\qquad\times\exp\left(-\int dt\,\overline{\eta}(t)\,S_F\,\eta(t)\right)\\
&= \mathcal{Z}(0,0)\,\exp\left(-\int dt\,\overline{\eta}(t)\,S_F\,\eta(t)\right)\,,
\end{aligned} \tag{4.6.31}$$

where $S_F(t-t')$ is the analog of the Dirac propagator in 3+1 dimensions,

$$\begin{aligned}
(i\,\partial_0-m_0)S_F(t-t') = i\,\delta(t-t') \quad\Rightarrow\quad S_F(t-t') &= e^{-im_0(t-t')} \qquad t>t'\\
&= 0 \qquad\qquad\quad t<t'\,,
\end{aligned} \tag{4.6.32}$$

and where

$$S_F\,\eta(t) = \int dt'\,S_F(t-t')\,\eta(t') \qquad \overline{\eta}\,S_F(t) = \int dt'\,\overline{\eta}(t')\,S_F(t'-t)\,. \tag{4.6.33}$$

Given the form of $\mathcal{Z}(\eta,\overline{\eta})$ we see that Wick's theorem applies and that all $2n$-point correlators reduce to a sum over products of n two-point correlators $S_F(t-t')$. The number of terms in the sum is determined by the different ways to pair up the $\psi(t_{2j-1})$ with the $\overline{\psi}(t_{2j})$. Hence the normalized correlator is given by

$$\begin{aligned}
&\frac{\langle T[\psi(t_1)\overline{\psi}(t_2)\psi(t_3)\overline{\psi}(t_4)\ldots\psi(t_{2n-1})\overline{\psi}(t_{2n})]\rangle}{\mathcal{Z}(0,0)}\\
&\quad= \prod_{j=1}^{n}\left(-i\,\frac{\delta}{\delta\overline{\eta}(t_{2j-1})}\right)\left(+i\,\frac{\delta}{\delta\eta(t_{2j})}\right)\exp\left(-\int dtdt'\,\overline{\eta}(t)\,S_F(t-t')\,\eta(t')\right)\Bigg|_{\eta=\overline{\eta}=0}\\
&\quad= \sum_{\sigma\in\mathcal{S}_n}(-1)^{[\sigma]}\prod_{j=1}^{n}S_F(t_{2j-1}-t_{2\sigma(j)})\,,
\end{aligned} \tag{4.6.34}$$

where $\mathcal{S}_n$ is the permutation group on n elements, σ is a particular element in the group, $(-1)^{[\sigma]}=1$ (-1) for an even (odd) permutation and $\sigma(j)$ is the result of the permutation on the j^{th} element.

The signs $(-1)^{[\sigma]}$ are crucial and must be handled with care in perturbation theory. In particular, we will have cause to consider loop diagrams which have correlators of the form

$$\langle T[:\overline{\psi}(t)\psi(t)::\overline{\psi}(t')\psi(t'):]\rangle = \mathcal{Z}(0)(-1)\mathrm{Tr}[(S_F(t-t'))^2] \tag{4.6.35}$$

This factor of -1 always appears when there is a fermion loop.

4.6.3 3+1 dimensional Dirac fermions

The setup here is almost the same as in $0+1$ dimensions, with the added complication that the fermion fields also have a Dirac index. We now have the functional differentials

$$\mathcal{D}\overline{\psi}\mathcal{D}\psi = \prod_x \prod_{\alpha=1}^{4} d\overline{\psi}_\alpha(x)\, d\psi_\alpha(x)\,. \tag{4.6.36}$$

where here $\overline{\psi}_\alpha(x) = \psi^\dagger(x)_\beta \gamma^0_{\beta\alpha}$. The functional integral is now

$$\mathcal{Z}(\eta,\overline{\eta}) = \int \mathcal{D}\overline{\psi}\,\mathcal{D}\psi \exp\left(i\int d^4x\left(\overline{\psi}(x)(i\partial\!\!\!/ - m + i\,\epsilon)\psi(x) + \overline{\eta}(x)\psi(x) + \overline{\psi}(x)\eta(x)\right)\right)\,. \tag{4.6.37}$$

where the sources also have Dirac indices. Completing the square and shifting integration variables, we then find

$$\begin{aligned}\mathcal{Z}(\eta,\overline{\eta}) &= \int \mathcal{D}\overline{\psi}\,\mathcal{D}\psi\, \exp\left(i\int d^4x\,(\overline{\psi}(x) - i\overline{\eta}S_F(x))(i\partial\!\!\!/ - m + i\,\epsilon)(\psi(x) - i\,S_F\eta(x))\right) \\ &\qquad\qquad \times \exp\left(-\int d^4x\,\overline{\eta}(x)\,S_F\,\eta(x)\right) \\ &= \mathcal{Z}(0,0)\, \exp\left(-\int d^4x d^4y\,\overline{\eta}(x)\,S_F(x-y)\,\eta(y)\right)\,,\end{aligned} \tag{4.6.38}$$

where $S_F(x)$ is given by (4.3.4) and

$$S_F\,\eta(x) = \int d^4y\, S_F(x-y)\,\eta(y) \qquad \overline{\eta}\,S_F(x) = \int d^4y\,\overline{\eta}(y)\,S_F(y-x)\,. \tag{4.6.39}$$

The normalized correlator is then

$$\begin{aligned}&\frac{\langle T[\psi_{\alpha_1}(x_1)\overline{\psi}_{\beta_1}(y_1)\psi_{\alpha_2}(x_2)\overline{\psi}_{\beta_2}(y_2)\dots\psi_{\alpha_n}(x_n)\overline{\psi}_{\beta_n}(y_n)]\rangle}{\mathcal{Z}(0,0)} \\ &\qquad = \prod_{j=1}^{n}\left(-i\,\frac{\delta}{\delta\overline{\eta}_{\alpha_j}(x_j)}\right)\left(+i\,\frac{\delta}{\delta\eta_{\beta_j}(y_j)}\right)\, \exp\left(-\int d^4xd^4y\,\overline{\eta}(x)\,S_F(x-y)\,\eta(y)\right) \\ &\qquad = \sum_{\sigma\in\mathcal{S}_n}(-1)^{[\sigma]}\prod_{j=1}^{n} S_F(x_j - y_{\sigma(j)})_{\alpha_j\beta_{\sigma(j)}}\,.\end{aligned} \tag{4.6.40}$$

Chapter 5

Scattering

In this chapter of the notes we discuss particle scattering. In particular, we will show that it is closely tied to the correlators we have computed in earlier lectures.

5.1 The S-matrix

A scattering problem assumes that we have an initial state in the far past with two widely separated particles, which scatters during some middle time period, and evolves into a state in the far future with two or more widely separated particles. The time development of this process is determined by the full interacting Hamiltonian, H. Given the incoming state, we wish to know the probability amplitude of finding the outgoing state.

When the particles are widely separated their interactions with each other are negligible. However, the particles still interact with the surrounding field leading to their mass-shifts and wave-function renormalizations. We call such particles "dressed". Let us suppose there is a Hamiltonian $H_{\rm NI}$ where the dressed particles do not interact with each other (we will explicitly demonstrate the existence of this Hamiltonian). Now let $|\varphi\rangle$ and $|\psi\rangle$ be two eigenstates of $H_{\rm NI}$, where each corresponds to a set of dressed particles with fixed momenta (as well as internal quantum numbers). Assuming the momentum and/or quantum numbers are different, we have that $\langle\psi|\varphi\rangle = 0$. In fact, if we were to consider the states at separate times then the inner product is still zero, assuming that $H_{\rm NI}$ is the time evolution operator.

But of course, $H_{\rm NI}$ is not the time evolution operator, H is. Given a state $|\Psi, -T\rangle$ at $t = -T$ it evolves unitarily to a state $|\Psi, +T\rangle$. Since we are assuming that the state starts widely separated in the far past, we can assume that Ψ equals one of the $H_{\rm NI}$ eigenstates, say φ, in the distant past. Thus $|\Psi, -T\rangle = |\varphi\rangle$. The state then evolves to the state in the far future

$$|\Psi, +T\rangle = |\varphi\rangle_{\rm in}\,, \tag{5.1.1}$$

where the subscript "in" lets us know that this was the state that started out looking like φ. But we also expect that the state in the far future looks like a sum over eigenstates of $H_{\rm NI}$, ψ_i,

$$|\Psi,+T\rangle=\sum_i c_i|\psi_i\rangle_{\rm out}\,,\tag{5.1.2}$$

since the states are assumed to be widely separated in the far future. The "out" indicates that these are states that evolve to ψ_i in the far future. Since this is the noninteracting state in the far future, its relation to the noninteracting state in the far past is $|\psi_i\rangle_{\rm out}=e^{-2iH_{\rm NI}T}|\psi_i\rangle$. Therefore, the probability amplitude to evolve from the widely separated state φ to the widely separated state ψ is

$${}_{\rm out}\langle\psi|\varphi\rangle_{\rm in}=\langle\psi|e^{2iH_{\rm NI}T}e^{-2iHT}|\varphi\rangle\;\equiv\langle\psi|S|\varphi\rangle\,.\tag{5.1.3}$$

The operator S is called the S-matrix which tells us how incoming states connect to outgoing states. The S-matrix is clearly unitary, satisfying $SS^\dagger=1$. The states $|\phi\rangle$ and $|\psi\rangle$ are the states at $t=-T$, but we can make the expression in (5.1.3) more symmetric by evolving the states to $t=0$ with the noninteracting Hamiltonian. If we define $|\phi\rangle_0=e^{-iH_{\rm NI}T}|\phi\rangle$ and $|\psi\rangle_0=e^{-iH_{\rm NI}T}|\psi\rangle$ then the probability amplitude is

$${}_{\rm out}\langle\psi|\varphi\rangle_{\rm in}={}_0\langle\psi|e^{iH_{\rm NI}T}e^{-2iHT}e^{iH_{\rm NI}T}|\phi\rangle_0={}_0\langle\psi|e^{-iH_{\rm NI}T}Se^{iH_{\rm NI}T}|\phi\rangle_0\,.\tag{5.1.4}$$

We then redefine the S-matrix to be its unitary transformation, $e^{-iH_{\rm NI}T}Se^{iH_{\rm NI}T}$, and further drop the 0 subscripts to write

$${}_{\rm out}\langle\psi|\varphi\rangle_{\rm in}=\langle\psi|e^{iH_{\rm NI}T}e^{-2iHT}e^{iH_{\rm NI}T}|\phi\rangle=\langle\psi|S|\phi\rangle\tag{5.1.5}$$

The S-matrix is usually written as

$$S=1+i\,\mathcal{T}\tag{5.1.6}$$

where the 1 corresponds to the "forward" part where there is no scattering (the outgoing state is the same as the ingoing evolved with H_{NI}), and $\mathcal{T}$ is called the T-matrix. If $\langle\psi|\phi\rangle=0$ then

$$\langle\psi|S|\varphi\rangle=i\langle\psi|\mathcal{T}|\varphi\rangle\,.\tag{5.1.7}$$

The unitarity of S directly leads to the condition

$$2\,{\rm Im}(\mathcal{T})=\mathcal{T}^\dagger\mathcal{T}\,.\tag{5.1.8}$$

It turns out that this relation is closely related to the optical theorem.

It is actually not convenient to normalize the states to unity. Leaving the states unnormalized, the probability for a transition from φ to ψ is (assuming $\langle\psi|\varphi\rangle=0$)

$$P_{\varphi\psi}=\frac{|\langle\psi|\mathcal{T}|\varphi\rangle|^2}{\langle\psi|\psi\rangle\langle\varphi|\varphi\rangle}\,.\tag{5.1.9}$$

The relation in (5.1.8) is also understood to be

$$\langle\psi|2{\rm Im}(\mathcal{T})|\varphi\rangle=\sum_a\langle\psi|\mathcal{T}^\dagger|\psi_a\rangle\frac{1}{\langle\psi_a|\psi_a\rangle}\langle\psi_a|\mathcal{T}|\varphi\rangle\,,\tag{5.1.10}$$

where the sum is over all states.

5.2 The LSZ reduction formula

We now show how to relate the S-matrix between two widely separated states to the correlators we have spent so much time studying. For definiteness, let us assume that we have a real scalar field with a ϕ^4 interaction. We further assume that the incoming state is given by

$$|\varphi\rangle = |\vec{k}_1, \vec{k}_2 \dots \vec{k}_n\rangle \tag{5.2.1}$$

while the outgoing state is

$$|\psi\rangle = |\vec{p}_1, \vec{p}_2 \dots \vec{p}_m\rangle \,. \tag{5.2.2}$$

For a one particle dressed state we have that up to a phase,

$$\langle 0|\phi_R(x)|\vec{k}\rangle = e^{-ik\cdot x} \tag{5.2.3}$$

where $\phi_R(x)$ is the renormalized scalar field. The $e^{-ik\cdot x}$ factor results from the translation symmetry while the overall normalization follows from the definition of $\phi_R(x)$ and the discussion around (3.2.42 – 3.2.45) of the third chapter of the notes. To see why translation symmetry implies the above form, we observe that

$$\langle 0|\phi_R(x+a)|\vec{k}\rangle = \langle 0|e^{ia\cdot P}\phi_R(x)e^{-ia\cdot P}|\vec{k}\rangle = \langle 0|1\cdot\phi_R(x)e^{-ia\cdot k}|\vec{k}\rangle = \langle 0|\phi_R(x)|\vec{k}\rangle\, e^{-ia\cdot k}\,, \tag{5.2.4}$$

where P^μ is the momentum operator, the generator of translations.

Notice that the relation in (5.2.2) is the same found for the free particle case,

$$\langle 0|\phi(x)|\vec{k}\rangle_{\text{free}} = e^{-ik\cdot x}\,. \tag{5.2.5}$$

In fact, for free particles we can extend this relation to

$$\begin{aligned}\langle \vec{p}_1, \dots \vec{p}_m| :\phi(x_1)\dots\phi(x_n)\phi(x_{n+1})\dots\phi(x_{n+m}): |\vec{k}_1, \dots \vec{k}_n\rangle_{\text{free}}& \\ = \sum_{\text{perms}} \exp\left(-i\sum_{j=1}^{n+m} k_j\cdot x_{\sigma(j)}\right)&\end{aligned} \tag{5.2.6}$$

where the fields are normal ordered, σ is an element of the permutation group on $n+m$ objects, and $k_{n+j} \equiv -p_j$. If $k_i \neq -k_j$ for all i and j then we can drop the normal ordering in (5.2.6). We also have the useful relation

$$\begin{aligned} e^{-ik\cdot x} &= \int d^4y \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x-y)} \frac{i}{q^2-m^2+i\epsilon}\,\frac{k^2-m^2+i\epsilon}{i}\, e^{-ik\cdot y} \\ &= \int d^4y\, G_F(x-y)\,\frac{k^2-m^2+i\epsilon}{i}\, e^{-ik\cdot y}\,. \end{aligned} \tag{5.2.7}$$

It is then a straightforward exercise to show using Wick's theorem that

$$\sum_{\text{perms}} \exp\left(-i\sum_{j=1}^{n+m} k_j\cdot x_{\sigma(j)}\right)$$
$$= \int\prod_{j=1}^{n+m}\left(d^4y_j\,\frac{k_j^2-m^2+i\epsilon}{i}\,e^{-ik_j\cdot y_j}\right)\langle T\left[\phi(y_1)\ldots\phi(y_{n+m})\ :\phi(x_1)\ldots\phi(x_{n+m})\ :\right]\rangle_{\text{free}}\,. \tag{5.2.8}$$

Again, if $k_i\neq -k_j$ for all i and j then we can drop the normal ordering since $\phi(y_j)$ must be contracted with the $\phi(x_k)$ fields to give a nonzero result.

We next note that the middle part of (5.1.5) is very similar to quantities we have previously considered. For example, we showed that

$$\langle 0|e^{-2iHT}|0\rangle = \langle 0|T\left[\exp\left(i\int d^4x\mathcal{L}_I\right)\right]|0\rangle_{\text{free}}\,, \tag{5.2.9}$$

where we had assumed that $H_0|0\rangle = 0$. We derived this relation using a path integral argument, but here we present a different argument that makes a generalization to other states besides the vacuum more transparent.

In quantum mechanics we are familiar with two different "pictures" for describing states and operators. In the Schrödinger picture, all time dependence is included in the states while the operators have no time dependence, unless the time dependence is explicit. Hence, we consider inner products of the form $\langle\psi,t|\mathcal{O}|\varphi,t\rangle$. In the Heisenberg picture there is no time dependence in the states, all time dependence is in the operators so that one has inner products of the form $\langle\psi|\mathcal{O}(t)|\varphi\rangle$. Of course, the two pictures are equivalent, since

$$\langle\psi|\mathcal{O}(t)|\varphi\rangle = \langle\psi|e^{iHt}\mathcal{O}e^{-iHt}|\varphi\rangle = \langle\psi,t|\mathcal{O}|\varphi,t\rangle\,. \tag{5.2.10}$$

What we have been doing up to now is a hybrid of the Schrödinger and Heisenberg pictures called the "interaction picture". In the interaction picture we put the time dependence into the operators, but only using the free Hamiltonian. In other words, the time dependence of an operator is defined to be $\mathcal{O}(t)\equiv e^{iH_0t}\mathcal{O}e^{-iH_0t}$. Let us apply this to the operator $e^{iH_0T}e^{-2iHT}e^{iH_0T}$. We first reëxpress this operator as

$$e^{iH_0T}e^{-2iHT}e^{iH_0T} = \lim_{\Delta t\to 0} e^{iH_0T}\left(e^{-iH_0\Delta t}e^{-iH_I\Delta t}\right)^{\frac{2T}{\Delta t}}e^{iH_0T}\,, \tag{5.2.11}$$

where $H=H_0+H_I$. We then rewrite every $e^{-iH_0\Delta t}$ as

$$e^{-iH_0\Delta t} = e^{-iH_0(t+\Delta t)}e^{iH_0t}\,, \tag{5.2.12}$$

where the t is determined by the position of the particular $e^{-iH_0\Delta t}$ in the product. Hence, we can rewrite (5.2.11) as

$$\begin{aligned} e^{iH_0T}e^{-2iHT}e^{iH_0T} &= \lim_{\Delta t\to 0} T\left[\prod_t\left(e^{+iH_0t}e^{-iH_I\Delta t}e^{-iH_0t}\right)\right] \\ &= \lim_{\Delta t\to 0} T\left[\prod_t e^{-iH_I(t)\Delta t}\right] = T\left[e^{-i\int dtH_I(t)}\right] = T\left[\exp\left(i\int d^4x\mathcal{L}_I\right)\right]\,, \end{aligned} \tag{5.2.13}$$

where the T in front of the square bracket is the time ordering symbol. Hence, we see that for *any* states we have that

$$\langle\psi|e^{iH_0T}e^{-2iTH}e^{iH_0T}|\varphi\rangle = \langle\psi|T\left[\exp\left(i\int d^4x\mathcal{L}_I(x)\right)\right]|\varphi\rangle\,, \tag{5.2.14}$$

where the time evolution of $\mathcal{L}_I(x)$ is understood to be determined by H_0.

We next claim that the noninteracting dressed states *are* the free particle states, provided that all counterterms are included in $\mathcal{L}_I$. To see this let us review the counterterm prescription. In the $0+1$ dimensional case we did not add counterterms (although we could have) since all perturbative corrections are finite. In this case, we are given a Lagrangian and then we find the physical masses and couplings. We do not make any changes to the Lagrangian in the process. Because no changes are made to the Lagrangian, the starting field $\phi(x)$ is the bare field $\phi_0(x)$ since it it is the field that appears with a canonical kinetic term in the Lagrangian. The renormalized field is $\phi_R(x) = Z^{-1/2}\phi(x)$, where Z is the residue of the pole at the physical mass.

In the $3+1$ case we worked backwards. We are given physical masses and couplings and we find the Lagrangian that gives us these physical masses and couplings. We do this by adding counterterms to the Lagrangian, after which it has the form

$$\mathcal{L}(x) = \int d^4x\,\tfrac{1}{2}(1+\delta_Z)\left(\partial_\mu\phi(x)\partial^\mu\phi(x) - (m^2+\delta_{m^2})\phi^2(x)\right) - \tfrac{1}{4!}(\lambda+\delta_\lambda)\phi^4(x)\,. \tag{5.2.15}$$

As you can see $\phi(x)$ is *not* the bare field. Instead, the bare field is $\phi_0(x) = Z^{1/2}\phi(x)$, $Z = 1+\delta_Z$, and the renormalized field is $\phi_R(x) = \phi(x)$. In fact, if you recall how the counterterms were chosen, they were done so that pole and residue for $\int d^4x\,e^{-ik\cdot x}\langle T[\phi(x)\phi(0)]\rangle$ stayed at m^2 and 1. In other words, $\phi(x)$ is the renormalized field.

With the counterterms included in the interaction part of the Lagrangian, turning off the interaction between the dressed particles also means turning off the counterterms, even though δ_Z and δ_{m^2} only contribute to the free field part of the Lagrangian. Removing the counterterms is the way the separated particles remain interacting with the field. Otherwise, the field that creates the normalized free particle states would be $\phi_0(x)$ and the particles would have mass squares of $m^2+\delta_{m^2}$.

Thus, it seems that all we have to do is replace $H_{\rm NI}$ in (5.1.5) with H_0 and then we are done. This is almost right, but it does not take into account the change in the vacuum energy due to the interaction. This change shifts the energy of all states by the same amount and leads to an extra phase that must be compensated for. To find the phase, we can consider the S-matrix for the vacuum. The vacuum is unique with an energy gap separating it from its lowest excited states. Therefore, $\langle 0|S|0\rangle = 1$. However, if we use (5.1.5) with $H_{\rm NI}$ replaced by H_0 we get

$$\langle 0|S|0\rangle \stackrel{?}{=} \langle 0|e^{iH_0T}e^{-2iHT}e^{iH_0T}|0\rangle = \langle 0|T\left[\exp\left(i\int d^4x\mathcal{L}_I(x)\right)\right]|0\rangle_{\rm free} \tag{5.2.16}$$

The right hand side is an extra phase that is common to all S-matrix elements with H_{NI} replaced with H_0. We can get rid of it by dividing all elements by the phase. Hence, the

S-matrix between the particle states is

$$\langle\vec{p}_1,\ldots\vec{p}_m|S|\vec{k}_1,\ldots\vec{k}_n\rangle = \frac{\langle\vec{p}_1,\ldots\vec{p}_m|T\left[\exp\left(i\int d^4x\mathcal{L}_I(x)\right)\right]|\vec{k}_1,\ldots\vec{k}_n\rangle_{\rm free}}{\langle 0|T\left[\exp\left(i\int d^4x\mathcal{L}_I(x)\right)\right]|0\rangle_{\rm free}}\,, \qquad (5.2.17)$$

where $\mathcal{L}_I$ contains the counterterms.

To complete the evalution of the S-matrix we need a slightly more general form of Wick's theorem, which we can more or less argue from the form that we had before. Recall from chapter 2 of the lecture notes that

$$\langle T[\phi(x_1)\phi(x_2)\ldots\phi(x_n)]\rangle_{\rm free} = \sum_{\rm pairs}\prod_{k=1}^{n/2}\langle T[\phi(x_{I(k)})\phi(x_{J(k)})]\rangle_{\rm free}\,. \qquad (5.2.18)$$

Let us now apply this to

$$\langle :\phi^N(y): T[\phi(x_1)\phi(x_2)\ldots\phi(x_n)] :\phi^M(z):\rangle_{\rm free}\,, \qquad (5.2.19)$$

where we assume that $y^0 > x_j^0 > z^0$ so that the correlator is time ordered. The fields at y and z are normal ordered so there are no Wick contractions between any two fields at y or any two fields at z. We will use the Wick bracket notation $\ldots\phi(x_j)\ldots\phi(x_k)\ldots$ to indicate the Wick contraction between the fields $\phi(x_j)$ and $\phi(x_k)$. Hence we have that

$$\begin{aligned}
&\langle :\phi^N(y): T[\phi(x_1)\phi(x_2)\ldots\phi(x_n)] :\phi^M(z):\rangle_{\rm free} = \\
&\quad\langle :\phi^N(y)::\phi(x_1)\phi(x_2)\ldots\phi(x_n)::\phi^M(z):\rangle_{\rm free} \\
&\quad+\sum_{j<k}\langle :\phi^N(y)::\phi(x_1)\ldots\phi(x_j)\ldots\phi(x_k)\ldots\phi(x_n)::\phi^M(z):\rangle_{\rm free} \\
&\quad+\sum_{\substack{j_1<k_1,\,j_2<k_2\\ j_1<j_2,k_1\neq j_2,k_2}}\langle :\phi^N(y)::\phi(x_1)\ldots\phi(x_{j_1})\ldots\phi(x_{j_2})\ldots\phi(x_{k_1})\ldots\phi(x_{k_2})\ldots\phi(x_n)::\phi^M(z):\rangle_{\rm free} \\
&\qquad+\ \ldots\ldots \qquad (5.2.20)
\end{aligned}$$

where the fields inside a pair of colons can only be contracted with fields outside the colons. We can now see that the structure in (5.2.20) is independent of the $:\phi^N(x):$ and $:\phi^M(y):$, so we drop them and write Wick's theorem as the operator statement for free scalar fields

$$\begin{aligned}
T[\phi(x_1)\phi(x_2)\ldots\phi(x_n)] =& :\phi(x_1)\phi(x_2)\ldots\phi(x_n): + \sum_{j<k}:\phi(x_1)\ldots\phi(x_j)\ldots\phi(x_k)\ldots\phi(x_n): \\
&+\sum_{\substack{j_1<k_1,\,j_2<k_2\\ j_1<j_2,k_1\neq j_2,k_2}}:\phi(x_1)\ldots\phi(x_{j_1})\ldots\phi(x_{j_2})\ldots\phi(x_{k_1})\ldots\phi(x_{k_2})\ldots\phi(x_n): +\ \ldots\ldots \qquad (5.2.21)
\end{aligned}$$

where the last set of dots contains three or more Wick contractions.

We now have all of the ingredients to write down the S-matrix. With very little loss of generality, we assume that $\vec{p}_i \neq \vec{k}_j$, in other words, all particles have at least some deflection in the scattering. This means that

$$\langle \vec{p}_1, \ldots \vec{p}_m | : \phi(x_1) \ldots \phi(x_N) : | \vec{k}_1, \ldots \vec{k}_n \rangle_{\text{free}} = 0\,, \tag{5.2.22}$$

unless $N = m + n$. We then expand $T\left[\exp\left(i\int d^4x \mathcal{L}_I(x)\right)\right]$ and use Wick's theorem in (5.2.21), keeping only those terms with $n + m$ normal ordered fields. Therefore, using (5.2.6) and (5.2.8), the S-matrix element can be written as

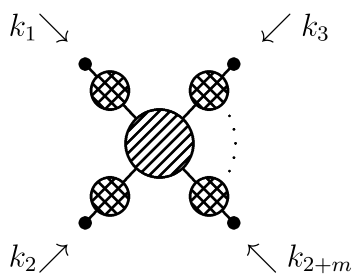
$$\begin{aligned}&\langle \vec{p}_1, \ldots \vec{p}_m | S | \vec{k}_1, \ldots \vec{k}_n \rangle \\ &= \int \prod_{j=1}^{n+m} \left(d^4 y_j \, \frac{k_j^2 - m^2 + i\epsilon}{i} \, e^{-ik_j \cdot y_j} \right) \frac{\left\langle T\left[\phi(y_1) \ \ldots \ \phi(y_{n+m}) \ \exp\left(i\int \ d^4x \mathcal{L}_I(x) \right) \right] \right\rangle_{\text{free}}}{\left\langle 0 | T\left[\exp\left(i\int d^4x \mathcal{L}_I(x) \right) \right] |0 \right\rangle_{\text{free}}}\end{aligned} \tag{5.2.23}$$

where all $\phi(y_j)$ are contracted with the fields in $\exp\left(i\int d^4x \mathcal{L}_I(x)\right)$ with all possible combinations. But since $k_i \neq -k_j$ the contraction of $\phi(y_i)$ with $\phi(y_j)$ is zero, hence we find

$$\langle \vec{p}_1, \ldots \vec{p}_m | i\,\mathcal{T} | \vec{k}_1, \ldots \vec{k}_n \rangle = \prod_{j=1}^{n+m} \frac{k_j^2 - m^2 + i\epsilon}{i} \, G^{(n+m)}(k_1, \ldots k_n, k_{n+1}, \ldots k_{n+m}) \tag{5.2.24}$$

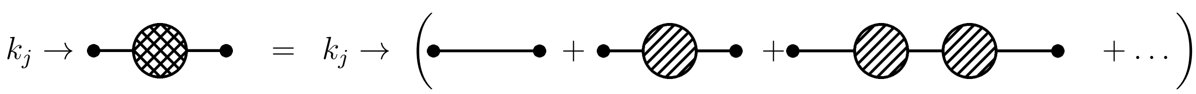
where $G^{(n+m)}(k_1, \ldots k_{n+m})$ is the renormalized $n+m$ correlator. We have replaced S with $i\,\mathcal{T}$ since the outgoing momenta are assumed not to equal the incoming momenta.

Specifying to the case that $n = 2$, we see that $G^{(2+m)}$ is connected, otherwise there would be $(2\pi)^4\delta^4(k_1 + k_j)$ or $(2\pi)^4\delta^4(k_2 + k_j)$ factors which are zero by assumption. The connected diagrams have the form



(5.2.25)

where



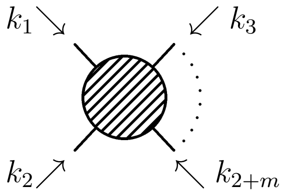
$$= \frac{i}{k_j^2 - m^2 - \Sigma(k_j^2) + i\epsilon}\,. \tag{5.2.26}$$

The k_j are the momenta for the incoming or outgoing particles, which are on-shell, *i.e.* $k_j^2 = m^2$. Since $\Sigma(m^2) = \frac{\partial}{\partial k^2}\Sigma(k^2)\Big|_{k^2=m^2} = 0$, the poles in the external momenta cancel out with the zeros in (5.2.24),

$$\left.\frac{k_j^2 - m^2 + i\epsilon}{k_j^2 - m^2 - \Sigma(k_j^2) + i\epsilon}\right|_{k_j^2=m^2} = 1\,. \tag{5.2.27}$$

We finally reach our desired result

$$\langle \vec{p}_1, \ldots \vec{p}_m | i\,\mathcal{T} | \vec{k}_1, \vec{k}_2\rangle = \tag{5.2.28}$$

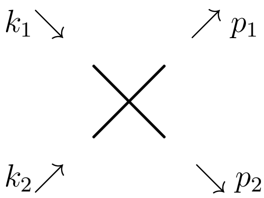


,

in other words, i times the $\mathcal{T}$ matrix is the *truncated*[1] correlator! This result is known as the Lehmann, Symanzik, Zimmermann (LSZ) reduction formula. You can now see why we were so interested in the truncated correlators.

As an example, consider the $2 \to 2$ scattering with incoming particles of momentum k_1 and k_2 and outgoing particles with momentum p_1 and p_2. To lowest order in the coupling λ the $\mathcal{T}$ matrix for this process is given by

$$\langle \vec{p}_1, \ldots \vec{p}_m | \mathcal{T} | \vec{k}_1, \vec{k}_2\rangle = -i$$



$$= -(2\pi)^4\delta^4(k_1 + k_2 - p_1 - p_2)\,\lambda\,. \tag{5.2.29}$$

Since any scattering process always has conservation of 4-momentum, it is convenient to define a new scattering matrix $\mathcal{M}$ such that

$$\mathcal{T} = (2\pi)^4\delta^4\left(\sum k_j\right)\mathcal{M}\,. \tag{5.2.30}$$

5.3 The cross section

Now that we have the amplitudes to go from incoming to outgoing states, we wish to relate this to quantities we can measure in the lab. In scattering experiments these are cross-sections.

Suppose we have a set of particles with flux f_1 aimed at another set of n_2 target particles. The flux is $\rho_1|v_1 - v_2|$, where ρ_1 is the density of the incoming particles and $v_1 - v_2$ is the relative velocity between the first set of particles and the second (we assume

[1]Peskin calls this amputated.

that the velocities are parallel). Then the interaction rate of the two sets of particles $\frac{dI}{dt}$ is proportional to $f_1 n_2$,

$$\frac{dI}{dt} = \sigma\, f_1 n_2\,. \tag{5.3.1}$$

Since $\frac{dI}{dt}$ has units of T^{-1} and $f_1 n_2$ has units of $T^{-1}L^{-2}$, σ must have units of an area. We call σ the cross-sectional area, or just cross-section, and indeed it would equal the cross-sectional area if the particles were hard classical spheres.

After a time T the total number of interactions is

$$I = \sigma\, f_1 n_2 T\,. \tag{5.3.2}$$

The total number of interactions is equal to the number of the first set of particles, n_1, multiplied by n_2, times the probability P_{12} that the particles will interact. Let us now suppose that our experiment is inside a volume V and that there is only one particle in each set. We then find that

$$\sigma = \frac{V}{|v_1 - v_2|T}\, P_{12}\,. \tag{5.3.3}$$

Hence to find the cross-section we need P_{12}.

We first assume that our two interacting particles are scalars and that in their interaction they produce two or more other scalars. The total number of produced scalars should be summed over since they all contribute to the interaction probability. However, one can also consider partial cross-sections where one distinguishes what type of final states one has. We will assume then the cross-section we compute corresponds to a final state with m particles.

Let the incoming state be $|\vec{k}_1, \vec{k}_2\rangle$ and the outgoing state be $|\vec{p}_1, \ldots, \vec{p}_m\rangle$. We do not necessarily assume that any of the particles come from the same scalar field, so they could have different masses, *etc.* According to (5.1.9) the probability to scatter into this state is

$$\begin{aligned} P_{2\to m} &= \frac{|\langle \vec{p}_1, \ldots, \vec{p}_m|\mathcal{T}|\vec{k}_1, \vec{k}_2\rangle|^2}{\langle \vec{k}_1, \vec{k}_2|\vec{k}_1, \vec{k}_2\rangle\langle \vec{p}_1, \ldots, \vec{p}_m|\vec{p}_1, \ldots, \vec{p}_m\rangle} \\ &= \frac{\left[(2\pi)^4\delta^4(P_{\rm out} - P_{\rm in})\right]^2}{(2k_1^0)(2k_2^0)[(2\pi)^3\delta^3(0)]^2\prod_j \left(2p_j^0(2\pi)^3\delta^3(0)\right)}\, |\langle \vec{p}_1, \ldots, \vec{p}_m|\mathcal{M}|\vec{k}_1, \vec{k}_2\rangle|^2\,, \end{aligned} \tag{5.3.4}$$

where $P_{\rm in} = k_1 + k_2$ and $P_{\rm out} = \sum_j p_j$. The several δ-functions evaluated at zero are strictly speaking infinite. However, we can deal with them as follows: We note that

$$(2\pi)^3\delta^3(\vec{k}) = \int d^3x e^{-i\vec{k}\cdot\vec{x}}\,. \tag{5.3.5}$$

If we set $\vec{k}=0$ then the right-hand side is the volume, which we cut-off at some arbitrarily large size V. Likewise, we have that

$$\left[(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\right]^2=\left[(2\pi)^4\delta^4(0)\right](2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})=VT(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\tag{5.3.6}$$

where T is the integrated time for the process.

The cross section will be inclusive, meaning that all possible final state momenta should be integrated over. The proper normalization for an integration over momentum in a volume V is $V\int\frac{d^3p}{(2\pi)^3}$, hence using (5.3.4) and (5.3.3) we finally obtain

$$\sigma_{2\to m}=\frac{S}{(2k_1^0)(2k_2^0)|v_1-v_2|}\int\prod_{j=1}^m\left(\frac{d^3p_j}{(2\pi)^3 2p_j^0}\right)(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\,|\langle\vec{p}_1,\dots,\vec{p}_m|\mathcal{M}|\vec{k}_1,\vec{k}_2\rangle|^2\,.\tag{5.3.7}$$

where S is a symmetry factor when some or all of the final state particles are identical. Notice that all factors of V and T have dropped out! Furthermore, this expression is Lorentz invariant, even though the first term might not look so. However, one can show that

$$(2k_1^0)(2k_2^0)|v_1-v_2|=2\,\left((s-m_1^2-m_2^2)^2-4m_1^2m_2^2\right)^{1/2}\tag{5.3.8}$$

where s is the Mandelstam variable $s=(k_1+k_2)^2$.

Let us consider the example of $2\to2$ scattering for scalars coming from a single real scalar field. In this case the outgoing particles are identical so the symmetry factor is $S=1/2$. The incoming particles have the identical mass $m_1=m_2=m$, hence to lowest order in the coupling the cross-section can be expressed as

$$\sigma_{2\to2}=\frac{\lambda^2}{4\sqrt{s(s-4m^2)}}\int\left(\frac{d^3p_1}{(2\pi)^3 2p_1^0}\right)\left(\frac{d^3p_2}{(2\pi)^3 2p_2^0}\right)(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\,,\tag{5.3.9}$$

where we used (5.2.29) for the $\mathcal{M}$-matrix. Let us now assume that we are in the Center of Momentum (COM) frame. In this case $\vec{k}_1=-\vec{k}_2$ and $\vec{p}_1=-\vec{p}_2$. It is common practice to drop the symmetry factor, but then restrict $\vec{p}_1$ to the forward direction. We then have

$$\begin{aligned}\sigma_{2\to2}&=\frac{\lambda^2}{2\sqrt{s(s-4m^2)}}\int_{\rm fwd}\left(\frac{d^3p_1}{(2\pi)^3(2E_1)^2}\right)(2\pi)\delta(2E_1-k_1^0-k_2^0)\\&=\frac{\lambda^2}{2\sqrt{s(s-4m^2)}}\int_{\rm fwd}\left(\frac{p_1^2dp_1\,d\Omega}{(2\pi)^3(2E_1)^2}\right)(2\pi)\delta(2E_1-k_1^0-k_2^0)\\&=\frac{\lambda^2}{2\sqrt{s(s-4m^2)}}\int_{\rm fwd}\left(\frac{p_1E_1dE_1\,d\Omega}{(2\pi)^3(2E_1)^2}\right)(2\pi)\frac{1}{2}\delta(E_1-\tfrac{1}{2}k_1^0-\tfrac{1}{2}k_2^0)\\&=\frac{\lambda^2}{2\sqrt{s(s-4m^2)}}\frac{p_1}{32\pi^2E_1}\int_{\rm fwd}d\Omega\,,\end{aligned}\tag{5.3.10}$$

where we set $E_1 = p_1^0$. In the COM frame $s = 4E_1^2$ while $s - 4m^2 = 4p_1^2$, which gives us

$$\sigma_{2\to 2} \quad = \quad \frac{\lambda^2}{64\pi^2\, s}\int_{\text{fwd}} d\Omega\,. \tag{5.3.11}$$

The integral over the solid angle gives 2π (since it is only in the forward direction), hence the cross-section is

$$\sigma_{2\to 2} \quad = \quad \frac{\lambda^2}{32\pi\, s}\,. \tag{5.3.12}$$

However, when measurements are made of the outgoing particles, the angles into which they scatter are generally known. So usually one expresses a *differential cross-section* $\dfrac{d\sigma_{2\to 2}}{d\Omega}$, which in this case is

$$\frac{d\sigma_{2\to 2}}{d\Omega} = \frac{\lambda^2}{64\pi^2\, s}\,. \tag{5.3.13}$$

Let us now compare this result to the scattering problem in quantum mechanics (*cf* chap. 11 of Griffiths). Here one has an incoming wave-function $\psi_{\text{in}}(\vec{r}) = e^{ikz}$ which scatters into a sum of partial waves

$$\psi_{\text{out}}(\vec{r}) \approx \frac{e^{ikr}}{r}\sum_{\ell=0}\frac{2l+1}{2i\,k}(\eta_\ell\, e^{2i\delta_\ell}-1)P_\ell(\cos\theta) = \frac{e^{ikr}}{r}\,F(\theta)\,, \tag{5.3.14}$$

where $P_\ell(\cos\theta)$ is the appropriate Legendre function, δ_ℓ is the phase-shift in the ℓ channel and η_ℓ=1 (η_ℓ<1) if the scattering is elastic (inelastic) in this channel. The scattering can be thought of as two-body in the COM frame where $\vec{r}$ is the relative separation between the two particles. The relative velocities between the particles is $|v_1 - v_2| = 2k/m$ if the velocities are nonrelativistic. For $k >> m$ then the relative velocity approaches 2. The flux per particle is the probability density multiplied by the relative velocity. Therefore, for the incoming flux this is $\psi^*_{\text{in}}(\vec{r})\psi_{\text{in}}(\vec{r})\,v = v$ up to a normalization factor. The flux per unit solid angle for the outgoing scattering is $\psi^*_{\text{out}}(\vec{r})\psi_{\text{out}}(\vec{r})r^2\,v = v|F(\theta)|^2$, up to the same normalization constant as the incoming flux. Therefore, the cross-section per unit solid angle is the outgoing flux/solid angle divided by the incoming flux,

$$\frac{d\sigma}{d\Omega} = \frac{\psi^*_{\text{out}}(\vec{r})\psi_{\text{out}}(\vec{r})r^2\,v}{\psi^*_{\text{in}}(\vec{r})\psi_{\text{in}}(\vec{r})\,v} = |F(\theta)|^2\,. \tag{5.3.15}$$

If we now compare (5.3.13) with (5.3.14) and (5.3.15) we see that the scalar particle scattering only has a contribution in the $\ell = 0$ channel, that is the s-wave channel. This is easily understood, the incoming and outgoing particles are spinless and the interaction is pointlike, meaning that there cannot be any orbital angular momentum. Thus, the total angular momentum is zero.

There is a physical limit to the cross-section. If we assume only s-wave scattering, then $F(\theta)$ is given by

$$F(\theta) = \frac{1}{2ik}\,(\eta_0\, e^{2i\delta_0}-1)\,. \tag{5.3.16}$$

$|F(\theta)|^2$ is maximum when $\eta_0 = 1$ and $\delta_0 = \frac{\pi}{2}$, in which case

$$|F(\theta)|^2 = \frac{1}{k^2}\,. \tag{5.3.17}$$

In the COM frame we have that $4\,k^2 = s - 4m^2$. Hence, we see that for high enough s (5.3.13) will violate this bound if $\lambda > 16\,\pi$. Such a breakdown would violate unitarity, which is a statement about the conservation of the probability current. However, it is also true that when $\lambda > 16\,\pi$, perturbation theory breaks down and the result in (5.3.13) cannot be trusted.

5.4 Scattering for fermions

5.4.1 LSZ for fermions

We proceed as in the scalar case. We first note that for an incoming fermion particle state $|\vec{k},+,s\rangle$, with momentum $\vec{k}$, charge $q = +1$ and spin s, the one-point function between the state and the vacuum for the renormalized field $\psi_R(x)$ is

$$\langle 0|\psi_R(x)|\vec{k},+,s\rangle = e^{-ik\cdot x}\,u^s(\vec{k})\,, \tag{5.4.1}$$

while for an anti-fermion we have

$$\langle 0|\overline{\psi}_R(x)|\vec{k},-,s\rangle = e^{-ik\cdot x}\,\overline{v}^s(\vec{k})\,. \tag{5.4.2}$$

For the outgoing states we find

$$\begin{aligned}\langle\vec{k},+,s|\overline{\psi}_R(x)|0\rangle &= e^{+ik\cdot x}\,\overline{u}^s(\vec{k})\\ \langle\vec{k},-,s|\psi_R(x)|0\rangle &= e^{+ik\cdot x}\,v^s(\vec{k})\,.\end{aligned} \tag{5.4.3}$$

Just like the scalars, these are the same relations one finds for the free particle case. Furthermore, the dressed fermions are the free fermions without the counterterms. Hence, a similar LSZ construction is possible for the fermions.

The combinatorics is messier for the fermions, so we shall limit ourselves to certain cases in the incoming and outgoing states. It's clear by angular momentum conservation that the number of fermions minus the number of anti-fermions cannot differ by an odd number between the incoming and outgoing states. If we include the vector charge, then the number of fermions minus anti-fermions is conserved[2]. For one fermion in the incoming state $|\vec{k},+,s\rangle_{\rm free} = a^{s\dagger}_{\vec{k}}|0\rangle_{\rm free}$ and one in the outgoing state ${}_{\rm free}\langle\vec{p},+,r| = {}_{\rm free}\langle 0|a^r_{\vec{p}}$ we have

$$\langle\vec{p},+,r|:\overline{\psi}_\alpha(x_1)\psi_\beta(x_2):|\vec{k},+,s\rangle_{\rm free} = e^{ip\cdot x_1}\overline{u}^r_\alpha(\vec{p})e^{-ik\cdot x_2}u^s_\beta(\vec{k})\,, \tag{5.4.4}$$

[2]If we have Majoranna fermions (problem 3.4 in Peskin) then a fermion is its own anti-fermion and $\psi\to e^{i\theta}\psi$ is no longer a good symmetry. Unless explicitly stated, we assume that the *fermion number* (fermions $-$ anti-fermions) is conserved.

while other combinations are given by

$$\begin{aligned}\langle\vec{p},-,r|:\overline{\psi}_\alpha(x_1)\psi_\beta(x_2):|\vec{k},-,s\rangle_{\rm free} &= -e^{ip\cdot x_2}v^r_\beta(\vec{p})e^{-ik\cdot x_1}\overline{v}^s_\alpha(\vec{k})\\ \langle\vec{0}|:\overline{\psi}_\alpha(x_1)\psi_\beta(x_2):|\vec{k}_1,+,s_1;\vec{k}_2,-,s_2\rangle_{\rm free} &= e^{-ik_1\cdot x_2}u^{s_1}_\beta(\vec{k}_1)e^{-ik_2\cdot x_1}\overline{v}^s_\alpha(\vec{k})\,.\end{aligned}\tag{5.4.5}$$

For two free incoming fermions (both with q=+), where the state is $|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle_{\rm free}=a^{s_1\dagger}_{\vec{k}_1}a^{s_2\dagger}_{\vec{k}_2}|0\rangle_{\rm free}$ and two outgoing free fermions ${}_{\rm free}\langle\vec{p}_1,+,r_1;\vec{p}_2,+,r_2|={}_{\rm free}\langle 0|a^{r_2}_{\vec{p}_2}a^{r_1}_{\vec{p}_1}$ we have the relation

$$\begin{aligned}&\langle\vec{p}_1,+,r_1;\vec{p}_2,+,r_2|:\overline{\psi}_{\alpha_1}(x_1)\overline{\psi}_{\alpha_2}(x_2)\psi_{\beta_1}(x_3)\psi_{\beta_2}(x_4):|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle_{\rm free}\\ &=\left(e^{ip_1\cdot x_1}\,\overline{u}^{r_1}_{\alpha_1}(\vec{p}_1)e^{ip_2\cdot x_2}\,\overline{u}^{r_2}_{\alpha_2}(\vec{p}_2)-e^{ip_1\cdot x_2}\,\overline{u}^{r_1}_{\alpha_2}(\vec{p}_1)e^{ip_2\cdot x_1}\,\overline{u}^{r_2}_{\alpha_1}(\vec{p}_2)\right)\\ &\qquad\times\left(e^{-ik_1\cdot x_4}\,u^{s_1}_{\beta_2}(\vec{k}_1)e^{-ik_2\cdot x_3}\,u^{s_2}_{\beta_1}(\vec{k}_2)-e^{-ik_1\cdot x_3}\,u^{s_1}_{\beta_1}(\vec{k}_1)e^{-ik_2\cdot x_4}\,u^{s_2}_{\beta_2}(\vec{p}_2)\right)\,,\end{aligned}\tag{5.4.6}$$

while for other combinations of in and out fermions the relations are

$$\begin{aligned}&\langle\vec{p}_1,-,r_1;\vec{p}_2,-,r_2|:\overline{\psi}_{\alpha_1}(x_1)\overline{\psi}_{\alpha_2}(x_2)\psi_{\beta_1}(x_3)\psi_{\beta_2}(x_4):|\vec{k}_1,-,s_1;\vec{k}_2,-,s_2\rangle_{\rm free}\\ &=\left(e^{ip_1\cdot x_3}\,v^{r_1}_{\beta_1}(\vec{p}_1)e^{ip_2\cdot x_4}\,v^{r_2}_{\beta_2}(\vec{p}_2)-e^{ip_1\cdot x_4}\,v^{r_1}_{\beta_2}(\vec{p}_1)e^{ip_2\cdot x_3}\,v^{r_2}_{\beta_1}(\vec{p}_2)\right)\\ &\qquad\times\left(e^{-ik_1\cdot x_2}\,\overline{v}^{s_1}_{\alpha_2}(\vec{k}_1)e^{-ik_2\cdot x_1}\,\overline{v}^{s_2}_{\alpha_1}(\vec{k}_2)-e^{-ik_1\cdot x_1}\,\overline{v}^{s_1}_{\alpha_1}(\vec{k}_1)e^{-ik_2\cdot x_2}\,\overline{v}^{s_2}_{\alpha_2}(\vec{p}_2)\right)\,,\end{aligned}\tag{5.4.7}$$

and

$$\begin{aligned}&\langle\vec{p}_1,+,r_1;\vec{p}_2,-,r_2|:\overline{\psi}_{\alpha_1}(x_1)\overline{\psi}_{\alpha_2}(x_2)\psi_{\beta_1}(x_3)\psi_{\beta_2}(x_4):|\vec{k}_1,+,s_1;\vec{k}_2,-,s_2\rangle_{\rm free}\\ &=\left(e^{ip_1\cdot x_1}\,\overline{u}^{r_1}_{\alpha_1}(\vec{p}_1)e^{-ik_2\cdot x_2}\,\overline{v}^{s_2}_{\alpha_2}(\vec{k}_2)-e^{ip_1\cdot x_2}\,\overline{u}^{r_1}_{\alpha_2}(\vec{p}_1)e^{-ik_2\cdot x_1}\,\overline{v}^{s_2}_{\alpha_1}(\vec{k}_2)\right)\\ &\qquad\times\left(e^{-ik_1\cdot x_3}\,u^{s_1}_{\beta_1}(\vec{k}_1)e^{ip_2\cdot x_4}\,v^{r_2}_{\beta_2}(\vec{p}_2)-e^{-ik_1\cdot x_4}\,u^{s_1}_{\beta_2}(\vec{k}_1)e^{ip_2\cdot x_3}\,v^{r_2}_{\beta_1}(\vec{p}_2)\right)\end{aligned}\tag{5.4.8}$$

We next write the analogue of (5.2.7) for fermion propagators

$$\begin{aligned}e^{-ik\cdot x} &= \int d^4y\int\frac{d^4q}{(2\pi)^4}e^{-iq\cdot(x-y)}\frac{i}{\not{q}-m+i\epsilon}\,\frac{\not{k}-m+i\epsilon}{i}\,e^{-ik\cdot y}\\ &= \int d^4y\,S_F(x-y)\,\frac{\not{k}-m+i\epsilon}{i}\,e^{-ik\cdot y}\,.\end{aligned}\tag{5.4.9}$$

and

$$\begin{aligned}e^{ip\cdot x} &= \int d^4y\int\frac{d^4q}{(2\pi)^4}e^{-iq\cdot(y-x)}\frac{\not{p}-m+i\epsilon}{i}\,\frac{i}{\not{q}-m+i\epsilon}\,e^{ip\cdot y}\\ &= \int d^4y\frac{\not{p}-m+i\epsilon}{i}\,S_F(y-x)\,e^{ip\cdot y}\,.\end{aligned}\tag{5.4.10}$$

It is then straightforward to show that

$$
\begin{aligned}
&\langle \vec{p}_1,+,r_1;\vec{p}_2,+,r_2|:\overline{\psi}_{\alpha_1}(x_1)\overline{\psi}_{\alpha_2}(x_2)\psi_{\beta_1}(x_3)\psi_{\beta_2}(x_4):|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle_{\rm free}\\
&=\int\prod_{j=1}^{2}(d^4y_je^{-ik_j\cdot y_j})\prod_{j=1}^{2}(d^4y_{j+2}e^{ip_j\cdot y_{j+2}})\\
&\times\Big\langle T\Big[\overline{u}^{r_2}(\vec{p}_2)\frac{\not{p}_2-m+i\epsilon}{i}\psi(y_4)\overline{u}^{r_1}(\vec{p}_1)\frac{\not{p}_1-m+i\epsilon}{i}\psi(y_3):\overline{\psi}_{\alpha_1}(x_1)\overline{\psi}_{\alpha_2}(x_2)\psi_{\beta_1}(x_3)\psi_{\beta_2}(x_4):\\
&\qquad\qquad\times\overline{\psi}(y_1)\frac{\not{k}_1-m+i\epsilon}{i}u^{s_1}(\vec{k}_1)\overline{\psi}(y_2)\frac{\not{k}_2-m+i\epsilon}{i}u^{s_2}(\vec{k}_2)\Big]\Big\rangle\,.
\end{aligned}
\tag{5.4.11}
$$

Using the same arguments as in the scalar case, the $\mathcal{T}$ matrix for fermion-fermion scattering is then

$$
\begin{aligned}
&\langle \vec{p}_1,+,r_1;\vec{p}_2,+,r_2|i\mathcal{T}|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle\\
&=\int\prod_{j=1}^{2}(d^4y_je^{-ik_j\cdot y_j})\prod_{j=1}^{2}(d^4y_{j+2}e^{ip_j\cdot y_{j+2}})\\
&\times\Big\langle T\Big[\overline{u}^{r_2}(\vec{p}_2)\frac{\not{p}_2-m+i\epsilon}{i}\psi(y_4)\overline{u}^{r_1}(\vec{p}_1)\frac{\not{p}_1-m+i\epsilon}{i}\psi(y_3)\exp\Big(i\textstyle\int d^4x\mathcal{L}_I(x)\Big)\\
&\qquad\qquad\times\overline{\psi}(y_1)\frac{\not{k}_1-m+i\epsilon}{i}u^{s_1}(\vec{k}_1)\overline{\psi}(y_2)\frac{\not{k}_2-m+i\epsilon}{i}u^{s_2}(\vec{k}_2)\Big]\Big\rangle\\
&\qquad\qquad\Big/\ \Big\langle T\Big[\exp\Big(i\textstyle\int d^4x\mathcal{L}_I(x)\Big)\Big]\Big\rangle
\end{aligned}
\tag{5.4.12}
$$

As in the scalar case, this is a truncated diagram, although with the added feature of having the Dirac wave-functions $u^s(\vec{k})$ attached. For anti-fermion anti-fermion scattering, one replaces $\overline{u}^{r_2}(\vec{p}_2)\frac{\not{p}_2-m+i\epsilon}{i}\psi(y_4)$ with $\overline{\psi}(y_4)\frac{-\not{p}_2-m+i\epsilon}{i}v^{r_2}(\vec{p}_2)$, *etc.* For fermion anti-fermion scattering we replace half the terms.

5.5 Feynman rules for fermions

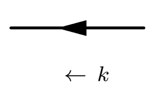
Before we can compute the truncated diagrams in the last section we need the Feynman rules for fermions. We will consider two types of interacting terms in this section, a Yukawa interaction with a real scalar and a four-fermion interaction.

5.5.1 Yukawa interactions

The fermion propagator is

$$
S_F(x)=\langle\psi(x)\overline{\psi}(0)\rangle=\int\frac{d^4k}{(2\pi)^4}\frac{i}{\not{k}-m+i\epsilon}\,e^{-ik\cdot x}\,,
\tag{5.5.1}
$$

In momentum space we draw the propagator diagram as



$$
=\frac{i}{\not{k}-m+i\epsilon}
\tag{5.5.2}
$$

The direction of the arrow matters since the propagator is not invariant under $k^\mu \to -k^\mu$. It points from the $\overline{\psi}$ to the ψ. Equivalently, it points in the direction of propagation for a fermion, or in the opposite direction of propagation for an anti-fermion.

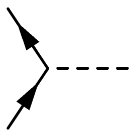
A Yukawa potential comes from the coupling between a scalar and fermion bilinears. In the simplest case where there is one real scalar and one Dirac fermion, the contribution to the action is given by

$$\mathcal{L}_Y = -g\,\phi\,\overline{\psi}\,\psi\,, \tag{5.5.3}$$

where g is called the Yukawa coupling. With fermions present, it is customary to draw scalar propagators using dashed lines,

$$= \frac{i}{k^2 - M^2 + i\epsilon} \tag{5.5.4}$$

where we use M for the scalar mass. The scalar-fermion-fermion vertex is drawn as



$$= -ig\,(2\pi)^4\delta^4(k_1 + k_2 + k_3) \tag{5.5.5}$$

where k_j is the momenta entering the vertex. In addition, we have the rules

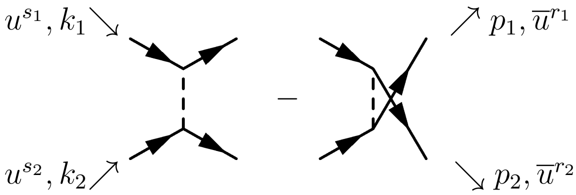
- Integrate over undetermined momenta
- Attach $u^s(\vec{k})$ $(\overline{v}^s(\vec{k}))$ for an incoming fermion (anti-fermion) with spin s. Attach $\overline{u}^r(\vec{k})$ $(v^r(\vec{k}))$ for an outgoing fermion (anti-fermion) with spin r.
- Multiply by (-1) for every fermion loop. Trace over Dirac indices for closed fermion loops.

To see where the last rule comes from, observe that a fermion loop arises from a series of contractions with the form

$$:\overline{\psi}(x_1)\psi(x_1)::\overline{\psi}(x_2)\psi(x_2):\ldots\cdots:\overline{\psi}(x_{n-1})\psi(x_{n-1})::\overline{\psi}(x_n)\psi(x_n):\ . \tag{5.5.6}$$

The neighboring contractions have the form $\psi(x_j)\overline{\psi}(x_{j+1}) = S_F(x_j - x_{j+1})$ while the outer one is $\overline{\psi}(x_1)\psi(x_n) = -\psi(x_n)\overline{\psi}(x_1) = -S_F(x_n - x_1)$. The neighboring propagators, which are 4×4 matrices, are multiplied together, with a final mulitplication by $-S_F(x_n - x_1)$. But since the Dirac index on $\overline{\psi}(x_1)$ is contracted with the index on $\psi(x_1)$, there is an overall trace on the product.

We now consider the amplitude for fermion-fermion scattering with the Yukawa potential. At lowest order in perturbation theory the relevant diagrams are



$$= (2\pi)^4\delta^4(k_1+k_2-p_1-p_2)(-ig)^2$$

$$\times\left(\frac{i\,\overline{u}^{r_1}(\vec{p}_1)u^{s_1}(\vec{k}_1)\overline{u}^{r_2}(\vec{p}_2)u^{s_2}(\vec{k}_2)}{(p_1-k_1)^2-M^2+i\epsilon}-\frac{i\,\overline{u}^{r_1}(\vec{p}_1)u^{s_2}(\vec{k}_2)\overline{u}^{r_2}(\vec{p}_2)u^{s_1}(\vec{k}_1)}{(p_1-k_2)^2-M^2+i\epsilon}\right) \tag{5.5.7}$$

where the relative minus sign comes from the exchange of the first and second outgoing fermions. Therefore, to leading order

$$\begin{aligned}&\langle\vec{p}_1,+,r_1;\vec{p}_2,+,r_2|\mathcal{M}|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle\\ &=-g^2\left(\frac{\overline{u}^{r_1}(\vec{p}_1)u^{s_1}(\vec{k}_1)\overline{u}^{r_2}(\vec{p}_2)u^{s_2}(\vec{k}_2)}{(p_1-k_1)^2-M^2+i\epsilon}-\frac{\overline{u}^{r_1}(\vec{p}_1)u^{s_2}(\vec{k}_2)\overline{u}^{r_2}(\vec{p}_2)u^{s_1}(\vec{k}_1)}{(p_1-k_2)^2-M^2+i\epsilon}\right).\end{aligned} \tag{5.5.8}$$

This expression has a pole in the t-channel when $t=(p_1-k_1)^2=M^2$, and a pole in the u-channel when $u=(p_1-k_2)^2=M^2$. However, since the p_i and k_i are on-shell, t and u satisfy $t\le 0$, $u\le 0$. Thus, the poles are never reached for physical particles.

To find the cross-section for this process, we assume that the final spin states are not determined, meaning that they should be summed over. Hence, the cross-section is

$$\begin{aligned}\sigma_{2\to2} &= \frac{1}{4\sqrt{s(s-4m^2)}}\sum_{r_1,r_2}\int\left(\frac{d^3p_1}{(2\pi)^3 2p_1^0}\right)\left(\frac{d^3p_2}{(2\pi)^3 2p_2^0}\right)(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\\ &\qquad\times\left|\langle\vec{p}_1,+,r_1;\vec{p}_2,+,r_2|\mathcal{M}|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle\right|^2,\end{aligned} \tag{5.5.9}$$

where we included the symmetry factor $S=1/2$ since the fermions are identical. We then use (4.1.35 – 4.1.37) from the fourth chapter of the notes to write

$$\begin{aligned}\sum_{r_1,r_2}\left|\overline{u}^{r_1}(\vec{p}_1)u^{s_1}(\vec{k}_1)\overline{u}^{r_2}(\vec{p}_2)u^{s_2}(\vec{k}_2)\right|^2 &= \overline{u}^{s_1}(\vec{k}_1)(\not{p}_1+m)u^{s_1}(\vec{k}_1)\overline{u}^{s_2}(\vec{k}_2)(\not{p}_2+m)u^{s_2}(\vec{k}_2)\\ &= 4(p_1\cdot k_1+m^2)(p_2\cdot k_2+m^2)=(t-4m^2)^2,\end{aligned} \tag{5.5.10}$$

where we used that $t=(p_1-k_1)^2=(p_2-k_2)^2$. Similarly,

$$\sum_{r_1,r_2}\left|\overline{u}^{r_1}(\vec{p}_1)u^{s_2}(\vec{k}_2)\overline{u}^{r_2}(\vec{p}_2)u^{s_1}(\vec{k}_1)\right|^2=(u-4m^2)^2\,, \tag{5.5.11}$$

where $u=(p_2-k_1)^2=(p_1-k_2)^2$. For the cross terms we encounter the combination

$$\begin{aligned}&\sum_{r_1,r_2}\overline{u}^{r_1}(\vec{p}_1)u^{s_1}(\vec{k}_1)\overline{u}^{r_2}(\vec{p}_2)u^{s_2}(\vec{k}_2)\overline{u}^{s_2}(\vec{k}_2)u^{r_1}(\vec{p}_1)\overline{u}^{s_1}(\vec{k}_1)u^{r_2}(\vec{p}_2)+\text{ c.c.}\\ &\qquad=\overline{u}^{s_2}(\vec{k}_2)(\not{p}_1+m)u^{s_1}(\vec{k}_1)\overline{u}^{s_1}(\vec{k}_1)(\not{p}_2+m)u^{s_2}(\vec{k}_2)+\text{ c.c.}\end{aligned} \tag{5.5.12}$$

To go further with this expression we will assume that scattering arises from a collision of two beams of unpolarized fermions. This means that the spins of the incoming fermions are random and should be averaged over. To take the average we sum over s_1 and s_2 and divide each sum by 2. Applying this to (5.5.12) we get

$$\begin{aligned}
&\frac{1}{4}\sum_{s_1,s_2}\overline{u}^{s_2}(\vec{k}_2)(\not{p}_1+m)u^{s_1}(\vec{k}_1)\overline{u}^{s_1}(\vec{k}_1)(\not{p}_2+m)u^{s_2}(\vec{k}_2)+\ \text{c.c.}\\
&=\frac{1}{4}\left(\mathrm{Tr}\left[(\not{p}_1+m)(\not{k}_1+m)(\not{p}_2+m)(\not{k}_2+m)\right]+\mathrm{Tr}\left[(\not{p}_1+m)(\not{k}_2+m)(\not{p}_2+m)(\not{k}_1+m)\right]\right)\\
&=\frac{1}{2}\mathrm{Tr}\left[(\not{p}_1+m)(\not{k}_1+m)(\not{p}_2+m)(\not{k}_2+m)\right]\,. \qquad (5.5.13)
\end{aligned}$$

We then use the Clifford algebra in (4.1.4) in the fourth chapter of the notes to derive the trace relations for the γ-matrices,

$$\begin{aligned}
\mathrm{Tr}[\gamma^\mu\gamma^\nu] &= \frac{1}{2}\mathrm{Tr}[\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu]=\eta^{\mu\nu}\mathrm{Tr}[1]=4\eta^{\mu\nu}\\
\mathrm{Tr}[\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma] &= \frac{1}{2}\mathrm{Tr}[\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma+\gamma^\nu\gamma^\lambda\gamma^\sigma\gamma^\mu]\\
&= \eta^{\mu\nu}\mathrm{Tr}[\gamma^\lambda\gamma^\sigma]+\frac{1}{2}\mathrm{Tr}[-\gamma^\nu\gamma^\mu\gamma^\lambda\gamma^\sigma+\gamma^\nu\gamma^\lambda\gamma^\sigma\gamma^\mu]\\
&= 4\eta^{\mu\nu}\eta^{\lambda\sigma}-\eta^{\mu\lambda}\mathrm{Tr}[\gamma^\nu\gamma^\sigma]+\frac{1}{2}\mathrm{Tr}[\gamma^\nu\gamma^\lambda(\gamma^\mu\gamma^\sigma+\gamma^\sigma\gamma^\mu)]\\
&= 4\left(\eta^{\mu\nu}\eta^{\lambda\sigma}-\eta^{\mu\lambda}\eta^{\nu\sigma}+\eta^{\mu\sigma}\eta^{\nu\lambda}\right)\,. \qquad (5.5.14)
\end{aligned}$$

The trace of an odd product of γ-matrices is zero because the Clifford algebra and the cyclicity of the trace reduces any combination to a sum of traces of a single γ-matrix multiplied by $\eta^{\mu\nu}$ factors. With these results and after some manipulations, the result in (5.5.13) reduces to

$$\begin{aligned}
&\frac{1}{2}\mathrm{Tr}\left[(\not{p}_1+m)(\not{k}_1+m)(\not{p}_2+m)(\not{k}_2+m)\right]\\
&=4\left(m^4+2m^2(p_1\cdot p_2+p_1\cdot k_1+p_1\cdot k_2)+p_1\cdot k_1 p_2\cdot k_2-p_1\cdot p_2 k_1\cdot k_2+p_1\cdot k_2 p_2\cdot k_1\right)\\
&=4s\,m^2-u\,t\,. \qquad (5.5.15)
\end{aligned}$$

Putting everything together, we get for the cross-section (assuming unpolarized fermions)

$$\begin{aligned}
\sigma_{2\to2} &= \frac{g^4}{4\sqrt{s(s-4m^2)}}\int\left(\frac{d^3p_1}{(2\pi)^3 2p_1^0}\right)\left(\frac{d^3p_2}{(2\pi)^3 2p_2^0}\right)(2\pi)^4\delta^4(P_{\text{out}}-P_{\text{in}})\\
&\quad\times\left(\frac{(t-4m^2)^2}{(t-M^2+i\epsilon)^2}+\frac{(u-4m^2)^2}{(u-M^2+i\epsilon)^2}-\frac{4sm^2-ut}{(t-M^2+i\epsilon)(u-M^2+i\epsilon)}\right) \qquad (5.5.16)
\end{aligned}$$

Assuming we are in the COM frame and following the steps in (5.3.10) we then find

$$\sigma_{2\to2} = \frac{g^4}{64\pi^2\,s}\int_{\text{fwd}}d\Omega\left(\frac{(t-4m^2)^2}{(t-M^2+i\epsilon)^2}+\frac{(u-4m^2)^2}{(u-M^2+i\epsilon)^2}-\frac{4sm^2-ut}{(t-M^2+i\epsilon)(u-M^2+i\epsilon)}\right)\,. \qquad (5.5.17)$$

Therefore, the differential cross-section is

$$\frac{d\sigma_{2\to2}}{d\Omega}=\frac{g^4}{64\pi^2\,s}\left(\frac{(t-4m^2)^2}{(t-M^2+i\epsilon)^2}+\frac{(u-4m^2)^2}{(u-M^2+i\epsilon)^2}-\frac{4sm^2-ut}{(t-M^2+i\epsilon)(u-M^2+i\epsilon)}\right)\tag{5.5.18}$$

One can express the Mandelstam variables t and u in terms of s and the polar angle,

$$\begin{aligned} t &= (4m^2-s)\,\sin^2\frac{\theta}{2}\,, & u&=(4m^2-s)\,\cos^2\frac{\theta}{2}\,,\\ t-4m^2 &= -4m^2\cos^2\frac{\theta}{2}-s\,\sin^2\frac{\theta}{2}\,, & u-4m^2&=-4m^2\sin^2\frac{\theta}{2}-s\,\cos^2\frac{\theta}{2}\,,\end{aligned}\tag{5.5.19}$$

and then insert these expressions into (5.5.18) to find the angular dependence of the differential cross-section.

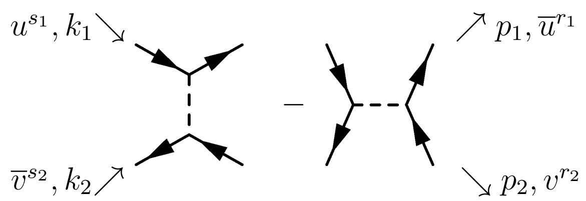
In the case where $M^2>>m^2$ and $M^2>>s,\ |t|,\ |u|$, the differential cross-section reduces to

$$\frac{d\sigma_{2\to2}}{d\Omega}=\frac{g^4}{64\pi^2\,s\,M^4}\left((s-4m^2)^2\left(\frac{3}{4}+\frac{1}{4}\cos^2\theta\right)+4m^2s\right)\,.\tag{5.5.20}$$

From the partial wave expansion in (5.3.14) we can see that this is a combination of angular momentum $\ell=0$ (constant in θ) and $\ell=1$ ($\cos^2\theta$) partial waves. This should be expected. At energies where $s,t,u<<M^2$, the interaction is effectively point-like (to be justified below), so angular momentum can only be coming from the spins of the particles. The particles are spin 1/2, so the total angular momentum can only be $J=0$ or $J=1$. The outgoing fermions can have total spin of $S=0$ or $S=1$. Hence to conserve angular momentum there could be an orbital angular momentum of $\ell=0$ or $\ell=1$. If the total incoming and outgoing spin is $S=1$, it might seem to be possible to have $\ell=2$ and still conserve angular momentum. However, the outgoing states cannot be in an $S=1$, $\ell=2$ state because this violates fermion statistics. Both $S=1$ and $\ell=2$ are symmetric under interchanges.

When s, t and u approach the order of M^2 then the angular dependence in the denominators in (5.5.18) becomes important and higher ℓ modes appear in the cross-section. What happens at these higher energies is that the fermion-fermion interaction is no longer point-like and so the incoming fermions can have significant orbital angular momentum.

Let us now turn to fermion anti-fermion scattering in the Yukawa theory. The relevant Feynman diagrams are



$$=(2\pi)^4\delta^4(k_1+k_2-p_1-p_2)(-ig)^2$$
$$\times\left(\frac{i\,\overline{u}^{r_1}(\vec p_1)u^{s_1}(\vec k_1)\overline{v}^{s_2}(\vec k_2)v^{r_2}(\vec p_2)}{(p_1-k_1)^2-M^2+i\epsilon}-\frac{i\,\overline{v}^{s_2}(\vec k_2)u^{s_1}(\vec k_1)\overline{u}^{r_1}(\vec p_1)v^{r_2}(\vec p_2)}{(k_1+k_2)^2-M^2+i\epsilon}\right)\tag{5.5.21}$$

Now the poles are in the t and the s channels. In fact, if $M^2 > 4m^2$ then the s-channel pole is attainable for on-shell particles. This leads to an enhancement of the cross-section at $s = M^2$, called a resonance. The resonance arises from the fermion anti-fermion colliding to form a physical scalar particle.

But the same Yukawa interaction that leads to creation of the scalar particle also leads to scalar particle decay into a fermion anti-fermion pair. The single particle wave-function for a zero momentum particle that decays is given by $\Psi(x) \sim e^{-iMt-\Gamma t/2}$, such that the probability density is $\Psi^\dagger(x)\Psi(x) \sim e^{-\Gamma t}$, where Γ^{-1} is the particle lifetime. We can see that the effect of the decay is to give the particle a complex mass, $M - i\Gamma/2$. This then moves the s-channel pole further down into the lower half-plane (it was already there to begin with because of the $i\,\epsilon$).

The amplitude in (5.5.21) is very similar to the amplitude in (5.5.7). This is because of a symmetry of the amplitudes called crossing symmetry. In this particular case, the transformation is $k_2 \leftrightarrow -p_2$, $r_2 \leftrightarrow s_2$. As far as the Dirac wave-functions are concerned, the transformation is $u^{s_2}(\vec{k}_2) \to v^{r_2}(\vec{p}_2)$, $\overline{u}^{r_2}(\vec{p}_2) \to \overline{v}^{s_2}(\vec{k}_2)$. You can see that the crossing symmetry is generated by twisting the bottom incoming and outgoing fermion lines. It is easy to see that this transformation takes the t-channel to the t-channel. With a little more perseverance one can also see that the twist will take the u-channel to the s-channel.

5.5.2 Particle decays

To find the decay rate for a particle we need the probability for a transition. Similarly to (5.3.4), the probability for a particle with 4-momentum k^μ to decay into m particles is (neglecting for the moment internal quantum numbers for the outgoing particles)

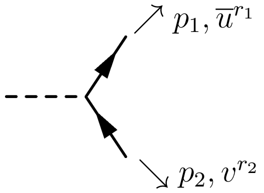
$$\begin{aligned} P_{1\to m} &= \frac{|\langle\vec{p}_1,\ldots,\vec{p}_m|\mathcal{T}|\vec{k}\rangle|^2}{\langle\vec{k}|\vec{k}\rangle\langle\vec{p}_1,\ldots,\vec{p}_m|\vec{p}_1,\ldots,\vec{p}_m\rangle} \\ &= \frac{\left[(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\right]^2}{(2k^0)[(2\pi)^3\delta^3(0)]\prod\limits_j\left(2p_j^0(2\pi)^3\delta^3(0)\right)}\,|\langle\vec{p}_1,\ldots,\vec{p}_m|\mathcal{M}|\vec{k}_1,\vec{k}_2\rangle|^2\,, \end{aligned} \tag{5.5.22}$$

If the final state is a fermion anti-fermion pair, then the rate to decay into any such configuration is the probability divided by the time over which the transition takes place,

$$\begin{aligned} \Gamma = \frac{P_{\phi\to\overline{\psi}\psi}}{T} &= \frac{1}{2k^0}\sum_{r_1,r_2}\int\left(\frac{d^3p_1}{(2\pi)^3 2p_1^0}\right)\left(\frac{d^3p_2}{(2\pi)^3 2p_2^0}\right)(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in}) \\ &\qquad\times\left|\langle\vec{p}_1,+,r_1;\vec{p}_2,-,r_2|\mathcal{M}|\vec{k}\rangle\right|^2\,, \end{aligned} \tag{5.5.23}$$

where we used (5.3.5), (5.3.6) and the standard measure for integration over the momentum given just above (5.3.7). The decay includes all possible final spin states for the fermion and anti-fermion, so the rate includes a sum over the spins. As in the case for the cross-section, all volume factors V have divided out.

To lowest order in g the $\mathcal{M}$-matrix is

$\langle \vec{p}_1, +, r_1; \vec{p}_2, -, r_2 | i\,\mathcal{M} | \vec{k}\rangle = \ k \rightarrow$ 

$= -i\, g\, \overline{u}^{r_1}(\vec{p}_1) v^{r_2}(p_2)\,. \quad (5.5.24)$

Therefore,

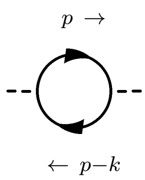
$$\begin{aligned}\sum_{r_1,r_2} \left| \langle \vec{p}_1, +, r_1; \vec{p}_2, -, r_2 | \mathcal{M} | \vec{k}\rangle \right|^2 &= g^2 \mathrm{Tr}(\not{p}_1 + m)(\not{p}_2 - m) = 4g^2\,(p_1{\cdot}p_2 - m^2) \\ &= 2\,g^2(M^2 - 4m^2)\,, \end{aligned} \tag{5.5.25}$$

where in the last step we used momentum conservation for the decay. If we go to the rest frame of the decaying particle, which is the COM frame, then $k^0 = M$ and the δ-function in (5.5.23) gives $\vec{p}_2 = -\vec{p}_1$. Therefore, the rate is

$$\begin{aligned}\Gamma &= \frac{g^2(M^2 - 4m^2)}{M} \int \frac{d^3p_1}{(2\pi)^3 (2p_1^0)^2}\,(2\pi)\delta(M - 2p_1^0) \\ &= \frac{g^2(M^2 - 4m^2)}{\pi\, M} \int \frac{p_1 p_1^0 dp_1^0}{(2p_1^0)^2}\,\delta(M - 2p_1^0) = \frac{g^2(M^2 - 4m^2)^{3/2}}{8\pi\, M^2}\,. \end{aligned} \tag{5.5.26}$$

5.5.3 Scalar self-energy

Here we describe the one-loop contribution to the scalar self-energy in the Yukawa theory. The relevant Feynman diagram is

$-\,i\Sigma_0(k^2) = \ k \rightarrow$ 

$= (-1)(-ig)^2 \int \frac{d^4p}{(2\pi)^4} \mathrm{Tr}\left[\frac{i}{\not{p} - m + i\epsilon}\,\frac{i}{\not{p} - \not{k} - m + i\epsilon}\right]$

(5.5.27)

where the (-1) is for the fermion loop. Notice the signs of the momentum of the second propagator. We are assuming that p runs through the top fermion propagator from left to right, hence by momentum conservation the bottom propagator has momentum $k - p$ from left to right. But the signs are determined by the direction of the arrow, which in this case goes right to left, so we use $p - k$. To evaluate the trace we write

$$\begin{aligned}\mathrm{Tr}\left[\frac{i}{\not{p} - m + i\epsilon}\,\frac{i}{\not{p} - \not{k} - m + i\epsilon}\right] &= -\frac{\mathrm{Tr}[(\not{p} + m)(\not{p} - \not{k} + m]}{(p^2 - m^2 + i\epsilon)((p-k)^2 - m^2 + i\epsilon)} \\ &= -\frac{4(p^2 - p{\cdot}k + m^2)}{(p^2 - m^2 + i\epsilon)((p-k)^2 - m^2 + i\epsilon)} \\ &= -\frac{2}{p^2 - m^2 + i\epsilon} - \frac{2}{(p-k)^2 - m^2 + i\epsilon} + \frac{2(k^2 - 4m^2)}{(p^2 - m^2 + i\epsilon)((p-k)^2 - m^2 + i\epsilon)} \end{aligned} \tag{5.5.28}$$

Substituting back into (5.5.27) we have

$$-i\Sigma_0(k^2) = -4g^2\int\frac{d^4p}{(2\pi)^4}\frac{1}{p^2-m^2+i\epsilon}+2g^2\int\frac{d^4p}{(2\pi)^4}\frac{k^2-4m^2}{(p^2-m^2+i\epsilon)((p-k)^2-m^2+i\epsilon)}\,, \tag{5.5.29}$$

where we shifted the integration variable for one of the integrals from p to $p-k$. Both integrals we have encountered before in our study of ϕ^4 theory, so we can simply borrow the results from there. The first integral is found in (3.3.16) of chapter 3, while the second integral is discussed in (3.4.9 – ImQ) of the same chapter . The first integral gives a mass-shift that can be cancelled by a counterterm. The second integral can be found by comparing to the first line of (3.4.9) in the third chapter and then using (3.4.11). Hence, we find

$$\Sigma_0(k^2) = \Delta m^2 - \frac{g^2\,(k^2-4m^2)}{8\,\pi^2}\left(\frac{1}{\epsilon}-\gamma_E+\log(4\pi)+\log\frac{\mu^2}{m^2}+Q\left(\frac{k^2}{m^2}\right)\right)\,, \tag{5.5.30}$$

with $Q(z)$ defined in (3.4.12) and Δm^2 the contribution from the first integral in (5.5.29). If we set $k^2 = M^2$ then $\Sigma_0(M^2)$ is the total mass-shift, the real part of which can be canceled by a counterterm. Our interest is in the imaginary part. Using (3.4.13) in the third chapter , we have that

$$\mathrm{Im}\left(Q\left(\frac{M^2}{m^2}+i\epsilon\right)\right) = \frac{\pi\sqrt{M^2-4m^2}}{M}\,, \tag{5.5.31}$$

Thererfore,

$$\Gamma = -\frac{\mathrm{Im}(\Sigma_0(M^2))}{M} = g^2\frac{(M^2-4m^2)^{3/2}}{8\pi M^2} \tag{5.5.32}$$

which agrees with (5.5.26).

5.5.4 Four fermion interaction

The other type of potential we consider is a 4-fermion interaction with the interaction Lagrangian

$$\mathcal{L}_{\overline{\psi}\psi\overline{\psi}\psi} = \frac{1}{2}G\,\overline{\psi}(x)\psi(x)\,\overline{\psi}(x)\psi(x)\,. \tag{5.5.33}$$

We actually will not go through the derivation of the cross-section, because we can get it directly from the Yukawa analysis.

The scalar part of the action for the Yukawa coupled theory is

$$S_{\text{scalar}} = \int d^4x\left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\,M^2\phi^2 - g\,\phi\overline{\psi}\psi\right)\,. \tag{5.5.34}$$

Since the scalar part is quadratic in ϕ we can do the functional integral for the scalars in the path integral. Hence,

$$\int\mathcal{D}\phi\, e^{iS_{\text{scalar}}} \sim \exp\left(-i\int d^4x d^4y\,\frac{1}{2}\,g^2\overline{\psi}(x)\psi(x)G_F(x-y)\overline{\psi}(y)\psi(y)\right)\,. \tag{5.5.35}$$

Therefore, the fermions have an effective action

$$S_{\text{eff}} = \int d^4x\,\overline{\psi}(x)(i\partial\!\!\!/ - m)\psi(x) - \frac{g^2}{2}\int d^4x d^4y\,\overline{\psi}(x)\psi(x)G_F(x-y)\overline{\psi}(y)\psi(y)\,, \tag{5.5.36}$$

where $G_F(x-y)$ is the Feynman propagator which leads to an effective non-local interaction between the fermions. The interaction only has a range of M^{-1}, as you can see by considering $x-y$ space-like separated, which for $|x-y|M >> 1$ has $G_F(x-y) \sim e^{-M|x-y|}$. Therefore, the effective potential between the fermions has an exponential suppression for distances greater than M^{-1}. Hence, if we consider scatterings with $s << M^2$ then the interaction is effectively point-like since the particle's deBroglie wave-length is much larger than M^{-1}.

If $s << M^2$, then we can simplify the effective action even more. In this case the scalar kinetic term is less important than the mass and Yukawa terms. Now we have

$$S_{\text{scalar}} \approx \int d^4x \left(-\frac{1}{2}\,M^2\phi^2 - g\,\phi\,\overline{\psi}\psi\right)\,. \tag{5.5.37}$$

and ϕ becomes a Lagrange multiplier. Solving for ϕ we get

$$\phi = -\frac{g}{M^2}\overline{\psi}(x)\psi(x) \tag{5.5.38}$$

Substituting this back into the action we reach the effective action for the fermions,

$$S_{\text{eff}} = \int d^4x \left(\overline{\psi}(x)(i\partial\!\!\!/ - m)\psi(x) + \frac{g^2}{2M^2}\overline{\psi}(x)\psi(x)\overline{\psi}(x)\psi(x)\right)\,. \tag{5.5.39}$$

Comparing with (5.5.33) we identify $G = \dfrac{g^2}{M^2}$. Now, if we assume that the four-fermion interaction in (5.5.39) *is* the action and not just the effective action for low energies, then from (5.5.20) we find that the differential cross-section to lowest order in G is

$$\frac{d\sigma_{2\to 2}}{d\Omega} = \frac{G^2}{64\pi^2\,s}\left((s-4m^2)^2\left(\frac{3}{4}+\frac{1}{4}\cos^2\theta\right)+4m^2s\right) \tag{5.5.40}$$

for any value of s.

But for high enough values of s (5.5.40) will run into problems because the cross-section grows linearly in s and eventually will cross the unitary bound that has the cross-section falling off in s^{-1}. So basically, a four-fermion interaction only makes sense at relatively low energies where there is no danger of crossing the unitary bound. Once s becomes large enough another theory needs to take over. You can see that the Yukawa theory does not have these problems, assuming g is in the perturbative regime, because the denominators in (5.5.18) also grow with s, such that at very high energies the term inside the large parentheses approaches 3 and the cross-section has an s^{-1} fall-off.

Why is a four-scalar interaction fine as far as the unitary bound is concerned, but problematic for four fermions? The answer has to do with the dimensions of the coupling. The scalar coupling λ is dimensionless. Hence, for very high energies when s is greater than any mass-scale, the cross-section must behave as $\sigma \sim \lambda^2 s^{-1}$ by dimensional arguments since there is no other scale around. However, G has dimension $D = -2$ since the fermions have $D = 3/2$. Hence, at very high energies the cross-section has to behave as $\sigma \sim G^2\,s$.

Chapter 6

Photons and QED

In this chapter of the lecture notes we begin our investigation of Quantum ElectroDynamics (QED). We will discuss gauge invariance and the electromagnetic field, quantization of the field and the Feyman rules for QED. We also consider some elementary (relatively) scattering amplitudes.

6.1 The gauge field

6.1.1 $U(1)$ gauge invariance

In the first and fourth chapter of these notes we discussed a global continuous symmetry for complex scalar fields and Dirac fermions. The transformation for both types of fields is

$$\begin{array}{rcl} \phi(x) & \to & e^{i\theta}\phi(x)\,, \qquad \phi^*(x)\to\phi^*(x)e^{-i\theta} \\ \psi(x) & \to & e^{i\theta}\psi(x)\,, \qquad \overline{\psi}(x)\to\overline{\psi}(x)e^{-i\theta}\,, \end{array} \tag{6.1.1}$$

where θ is constant in x.

If we attempt to make this transformation *local*, where $\theta=\theta(x)$, then the kinetic terms $\partial_\mu\phi^*\partial^\mu\phi$ and $i\overline{\psi}\partial\!\!\!/\psi$ are not invariant, but instead transform as

$$\begin{array}{rcl} \partial_\mu\phi^*\partial^\mu\phi & \to & (\partial_\mu - i(\partial_\mu\theta))\phi^*(\partial_\mu + i(\partial_\mu\theta))\phi \\ i\overline{\psi}(x)\partial\!\!\!/\psi & \to & i\overline{\psi}(\partial\!\!\!/ + i(\partial\!\!\!/\theta))\psi \end{array} \tag{6.1.2}$$

because of the derivatives acting on the $e^{i\theta(x)}$ factors. We can restore the symmetry if we replace the derivative ∂_μ with the *covariant deriative* D_μ, where

$$D_\mu = \partial_\mu + ie\,A_\mu \qquad . \tag{6.1.3}$$

Under the transformation A_μ transforms as

$$A_\mu \to A_\mu - \frac{1}{e}\partial_\mu\theta\,, \tag{6.1.4}$$

and we see that the $\partial_\mu\theta(x)$ terms cancel, leaving the Lagrangian invariant.

These local transformations are called $U(1)$ gauge transformations. At each spacetime point the transformations are elements of the $U(1)$ group, where $U(1)$ refers to one dimensional unitary transformations, in other words, multiplication by a phase. A_μ is called the gauge field and it appears in its own kinetic term

$$-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, \tag{6.1.5}$$

where $F_{\mu\nu}$ is the field strength,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{6.1.6}$$

The six components of the electric and magnetic fields sit in the six nonzero entries of $F_{\mu\nu}$, with

$$E_i = F_{0i} \qquad B_i = -\frac{1}{2}\varepsilon_{ijk}F_{jk}\,. \tag{6.1.7}$$

$F_{\mu\nu}$ is invariant under the gauge transformation, hence so are $\vec{E}$, $\vec{B}$ and the kinetic term. The parameter e is the gauge coupling. It is sometimes convenient to make a field redefinition $A_\mu \to \frac{1}{e}A_\mu$, in which case the coupling only appears as an overall $\frac{1}{e^2}$ factor in front of the kinetic term.

If we assume that the gauge field couples to a Dirac fermion field, then the Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\overline{\psi}(\not{D} - m)\psi\,, \tag{6.1.8}$$

and the equations of motion for the gauge field are

$$\begin{aligned} &\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu A_\nu} - \frac{\partial\mathcal{L}}{\partial A_\nu} = 0 \\ \Rightarrow \qquad &-\partial_\mu F^{\mu\nu} + j^\nu = 0\,, \end{aligned} \tag{6.1.9}$$

where j^ν is the electromagnetic current,

$$j^\nu = e\,\overline{\psi}\gamma^\nu\psi\,. \tag{6.1.10}$$

Hence we see that this is the vector current we discussed previously, multiplied by the coupling. Using results in the fourth chapter, we find that

$$[Q, a^{s\dagger}_{\vec{k}}] = e\,a^{s\dagger}_{\vec{k}} \qquad [Q, b^{s\dagger}_{\vec{k}}] = -e\,b^{s\dagger}_{\vec{k}}\,, \tag{6.1.11}$$

where $Q = \int d^3x j^0$. Therefore, we see that the fermions have one unit of electric charge e, while the antifermions have electric charge $-e$. Peskin uses $e = -|e|$, that is the fermion has negative electric charge and hence can be identified with the electron[1]. The anti-fermion has positive electric charge and is the positron. Other textbooks use the opposite convention. To cause as little confusion as possible, we will stick with the Peskin convention.

[1] It can also be a muon or tau particle, but we will usually assume that it is the electron unless stated otherwise.

6.1.2 Quantization under Lorenz gauge

If we assume that the currents are zero, then the equations of motion for A_μ become

$$\begin{aligned} \partial^2 A_\nu - \partial_\nu \partial_\mu A^\mu &= 0 \\ \Rightarrow \quad (\eta_{\nu\mu}\partial^2 - \partial_\nu\partial_\mu)A^\mu &= 0\,. \end{aligned} \tag{6.1.12}$$

The physics should be invariant under gauge transformations, so we can "choose a gauge" where the equations simplify. For instance, we can always make a gauge transformation such that $\partial_\mu A^\mu = 0$ after the transformation; one chooses the gauge parameter θ such that $\partial^2\theta = e\,\partial_\mu A^\mu$. This is called the Lorenz gauge[2]. The nice thing about the Lorenz gauge is that it keeps the Lorentz invariance manifest.

We want to preserve the Lorentz covariance after quantizing the gauge fields. This means that we want to find a set of a canonical momenta $\Pi^\mu(x)$, such that

$$[A_\mu(\vec{x},t), \Pi^\nu(\vec{y},t)] = i\delta_\mu{}^\nu \delta^3(\vec{x}-\vec{y})\,. \tag{6.1.13}$$

However, the canonical momenta are given by

$$\Pi^\nu(x) = \frac{\partial \mathcal{L}}{\partial \partial_0 A_\nu} = -F^{0\nu}\,. \tag{6.1.14}$$

Therefore,

$$\Pi^j(x) = E_j(x)\,, \qquad \Pi^0(x) = 0\,, \tag{6.1.15}$$

and so it seems that we cannot construct the above commutation relation for A_0 because we do not have a Π^0.

We can circumvent this difficulty by adding the additional "gauge fixing" term to the Lagrangian,

$$\mathcal{L}_{\rm gf} = -\frac{1}{2}\left(\partial_\mu A^\mu\right)^2\,. \tag{6.1.16}$$

After integrating by parts this becomes

$$\mathcal{L}_{\rm gf} = -\frac{1}{2}\left(\partial_0 A_0\right)\left(\partial^0 A^0\right) - \partial_0 A_i \partial^i A^0 - \frac{1}{2}\left(\partial_i A^i\right)^2\,. \tag{6.1.17}$$

The full Lagrangian for the gauge fields is then

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{L}_{\rm gf} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu\,, \tag{6.1.18}$$

and so the canonical momenta have the more standard form

$$\Pi^\mu(x) = \frac{\partial \mathcal{L}}{\partial \partial_0 A_\mu} = -\partial^0 A^\mu(x)\,. \tag{6.1.19}$$

[2] Ludvig *Lorenz* (1829-1891) was a Danish physicist and should not be confused with Hendrik *Lorentz* (1853-1928) who was Dutch. Many older textbooks (including Peskin) have called it the Lorentz gauge, which is now thought to be a typo that promulgated through history. Lorenz introduced his gauge soon after Maxwell introduced his equations, when Lorentz was only 14 years old.

Furthermore, the equations of motion from $\mathcal{L}$ with $\mathcal{L}_{\rm gf}$ included are now

$$\partial^2 A_\nu = 0\,, \tag{6.1.20}$$

the equations one finds in the Lorenz gauge. Thus $\mathcal{L}_{\rm gf}$ fixes the gauge to the Lorenz gauge. The gauge-fixing term is not gauge invariant, but this is expected; it cannot be gauge invariant if its effect is to pick a particular gauge. Let us look at this another way. The equations of motion are found by extremizing the action under variations of A_μ. Suppose we restrict the variations in A_μ to pure gauge transformations, $\delta A_\mu = \delta\widetilde{A}_\mu = \partial_\mu\delta\theta$. $F_{\mu\nu}$ is independent of such transformations, so the variation of the action is

$$\delta S = \int d^4x(-\partial_\mu A^\mu)\partial_\nu\delta\widetilde{A}^\nu = \int d^4x(-\partial_\nu\partial_\mu A^\mu)\delta\widetilde{A}^\nu\,. \tag{6.1.21}$$

Hence we get that $\partial_\mu A^\mu$ is a constant if $\delta S = 0$. If we assume that A_μ is not divergent as $x^\nu \to \infty$, then the constant is zero.

From the equations of motion in (6.1.20) we see that each component of A_μ satisfies a massless Klein-Gordon equation. Using this as well as the commutation relations in (6.1.13) and the canonical momenta in (6.1.19) we can go straight to the second quantization for the gauge field,

$$A_\mu(x) = \int \frac{d^3k}{(2\pi)^3\,2\,k^0}\left(a_{\mu\,\vec{k}}\,e^{-ik\cdot x} + a^\dagger_{\mu\,\vec{k}}\,e^{+ik\cdot x}\right)\,, \tag{6.1.22}$$

where $a_{\mu\,\vec{k}}$ and $a^\dagger_{\mu\,\vec{k}}$ are the annihilation and creation operators for the gauge field. From the equations of motion we have that $k^0 = |\vec{k}|$. The commutation relations for the annihilation and creation operators read

$$[a_{\mu\,\vec{k}}, a^\dagger_{\nu\,\vec{k}'}] = -2k^0(2\pi)^3\delta^3(\vec{k}-\vec{k}')\eta_{\mu\nu}\,. \tag{6.1.23}$$

Note that the commutation relations for the time component have the opposite sign as that for a real scalar field. This is because the kinetic term for A_0 in (6.1.18) has the opposite sign.

We can then construct the Hilbert space as a Fock space where $a_{\mu\vec{k}}|0\rangle = 0$ and all other states are generated by $a^\dagger_{\mu\vec{k}}$. But the presence of $\eta_{\mu\nu}$ in the commutation relations raises some issues. Most worrisome are the existence of negative normed states, for example, the state $a^\dagger_{0\,\vec{k}}|0\rangle$, which would seem to be a complete disaster. The resolution of this difficulty is that the Hilbert space contains not only the physical states, but nonphysical states as well. In particular, the physical states satisfy the gauge condition,

$$\langle \text{phys}|\partial_\mu A^\mu(x)|\text{phys}\rangle = 0\,. \tag{6.1.24}$$

Since this is true for any physical state, we can rewrite the gauge condition as

$$\langle \text{phys}'|\partial_\mu A^\mu(x)|\text{phys}\rangle = 0\,. \tag{6.1.25}$$

for any two physical states. We then write a state in the Fock space as

$$|\vec{k}_1, \xi_1; \vec{k}_2, \xi_2; \ldots \vec{k}_n, \xi_n\rangle = \prod_{j=1}^{n} \xi_j^\mu a^\dagger_{\mu \vec{k}_j} |0\rangle\,, \tag{6.1.26}$$

where the ξ_j^μ are the polarization vectors. We then see that gauge condition (6.1.25) requires

$$k_j \cdot \xi_j = 0\,. \tag{6.1.27}$$

This then gets rid of the negative normed states.

However, there are still zero normed states which survive the condition in (6.1.27). This will happen if one or more of the polarizations satisfies $\xi_j^\mu \sim k_j^\mu$, in which case $\xi_j \cdot \xi_j = 0$. These states are here because the Lorenz gauge condition does not completely fix the gauge; we can make the gauge transformation $A_\mu \to A_\mu - \frac{1}{e}\partial_\mu\theta$ and still preserve the Lorenz condition if $\partial^2\theta = 0$. The particles with zero norm are the quanta of this pure gauge field. Hence, if we start with a physical state and then perform a gauge transformation, we get back the original state plus a state with zero norm. It is in this way that physical measurements like expectation values *etc.* remain gauge invariant.

This leaves two independent polarizations for ξ^μ, both of which are transverse to k^μ. It is convenient to write the polarization vectors as $\xi_\mu^{(\lambda)}$ and to write the gauge fields as

$$A_\mu(x) = \sum_{\lambda=1,2} \int \frac{d^3k}{(2\pi)^3\, 2\, k^0}\, \xi_\mu^{(\lambda)}(\vec{k}) \left(a_{\vec{k}}^{(\lambda)}\, e^{-ik\cdot x} + a_{\vec{k}}^{(\lambda)\dagger}\, e^{+ik\cdot x} \right)\,, \tag{6.1.28}$$

where the sum over λ corresponds to the two photon polarizations. The commutation relations are

$$[a_{\vec{k}}^{(\lambda)}, a_{\vec{k}'}^{(\lambda')\dagger}] = 2k^0(2\pi)^3\delta^3(\vec{k}-\vec{k}')\delta^{\lambda\lambda'}\,, \tag{6.1.29}$$

and the reduced Fock space constructed from the $a_{\vec{k}}^{(\lambda)\dagger}$ has positive definite norm. The polarization vectors not only satisfy $k \cdot \xi^{(\lambda)} = 0$, but can be chosen to be transverse in the spatial components as well, $\vec{k} \cdot \vec{\xi}^{(\lambda)} = 0$.

It is not enough to reduce the states to the positive normed states. We will also need to show that once interactions are included, states with zero or negative norm will not be produced. It turns out that such occurrences are prevented by a feature found in all gauge theories called a Ward identity, which we will come back to in chapters 7 and 8.

6.1.3 The gauge field propagator

In the free theory we can consider the time ordered correlator

$$G_{\mu\nu}(k) = \int d^4x\, e^{ik\cdot x} \langle T[A_\mu(x)A_\nu(0)]\rangle\,. \tag{6.1.30}$$

The correlator has a structure similar to the real scalar field, with an extra factor coming from the Lorentz indices. Using the commutation relations in (6.1.23), one finds that the correlator is

$$G_{\mu\nu}(k) = \frac{-i\eta_{\mu\nu}}{k^2 + i\,\epsilon}\,. \tag{6.1.31}$$

However, there are other gauge equivalent forms of the correlator. To see this, we define the projector,

$$P_{\mu\nu} = \eta_{\mu\nu} - \partial^{-2}\partial_\mu\partial_\nu\,. \tag{6.1.32}$$

If we square $P_{\mu\nu}$ we find

$$P_{\mu\nu}P^\nu{}_\lambda = \eta_{\mu\lambda} - 2\,\partial^{-2}\partial_\mu\partial_\lambda + \partial^{-2}\partial_\mu\partial_\lambda = P_{\mu\lambda}\,, \tag{6.1.33}$$

demonstrating that it is a projector. We then let $\widetilde{A}_\mu(x) = P_{\mu\nu}A^\nu(x)$, where it is clear that

$$\partial_\mu\widetilde{A}^\mu = \partial_\mu P^\mu{}_\nu A^\nu = \partial_\mu A^\mu - \partial_\nu A^\nu = 0\,, \tag{6.1.34}$$

hence satisfying the Lorenz gauge condition. If we replace $A_\mu(x)$ with $\widetilde{A}_\mu(x)$ in the correlator, we then find

$$G_{\mu\nu}(k) = \frac{-i(\eta_{\mu\nu} - k_\mu k_\nu/k^2)}{k^2 + i\,\epsilon}\,. \tag{6.1.35}$$

where $\eta_{\mu\nu} - k_\mu k_\nu/k^2$ is the projector in momentum space. If we also replace $A_\nu(0)$ with $\widetilde{A}_\mu(0)$ then we get the same thing since $P_{\mu\nu}$ is a projector.

It is important to note that whichever correlator we use, one gets the same result for $\xi^\mu G_{\mu\nu}(k)$ if $\xi \cdot k = 0$. In fact, we could have used

$$G_{\mu\nu}(k) = \frac{-i(\eta_{\mu\nu} - (1-\lambda)k_\mu k_\nu/k^2)}{k^2 + i\,\epsilon}\,. \tag{6.1.36}$$

where λ is any real number. The case where $\lambda = 1$ is called Feynman gauge (although it is not really a gauge choice), while $\lambda = 0$ is called Landau gauge.

We can derive this last expression from a path integral. If we go back to (6.1.21), we see that the overall coefficient in front of $\mathcal{L}_{\rm gf}$ does not matter as far as fixing to Lorenz gauge. Hence, if we now let

$$\mathcal{L}_{\rm gf} = -\frac{1}{2\lambda}(\partial_\mu A^\mu)^2\,, \tag{6.1.37}$$

then the new action is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{L}_{\rm gf} = -\frac{1}{2}\partial_\mu A_\nu\partial^\mu A^\nu + \frac{1}{2}\left(1 - \frac{1}{\lambda}\right)\partial_\mu A^\mu\partial_\nu A^\nu\,. \tag{6.1.38}$$

The path integral with a background source $j^\mu(x)$ is

$$\mathcal{Z}(j) = \int \mathcal{D}A \exp\left(i\int d^4x\mathcal{L} + A_\mu j^\mu\right)\,, \tag{6.1.39}$$

where j^μ has the same coupling to A_μ as the electromagnetic current. Since $\mathcal{L}$ is quadratic in A_μ, the path integral is gaussian and can be computed. Integrating by parts, we can write $\mathcal{L}$ as

$$\mathcal{L} = \frac{1}{2}A_\mu\left(\eta^{\mu\nu}\partial^2 - \left(1-\frac{1}{\lambda}\right)\partial^\mu\partial^\nu\right)A_\nu\,. \tag{6.1.40}$$

Completing the square, the path integral becomes

$$\mathcal{Z}(j) = \mathcal{Z}(0)\exp\left(-\frac{1}{2}\int d^4xd^4y\, j^\mu(x)G_{\mu\nu}(x-y)j^\nu(y)\right)\,, \tag{6.1.41}$$

where $G_{\mu\nu}(x-y)$ is the Green's function satisfying

$$\left(\eta^{\mu\nu}\partial^2 - \left(1-\frac{1}{\lambda}\right)\partial^\mu\partial^\nu\right)G_{\nu\lambda}(x-y) = i\,\delta^\mu{}_\lambda\delta^4(x-y)\,. \tag{6.1.42}$$

Converting to momentum space, the Green's function satisfies

$$\left(-\eta^{\mu\nu}k^2 + \left(1-\frac{1}{\lambda}\right)k^\mu k^\nu\right)G_{\nu\lambda}(k) = i\,\delta^\mu{}_\lambda\,, \tag{6.1.43}$$

which has the solution

$$G_{\mu\nu}(k) = \frac{-i(\eta_{\mu\nu}-(1-\lambda)k_\mu k_\nu/k^2)}{k^2}\,. \tag{6.1.44}$$

The time-ordered correlator is then

$$\langle T[A_\mu(x)A_\nu(y)]\rangle = \mathcal{Z}^{-1}(0)\frac{-i\delta}{\delta j^\mu(x)}\frac{-i\delta}{\delta j^\nu(y)}\mathcal{Z}(j)\bigg|_{j^\lambda=0} = G(x-y)\,, \tag{6.1.45}$$

which agrees with (6.1.35).

Note that $\lambda = \infty$ leads to a singular propagator. This can be understood as follows: when $\lambda = \infty$, $\frac{1}{\lambda} = 0$ and there is no gauge fixing term in the Lagrangian. The Lagrangian is then invariant under pure gauge transformations meaning that the operator $\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu$ has some of its eigenvalues equal to zero (these are called zero-modes), making it noninvertible. We can then only make sense of the path integral if we restrict the integrations over A_μ to a particular "gauge slice". This removes the zero-modes. More details can be found in chapter 9 of Peskin.

We close this section by noting that the path integral can also be written as

$$\mathcal{Z}(j) = \mathcal{Z}(0)\exp\left(-\frac{1}{2}\int\frac{d^4k}{(2\pi)^4}\,j^\mu(-k)\,\frac{-i(\eta_{\mu\nu}-(1-\lambda)k_\mu k_\nu/k^2)}{k^2}\,j^\nu(k)\right)\,. \tag{6.1.46}$$

The source terms in (6.1.39) are gauge invariant only if the current is conserved, $\partial_\mu j^\mu(x) = 0$. Hence, for physical electromagnetic sources we have that $k_\mu j^\mu(k) = 0$ and the path integral is independent of λ.

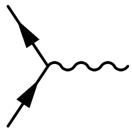
6.2 The Feynman rules in QED

6.2.1 The rules

The symbol for the photon propagator is the wavy line:

$$\mu \quad \nu = \frac{-i\eta_{\mu\nu}}{k^2+i\epsilon}\,, \tag{6.2.1}$$

where we use Feynman gauge with $\lambda = 1$ to make it as simple as possible. The μ and ν are the incoming and outgoing components of the gauge field. The gauge field couples to the Dirac fields through the covariant derivative in (6.1.3). The vertex contribution to the Feynman diagrams from this term is



$$\mu = -i\, e\, \gamma^\mu\, (2\pi)^4 \delta^4(k_1+k_2+k_3) \tag{6.2.2}$$

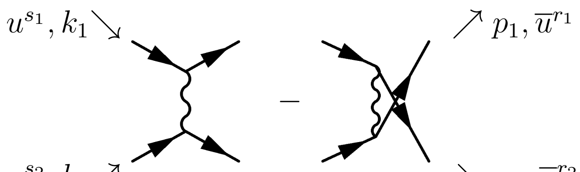
where the k_i are the three incoming momenta. To these rules as well as the previous rules for fermions we also add

- Multiply by a polarization vector $\xi_\mu^{(\lambda)}(\vec{k})$ for each external photon line.
- Sum over contracted Lorentz indices.

Let us now apply these rules to two of the simpler tree-level processes.

6.2.2 Electron-electron elastic scattering

The first is electron-electron elastic scattering. The Feynman diagrams for this process are similar to our fermion-fermion scattering with scalar exchange, and are given by



$$\begin{aligned}
&= (2\pi)^4\delta^4(k_1+k_2-p_1-p_2)(-ie)^2 \\
&\times\left(\frac{-i\,\eta_{\mu\nu}\,\overline{u}^{r_1}(\vec{p}_1)\gamma^\mu u^{s_1}(\vec{k}_1)\overline{u}^{r_2}(\vec{p}_2)\gamma^\nu u^{s_2}(\vec{k}_2)}{(p_1-k_1)^2+i\epsilon} - \frac{-i\eta_{\mu\nu}\,\overline{u}^{r_1}(\vec{p}_1)\gamma^\mu u^{s_2}(\vec{k}_2)\overline{u}^{r_2}(\vec{p}_2)\gamma^\nu u^{s_1}(\vec{k}_1)}{(p_1-k_2)^2+i\epsilon}\right) \\
&= (2\pi)^4\delta^4(k_1+k_2-p_1-p_2)\langle\vec{p}_1,+,r_1;\vec{p}_2,+,r_2|i\mathcal{M}|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle\,.
\end{aligned} \tag{6.2.3}$$

Instead of doing the entire calculation, let us pick out the terms that differ from the Yukawa case and change these accordingly. Of course, we have that g is replaced with e and now there are also extra γ-matrices. We also replace the scalar mass M with zero.

The cross-section is still given by

$$\begin{aligned}\sigma_{e^-e^+\to e^-e^+} &= \frac{1}{4\sqrt{s(s-4m^2)}}\sum_{r_1,r_2}\int\left(\frac{d^3p_1}{(2\pi)^3 2p_1^0}\right)\left(\frac{d^3p_2}{(2\pi)^3 2p_2^0}\right)(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\\ &\qquad\times\left|\langle\vec p_1,+,r_1;\vec p_2,+,r_2|\mathcal{M}|\vec k_1,+,s_1;\vec k_2,+,s_2\rangle\right|^2 .\\ &= \frac{1}{64\pi^2\, s}\int_{\rm fwd} d\Omega\sum_{r_1,r_2}\left|\langle\vec p_1,+,r_1;\vec p_2,+,r_2|\mathcal{M}|\vec k_1,+,s_1;\vec k_2,+,s_2\rangle\right|^2\end{aligned} \tag{6.2.4}$$

When squaring the $\mathcal{M}$-matrix we encounter three types of terms from the numerators. The first is

$$\begin{aligned}&\sum_{r_1,r_2}\left|\overline{u}^{r_1}(\vec p_1)\gamma^\mu u^{s_1}(\vec k_1)\overline{u}^{r_2}(\vec p_2)\gamma_\mu u^{s_2}(\vec k_2)\right|^2\\ &\qquad = \overline{u}^{s_1}(\vec k_1)\gamma^\mu(\not p_1+m)\gamma^\nu u^{s_1}(\vec k_1)\overline{u}^{s_2}(\vec k_2)\gamma_\mu(\not p_2+m)\gamma_\nu u^{s_2}(\vec k_2)\,.\end{aligned} \tag{6.2.5}$$

We assume that the incoming electrons are unpolarized, meaning that we will average over the spins. The spin averaging then gives us

$$\begin{aligned}&\frac14\sum_{s_1,s_2}\overline{u}^{s_1}(\vec k_1)\gamma^\mu(\not p_1+m)\gamma^\nu u^{s_1}(\vec k_1)\overline{u}^{s_2}(\vec k_2)\gamma_\mu(\not p_2+m)\gamma_\nu u^{s_2}(\vec k_2)\\ &\qquad = \frac14{\rm Tr}[(\not k_1+m)\gamma^\mu(\not p_1+m)\gamma^\nu]{\rm Tr}[(\not k_2+m)\gamma_\mu(\not p_2+m)\gamma_\nu]\end{aligned} \tag{6.2.6}$$

Using the rules for traces of γ-matrices we have

$$ {\rm Tr}[(\not k+m)\gamma^\mu(\not p+m)\gamma^\nu] = 4\left(\eta^{\mu\nu}(m^2-p\cdot k)+p^\mu k^\nu+k^\mu p^\nu\right)\,. \tag{6.2.7}$$

Hence we have that the expression in (6.2.6) is

$$4\left(t^2+t(2m^2-t)+\frac12(s-2m^2)^2+\frac12(u-2m^2)^2\right)=2t^2-4us+16m^4\,. \tag{6.2.8}$$

By symmetry, the second type of term after averaging the incoming spins is

$$\frac14\sum_{r_1,r_2,s_1,s_2}\left|\overline{u}^{r_2}(\vec p_2)\gamma^\mu u^{s_1}(\vec k_1)\overline{u}^{r_1}(\vec p_1)\gamma_\mu u^{s_2}(\vec k_2)\right|^2 = 2u^2-4ts+16m^4\,. \tag{6.2.9}$$

Finally, the spin averaged cross terms are

$$\begin{aligned}&\frac14\sum_{r_1,r_2,s_1,s_2}\overline{u}^{r_1}(\vec p_1)\gamma^\mu u^{s_1}(\vec k_1)\overline{u}^{r_2}(\vec p_2)\gamma_\mu u^{s_2}(\vec k_2)\overline{u}^{s_2}(\vec k_2)\gamma_\nu u^{r_1}(\vec p_1)\overline{u}^{s_1}(\vec k_1)\gamma^\nu u^{r_2}(\vec p_2)+\text{ c.c.}\\ &\qquad = \frac14{\rm Tr}[(\not p_1+m)\gamma^\mu(\not k_1+m)\gamma^\nu(\not p_2+m)\gamma_\mu(\not k_2+m)\gamma_\nu]+\text{ c.c.}\,.\end{aligned} \tag{6.2.10}$$

To simplify this expression we require the following identities

$$\begin{aligned}&\gamma^\mu\gamma^\nu\gamma_\mu = -\gamma^\nu\gamma^\mu\gamma_\mu+2\eta^{\mu\nu}\gamma_\mu = -4\gamma^\nu+2\gamma^\nu=-2\gamma^\nu\\ &\gamma^\mu\gamma^\nu\gamma^\lambda\gamma_\mu = -\gamma^\nu\gamma^\mu\gamma^\lambda\gamma_\mu+2\eta^{\mu\nu}\gamma^\lambda\gamma_\mu = 2\gamma^\nu\gamma^\lambda+2\gamma^\lambda\gamma^\nu = 4\eta^{\nu\lambda}\\ &\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma\gamma_\mu = -\gamma^\nu\gamma^\mu\gamma^\lambda\gamma^\sigma\gamma_\mu+2\eta^{\mu\nu}\gamma^\lambda\gamma^\sigma\gamma_\mu = -4\gamma^\nu\eta^{\lambda\sigma}+2\gamma^\lambda\gamma^\sigma\gamma^\nu = -2\gamma^\sigma\gamma^\lambda\gamma^\nu\,.\end{aligned} \tag{6.2.11}$$

Hence we have that

$$\gamma^\mu(\not{k}_1+m)\gamma^\nu(\not{p}_2+m)\gamma_\mu = -2m^2\gamma^\nu + 4m(k_1^\nu+p_2^\nu) - 2\not{p}_2\gamma^\nu\not{k}_1\,. \tag{6.2.12}$$

We then use

$$\begin{aligned}
&\frac{1}{4}\mathrm{Tr}[(\not{p}_1+m)(-2m^2\gamma^\nu)(\not{k}_2+m)\gamma_\nu] = 4m^2p_1\cdot k_2 - 8m^4 = -2u\,m^2 - 4m^4\\
&\frac{1}{4}\mathrm{Tr}[(\not{p}_1+m)(\not{k}_2+m)4m(\not{k}_1+\not{p}_2)] = 4m^2(p_1+k_2)\cdot(p_2+k_1) = 4m^2(s-t)\\
&\frac{1}{4}\mathrm{Tr}[(\not{p}_1+m)(-2\not{p}_2\gamma^\nu\not{k}_1)(\not{k}_2+m)\gamma_\nu]\\
&\qquad = -8p_1\cdot p_2k_1\cdot k_2 + 4m^2k_1\cdot p_2 = -2(s-2m^2)^2 - 2m^2u + 4m^4\,.
\end{aligned} \tag{6.2.13}$$

to find

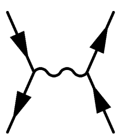
$$\frac{1}{4}\mathrm{Tr}[(\not{p}_1+m)\gamma^\mu(\not{k}_1+m)\gamma^\nu(\not{p}_2+m)\gamma_\mu(\not{k}_2+m)\gamma_\nu] + \text{ c.c.} = -4(t+u)^2 + 16m^4\,. \tag{6.2.14}$$

Putting everything together, the square of the $\mathcal{M}$-matrix, including the sum over the final spins and spin averaging for the initial spins, is

$$\begin{aligned}
&\frac{1}{4}\sum_{r_1,r_2,s_1,s_2}\left|\langle\vec{p}_1,+,r_1;\vec{p}_2,+,r_2|\mathcal{M}|\vec{k}_1,+,s_1;\vec{k}_2,+,s_2\rangle\right|^2\\
&\qquad = e^4\left(\frac{2t^2-4us+16m^4}{(t+i\epsilon)^2} + \frac{2u^2-4ts+16m^4}{(u+i\epsilon)^2} + \frac{4(t+u)^2-16m^4}{(t+i\epsilon)(u+i\epsilon)}\right)\,.
\end{aligned} \tag{6.2.15}$$

6.2.3 $e^-e^+ \to \mu^-\mu^+$

We next consider electron-positron annihilation into a muon-antimuon pair. The reason why we consider this particular process is that it is simpler than $e^-e^+\to e^-e^+$ because there is no t-channel photon exchange. The only diagram is



$$= (2\pi)^4\delta^4(k_1+k_2-p_1-p_2)(-ie)^2\frac{-i\,\overline{v}^{s_2}(\vec{k}_2)\gamma^\mu u^{s_1}(\vec{k}_1)\overline{u}^{r_1}(\vec{p}_1)\gamma_\mu v^{r_2}(\vec{p}_2)}{(k_1+k_2)^2+i\epsilon} \tag{6.2.16}$$

Proceeding as before, the differential cross-section for unpolarized electrons and positrons in the COM frame is

$$\sigma_{e^-e^+\to\mu^-\mu^+} = \frac{e^4}{64\pi^2\,s}\sqrt{\frac{s-4M^2}{s-4m^2}}\int d\Omega\,\frac{\frac{1}{4}\mathrm{Tr}[(\not{k}_1+m)\gamma^\mu(\not{k}_2-m)\gamma^\nu]\mathrm{Tr}[(\not{p}_1+M)\gamma_\mu(\not{p}_2-M)\gamma_\nu]}{(s+i\epsilon)^2}\,, \tag{6.2.17}$$

where M is the muon mass and the integration is over the entire 4π of the solid angle since the outgoing particles are not identical. The additional square root prefactor arises because the magnitude of the spatial momentum is not the same for the incoming and outgoing particles. The muon is 200 times heavier than the electron, so we can set $m = 0$ for the rest of the calculation, since $s > 4M^2 >> 4m^2$. Hence we have

$$\begin{aligned}
&\frac{1}{4}\mathrm{Tr}[(\not{k}_1+m)\gamma^\mu(\not{k}_2-m)\gamma^\nu]\mathrm{Tr}[(\not{p}_1+M)\gamma_\mu(\not{p}_2-M)\gamma_\nu] \\
&\quad\approx 4(k_1^\mu k_2^\nu - k_1\cdot k_2\eta^{\mu\nu} + k_1^\nu k_2^\mu)(p_{1\mu}p_{2\nu} - (p_1\cdot p_2 + M^2)\eta_{\mu\nu} + p_{1\nu}p_{2\mu}) \\
&\quad= 4(2k_1\cdot p_1 k_2\cdot p_2 + 2k_1\cdot p_2 k_2\cdot p_1 + 2M^2 k_1\cdot k_2) \approx 2((t-M^2)^2 + (u-M^2)^2 + 2sM^2) \\
&\quad\approx 2(s^2 - 2ut + 2M^4) = 2s^2 + s(s-4M^2)\sin^2\theta\,, \qquad (6.2.18)
\end{aligned}$$

where θ is the angle between the direction of the μ^- and the e^-. In deriving this we used that

$$\begin{aligned}
s &\approx 2k_1\cdot k_2 \approx 2p_1\cdot p_2 + 2M^2 \\
t &\approx M^2 - 2k_1\cdot p_1 \approx M^2 - \frac{1}{2}(s - \sqrt{s}\sqrt{s-4M^2}\cos\theta) \\
u &\approx M^2 - 2k_1\cdot p_2 \approx M^2 - \frac{1}{2}(s + \sqrt{s}\sqrt{s-4M^2}\cos\theta) \\
2M^2 &\approx s + t + u\,. \qquad (6.2.19)
\end{aligned}$$

The final result for the differential cross-section is then

$$\frac{d\sigma_{e^-e^+\to\mu^-\mu^+}}{d\Omega} = \frac{e^4\sqrt{s-4M^2}}{64\pi^2\, s^{5/2}}\,(2s + (s-4M^2)\sin^2\theta)\,. \qquad (6.2.20)$$

Chapter 7

QED Continued

In this chapter we continue our study of QED. We first discuss the photon helicity where we argue that there are ± 1 helicity states but not a 0 helicity state. We show that this is consistent for a Lorentz invariant theory by proving that the helicity of a massless particle is a Poincaré invariant. We then introduce the Schwinger-Dysan equation and use it to derive the Ward identity. Next we show how the Ward identity can be used to simplify amplitudes with external photons. We then consider examples of amplitudes with external photons.

7.1 The photon helicity

In the previous chapter of the notes we argued that the photon has two physical polarizations. In this section we show that the independent polarizations have helicities ± 1. Recall that the helicity is the spin component parallel to a particle's spatial momentum.

A photon polarization vector is given by $\xi_\mu^{(\lambda)}(k)$ where λ refers to one of the two physical polarizations. The polarization vector satisfies $\xi^{(\lambda)} \cdot k = 0$. Moreover, under a Lorentz transformation it transforms as $\xi_\mu^{(\lambda)}(k) \to \xi'^{(\lambda)}_\mu(k) = \Lambda^{-1\nu}{}_\mu \xi_\nu^{(\lambda)}(\Lambda^{-1}k)$. Let us now suppose that k^μ is given by $(k, 0, 0, k)$ and that the physical polarizations satisfy $\vec{\xi}^{(\lambda)} \cdot \vec{k} = 0$. We then consider a Lorentz transformation which is a rotation about the $\hat{z}$ axis. Since $\vec{\xi}^{(\lambda)}$ is a three-vector orthogonal to $\hat{z}$, the two independent components can be chosen to be eigenvectors of L_z with eigenvalues ± 1. The $+$ component is right circularly polarized, while the $-$ component is left circularly polarized. In other words, the $+$ polarization corresponds to a photon with helicity $+1$ while the $-$ polarization corresponds to a photon with helicity -1.

You now might be worried that there are only two components for the polarization and not three. The spin 1 representation of the rotation group contains three states, with L_z eigenvalues $+1, 0, -1$. The component with $L_z = 0$ would be the vector pointing in the $\hat{z}$ direction (the longitudinal direction), which the photon polarizations are missing since it does not satisfy $\xi \cdot k = 0$. The reason why this is not a problem is because the photon is massless. Let us show that if the photon did have a mass, it would be necessary to have a spin 0 component in a Lorentz invariant theory. Assuming the photon had a mass, it would be possible to boost in the $\hat{z}$ direction to a frame where the photon is at rest. Then we could rotate about the $\hat{x}$ direction, which will not change $\vec{k}$ since it is now zero. However, this will transform the $L_z = 1$ polarization into a linear combination of polarizations which includes the $L_z = 0$ polarization.

Let us now make these arguments more precise. The full set of space-time invariances for a translationally and Lorentz invariant quantum field theory is the Poincaré group, generated by P_μ and $J_{\mu\nu}$. These generators satisfy the commutation relations

$$\begin{aligned} [P_\mu, P_\nu] &= 0 \\ [J_{\mu\nu}, P_\lambda] &= -i(\eta_{\mu\lambda}P_\nu - \eta_{\nu\lambda}P_\mu) \\ [J_{\mu\nu}, J_{\lambda\sigma}] &= i(\eta_{\mu\sigma}J_{\nu\lambda} - \eta_{\mu\lambda}J_{\nu\sigma} - \eta\nu\sigma J_{\mu\lambda} + \eta_{\nu\lambda}J_{\mu\sigma}) . \end{aligned} \tag{7.1.1}$$

This algebra has two independent *Casimir* operators, operators that commute with every generator in the Poincaré algebra. Obviously, the Casimirs must be Lorentz invariant in order to commute with $J_{\mu\nu}$. Clearly P^2 is a Casimir since it also commutes with P_μ, but the other Casimir is less well known. To describe this one, we define the *Pauli-Lubanski* vector W_μ,

$$W_\mu \equiv \frac{1}{2}\epsilon_{\mu\nu\lambda\sigma}P^\nu J^{\lambda\sigma} . \tag{7.1.2}$$

We can easily check that P_μ commutes with W_ν,

$$[P_\mu, W_\nu] = \frac{1}{2}\epsilon_{\nu\rho\lambda\sigma}P^\rho[P_\mu, J^{\lambda\sigma}] = +i\,\epsilon_{\nu\rho\mu\sigma}P^\rho P^\sigma = 0 , \tag{7.1.3}$$

therefore $W^2 = W_\mu W^\mu$ is also a Casimir, which is clearly independent from P^2 since it contains Lorentz generators. In terms of the Poincaré generators it is given by

$$W^2 = -\frac{1}{2}(J_{\mu\nu}J^{\mu\nu}P^2 - 2J^{\mu\lambda}J_{\nu\lambda}P_\mu P^\nu) . \tag{7.1.4}$$

The irreducible representations of the Poincaré group are labeled by the values of the Casimirs. The trivial representation is the vacuum $|0\rangle$ which has $P_\mu|0\rangle = J_{\mu\nu}|0\rangle = 0$. The first set of nontrivial representations start with the one particle states. Suppose we have a one particle state $|k^\mu, \sigma\rangle$ where k^μ is its 4-momentum and σ is an internal spin. Clearly, $P^2|k^\mu, \sigma\rangle = m^2|k^\mu, \sigma\rangle$ where m^2 is the invariant mass-squared. To find the other Casimir, we first boost to the particle's rest frame, $k^\mu_R = (m, 0, 0, 0)$. This will not change the value of W^2 since it commutes with all Poincaré generators. In the rest frame we have that

$$\begin{aligned} W^2|k^\mu_R, \sigma\rangle &= -\frac{1}{2}(J_{\mu\nu}J^{\mu\nu}m^2 - 2J^{0\lambda}J_{0\lambda}m^2)|k^\mu_R, \sigma\rangle \\ &= -\frac{m^2}{2}J_{ij}J^{ij}|k^\mu_R, \sigma\rangle = -m^2\vec{J}^2|k^\mu_R, \sigma\rangle = -m^2 s(s+1)|k^\mu_R, \sigma\rangle , \end{aligned} \tag{7.1.5}$$

where s is the total spin. Hence when acting on a particle at rest W^2 is proportional to the usual three-dimensional angular momentum invariant $\vec{J}^2$. When the particle is not at rest, W^2 is no longer proportional to $\vec{J}^2$, but the eigenvalue is the same.

It is clear that for massive particles P_μ is time-like while W_μ is space-like. However, for massless particles both 4-vectors are light-like since they are each proportional to m^2. Since $P_\mu W^\mu$ is identically zero, this means that $W^\mu = hP^\mu$ for massless particles, where h is a constant. As an operator statement we can write this as

$$(hP_\nu - W_\nu)|k^\mu, \sigma\rangle = 0 \tag{7.1.6}$$

This statement is invariant under Poincaré transformations since

$$\begin{aligned} [P_\mu, hP_\nu - W_\nu] &= 0 \\ [J_{\mu\lambda}, hP_\nu - W_\nu] &= -i\,\eta_{\mu\nu}(hP_\lambda - W_\lambda) + i\,\eta_{\lambda\nu}(hP_\mu - W_\mu)\,, \end{aligned} \tag{7.1.7}$$

hence the constant h is an invariant. If we choose a frame where $k^\mu = (k, 0, 0, k)$, then the time componenent of (7.1.6) gives

$$(hk + kJ^{12})|k^\mu, \sigma\rangle = 0\,. \tag{7.1.8}$$

Since $-J^{12}$ is the component of angular momentum along the $\hat{z}$ direction (see (4.1.8) in the fourth chapter), we see that h is the helicity. Therefore, a massless particle may have a single helicity in a Poincaré invariant theory. We already saw this happening for Weyl fermions where the massless fermions could be restricted to $h = +1/2$ only.

However, CPT invariance also forces us to include $-h$ helicity states since helicity changes sign under P but not under C and T. The photon is also its own charge conjugate, hence if helicity h is present there must also exist a massless photon with helicity $-h$. We have already seen that that $h = \pm 1$ for the photon. But $h = 0$ is not required by Poincaré invariance or by CPT.

It should be evident that if the photon is to only have two physical polarizations, it must remain massless, even when interactions are included. In order to make a massive photon, one must also introduce an additional degree of freedom to make the longitudinal polarization. Naively, one might think it is possible to add a mass simply by including the mass-term $\frac{1}{2}m^2 A_\mu A^\mu$. However, such a term will clearly break gauge invariance.

7.2 The Ward identity

Before showing the Ward identity[1], we first derive another equation called the Schwinger-Dyson equation. We will derive it explicitly only for a single real scalar field, but it is easily generalizable to a generic field.

Suppose we consider the correlator

$$\langle T[\phi(x_1)\phi(x_2)\ldots\phi(x_n)]\rangle = \mathcal{Z}^{-1}\int \mathcal{D}\phi\,\phi(x_1)\phi(x_2)\ldots\phi(x_n)e^{iS(\phi)}\,, \tag{7.2.1}$$

[1]In a previous version I had called this the Ward-Takahashi identity. There is some variation in the literature as to what to call the form of the identity discussed in this section, but here I will use Peskin's nomenclature.

Since all field configurations are integrated over, the path integral is invariant under an infinitesimal shift of the field variables $\phi(x) \to \phi(x) + \delta\phi(x)$. Hence we have that

$$\int \mathcal{D}\phi \left(i\frac{\delta S}{\delta\phi(x)} \right) \phi(x_1)\phi(x_2)\ldots\phi(x_n)e^{iS(\phi)} + \sum_{j=1}^{n} \delta^4(x-x_j)\phi(x_1)\ldots\cancel{\phi(x_j)}\ldots\phi(x_n)e^{iS(\phi)} = 0\,. \tag{7.2.2}$$

where the field under the slash is removed. This equation is called the Schwinger-Dyson (S-D) equation. The second set of terms with the δ-functions are known as contact terms. The derivative of the action gives

$$-\,\partial^2\phi(x) - m^2\phi(x) + \mathcal{L}'_{\rm int}(\phi(x)) \tag{7.2.3}$$

which is the lefthand side of the equations of motion (the righthand side being zero). Hence the S-D equation can be written as

$$\langle T[(-\partial^2\phi(x) - m^2\phi(x) + \mathcal{L}'_{\rm int}(\phi(x))\phi(x_1)\phi(x_2)\ldots\phi(x_n)]\rangle = \text{contact terms}\,, \tag{7.2.4}$$

which tells us the equations of motion are satisfied inside correlators, up to contact terms.

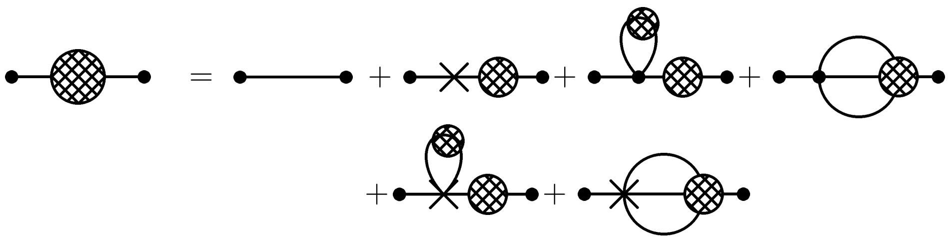
To get some idea about the information in the S-D equation, let us suppose that $n = 1$. We then Fourier transform the x and the x_1 coordinates to give

$$(2\pi)^4\delta^4(k+k_1)\left((k^2-m^2)G(k^2) + \int d^4x\, e^{ik\cdot x}\langle T[\mathcal{L}'_{\rm int}(\phi(x))\phi(0)]\rangle\right) = i\,(2\pi)^4\delta^4(k+k_1)\,, \tag{7.2.5}$$

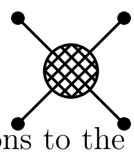
where the righthand side is the contribution of the contact term. Therefore,

$$G(k^2) = \frac{i}{k^2-m^2} + \frac{i}{k^2-m^2}\int d^4x\, e^{ik\cdot x}\langle T[i\,\mathcal{L}'_{\rm int}(\phi(x))\phi(0)]\rangle\,. \tag{7.2.6}$$

Since $\mathcal{L}_{\rm int}$ contains the counterterms, diagramatically (7.2.6) looks like



(7.2.7)

Recall from the third chapter that  contains not only the full correction to the vertex but also the full corrections to the propagators on its external legs. You can convince yourself of the veracity of (7.2.7) by considering the first few diagrams that contribute on the left and right sides of the equation.

Let us now apply the S-D equation to a scattering amplitude. We suppose that there are $n+1$ scalars involved in the amplitude which each have a momentum k_j (some of the $n+1$ are incoming and some are outgoing). The relevant $\mathcal{T}$-matrix is then

$$\int d^4x\, e^{ik\cdot x}(k^2-m^2)\int\prod_{j=1}^{n}d^4x_j\, e^{ik_j\cdot x_j}(k_j^2-m^2)\langle T[\phi(x)\phi(x_1)\phi(x_2)\dots\phi(x_n)]\rangle\,. \tag{7.2.8}$$

Integrating by parts twice on the x integration and using the S-D equation this becomes

$$\int d^4x\, e^{ik\cdot x}\int\prod_{j=1}^{n}d^4x_j\, e^{ik_j\cdot x_j}(k_j^2-m^2)\langle T[-\mathcal{L}_{\rm int}(\phi(x))\phi(x_1)\phi(x_2)\dots\phi(x_n)]\rangle + \text{ contact terms}\,. \tag{7.2.9}$$

We now claim that the contact terms cannot contribute to the scattering amplitude. Each contact term will be missing one of the external fields $\phi(x_j)$, hence the diagram will not have the propagator factor $\frac{1}{k_j^2-m^2-\Sigma(k_j^2)}$. But the numerator factor $k_j^2-m^2$ is still present and since the particles are on-shell, this will give zero.

Let us now apply these ideas to QED. If we have an external photon with polarization vector ξ_μ in a scattering amplitude, then its LSZ reduction to a $\mathcal{T}$-matrix element is

$$\langle f|\mathcal{T}|i\rangle = \int d^4x\, e^{ik\cdot x}\, k^2\xi^\mu\dots\langle T[A_\mu(x)\dots]\rangle\,, \tag{7.2.10}$$

where the dots signify all the other fields and integration variables and f and i stand for the initial and final states. We could have generalized $k^2\xi^\mu$ to $(k^2\eta^{\mu\nu}-(1-1/\lambda)k^\mu k^\nu)\xi_\nu$, but the second term does not contribute since $\xi\cdot k=0$. If we now integrate by parts twice the $\mathcal{T}$-matrix element becomes

$$\langle f|\mathcal{T}|i\rangle = \int d^4x\, e^{ik\cdot x}\,\xi^\mu\dots\langle T[-\partial^2A_\mu(x)\dots]\rangle\,, \tag{7.2.11}$$

If we choose the Lorenz gauge, then the equations of motion are $\partial^2A_\mu=j_\mu$, hence using the S-D equation we can replace the $\mathcal{T}$-matrix with

$$\langle f|\mathcal{T}|i\rangle = \int d^4x\, e^{ik\cdot x}\,\xi^\mu\dots\langle T[-j_\mu(x)\dots]\rangle\,, \tag{7.2.12}$$

since the contact terms do not contribute.

We can now show an important result. Suppose that $\xi^\mu\sim k^\mu$. We have already seen that such a photon would correspond to a state with zero norm. We then have that the $\mathcal{T}$-matrix is proportional to

$$\langle f|\mathcal{T}|i\rangle \sim \int d^4x\, e^{ik\cdot x}\,k^\mu\dots\langle T[-j_\mu(x)\dots]\rangle = \int d^4x\, e^{ik\cdot x}\,\dots\langle T[-i\partial^\mu j_\mu(x)\dots]\rangle\,, \tag{7.2.13}$$

where we again integrated by parts. However, gauge invariance requires that A_μ can only couple to a conserved current, therefore (7.2.13) is zero.

Going back to (7.2.10) we observe that we can write the $\mathcal{T}$-matrix element as

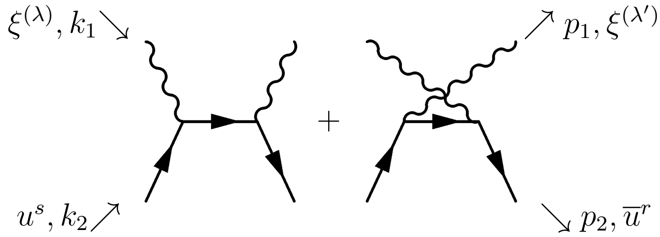
$$\langle f|\mathcal{T}|i\rangle = (2\pi)^4(P_{\text{out}} - P_{\text{in}})\langle f|\mathcal{M}|i\rangle = (2\pi)^4(P_{\text{out}} - P_{\text{in}})\xi_\mu\mathcal{M}^\mu\,. \tag{7.2.14}$$

Hence, we have shown that $k_\mu\mathcal{M}^\mu = 0$. This result is called the Ward identity and it means that the probability is zero that a scattering amplitude will produce a zero norm state. Therefore, it is consistent to reduce our Hilbert space to the physical reduced space with positive normed states. Since a gauge transformation leads to a shift in the polarization vector by a vector proportional to k^μ, the Ward identity also shows that amplitudes are invariant under gauge transformations. In the next chapter we will give a generalization of the Ward identity called the Ward-Takahashi identity which can be applied to correlators in general and not just those that appear in a $\mathcal{T}$-matrix where the external particles are on-shell.

7.3 Amplitudes with external photons

7.3.1 Compton scattering

Compton scattering is the scattering of a photon off of an electron (or any other charged particle). The relevant tree level Feynman diagrams are



$$= (2\pi)^4\delta^4(k_1+k_2-p_1-p_2)$$

$$\times(-ie)^2\left[\overline{u}^r(\vec{k}_2)\not{\xi}^{(\lambda')}\frac{i}{\not{k}_1+\not{k}_2-m+i\epsilon}\not{\xi}^{(\lambda)}u^s(\vec{k}_1) + \overline{u}^r(\vec{k}_2)\not{\xi}^{(\lambda)}\frac{i}{\not{k}_2-\not{p}_1-m+i\epsilon}\not{\xi}^{(\lambda')}u^s(\vec{k}_1)\right] \tag{7.3.1}$$

We will continue to assume that the final spins are summed over and that the incoming electron is unpolarized, meaning that this spin is averaged. Squaring matrix elements we then get the following combinations:

$$\frac{1}{2}\sum_{r,s}|\overline{u}^r(\vec{p}_2)\not{\xi}^{(\lambda')}\frac{i}{\not{k}_1+\not{k}_2-m+i\epsilon}\not{\xi}^{(\lambda)}u^s(\vec{k}_1)|^2$$
$$= \frac{1}{2}\frac{\text{Tr}[(\not{p}_2+m)\not{\xi}^{(\lambda')}(\not{k}_1+\not{k}_2+m)\not{\xi}^{(\lambda)}(\not{k}_2+m)\not{\xi}^{(\lambda)*}(\not{k}_1+\not{k}_2+m)\not{\xi}^{(\lambda')*}]}{((k_1+k_2)^2-m^2)^2} \tag{7.3.2}$$

$$\frac{1}{2}\sum_{r,s}|\overline{u}^r(\vec{p}_2)\not{\xi}^{(\lambda)}\frac{i}{\not{k}_2-\not{p}_1-m+i\epsilon}\not{\xi}^{(\lambda')}u^s(\vec{k}_1)|^2$$
$$= \frac{1}{2}\frac{\text{Tr}[(\not{p}_2+m)\not{\xi}^{(\lambda)}(\not{k}_2-\not{p}_1+m)\not{\xi}^{(\lambda')}(\not{k}_2+m)\not{\xi}^{(\lambda')*}(\not{k}_1-\not{p}_1+m)\not{\xi}^{(\lambda)*}]}{((k_2-p_1)^2-m^2)^2} \tag{7.3.3}$$

$$\begin{aligned}
&\frac{1}{2}\sum_{r,s}\overline{u}^r(\vec{p}_2)\not{\xi}^{(\lambda')}\frac{i}{\not{k}_1+\not{k}_2-m+i\epsilon}\not{\xi}^{(\lambda)}u^s(\vec{k}_1)\overline{u}^s(\vec{k}_2)\not{\xi}^{(\lambda')*}\frac{-i}{\not{k}_2-\not{p}_1-m+i\epsilon}\not{\xi}^{(\lambda)*}u^r(\vec{k}_2)+\text{ c.c.}\\
&=\frac{1}{2}\frac{\mathrm{Tr}[(\not{p}_2+m)\not{\xi}^{(\lambda')}(\not{k}_1+\not{k}_2+m)\not{\xi}^{(\lambda)}(\not{k}_2+m)\not{\xi}^{(\lambda')*}(\not{k}_2-\not{p}_1+m)\not{\xi}^{(\lambda)*}]}{((k_1+k_2)^2-m^2)((k_2-p_1)^2-m^2)}+\text{ c.c.}\,,
\end{aligned} \tag{7.3.4}$$

where $\not{\xi}^{(\lambda)*}\equiv\xi^{(\lambda)*}_\mu\gamma^\mu$. We dropped the $i\,\epsilon$ terms in the denominator because the poles are never reached for on-shell particles.

To reduce these expressions further, we will also sum over the final state polarization of the photon and average over the incoming photon's polarization. We then use the following trick originally proposed by Feynman: In terms of one of the external photons, we can write the square of the $\mathcal{M}$-matrix element as

$$\xi^{(\lambda)}_\mu\mathcal{M}^\mu\xi^{(\lambda)*}_\nu\mathcal{M}^{\nu*}\,. \tag{7.3.5}$$

The two polarizations are normalized to be

$$\xi^{(\lambda)}\cdot\xi^{(\lambda')*}=-\delta^{\lambda\lambda'}\,, \tag{7.3.6}$$

hence they are space-like and span a two-dimensional space within the four dimensional space-time. Therefore the sum over polarizations should be given by

$$\sum_{\lambda=1,2}\xi^{(\lambda)}_\mu\xi^{(\lambda)*}_\nu=-\eta_{\mu\nu\perp}\,, \tag{7.3.7}$$

where ${\eta_\mu}^\nu_\perp$ is a rank two projector in the four-dimensional space-time[2]. Four dimensional Minkowski space can be spanned by two light-like vectors and two space-like vectors. Choosing $\xi^{(1)}_\mu$ and $\xi^{(2)}_\mu$ to be the space-like vectors and k_μ to be one of the light-like vectors, we choose another light-like vector $\bar{k}_\mu$, such that $\xi^{(\lambda)}\cdot\bar{k}=0$ [3] and $k\cdot\bar{k}\neq0$ [4]. It then follows that the projector must satisfy $k^\nu\eta_{\mu\nu\perp}=\bar{k}^\nu\eta_{\mu\nu\perp}=0$, $\xi^{(\lambda)\mu}\eta_{\mu\nu\perp}=\xi^{(\lambda)}_\nu$. It is straightforward to verify that

$$\eta_{\mu\nu\perp}=\eta_{\mu\nu}-\frac{k_\mu\bar{k}_\nu+k_\nu\bar{k}_\mu}{k\cdot\bar{k}} \tag{7.3.8}$$

is a rank two projector that satisfies these conditions. The expression in (7.3.8) is a rather awkward combination to have to carry around. But now we remember that $\xi^{(\lambda)}_\mu$ is contracted with $\mathcal{M}^\mu$ and by the Ward identity, $k_\mu\mathcal{M}^\mu=0$. Hence the second term in (7.3.8) does not contribute. Thus, we can replace $\eta_{\mu\nu\perp}$ with $\eta_{\mu\nu}$ in (7.3.7) and still get the correct answer.

Thus, if we sum over polarizations, we can make the replacement in (7.3.2), (7.3.3) and (7.3.4),

$$\sum_{\lambda=1,2}\ldots\not{\xi}^{(\lambda)}\ldots\not{\xi}^{(\lambda)*}\cdots=-\ldots\gamma^\mu\ldots\gamma_\mu\cdots\,. \tag{7.3.9}$$

[2]Recall that a projector squares to itself, hence its eigenvalues can only be 1 or 0. The rank is the number of 1's.

[3]For example, if $\vec{\xi}^{(\lambda)}\cdot\vec{k}=0$, we can choose $\bar{k}^\mu=(|\vec{k}|,-\vec{k})$.

[4]Somewhat counterintuitively, if $k\cdot\bar{k}=0$ then they are not independent 4-vectors.

We then use the identities in (6.2.11) of the sixth chapter of the notes. This results in

$$\begin{aligned}
&\frac{1}{2}\sum_{\lambda,\lambda'=1,2}\frac{1}{2}\sum_{r,s}|\overline{u}^r(\vec{p}_2)\not{\epsilon}^{(\lambda')}\frac{i}{\not{k}_1+\not{k}_2-m+i\epsilon}\not{\epsilon}^{(\lambda)}u^s(\vec{k}_1)|^2\\
&=\frac{1}{16}\frac{\mathrm{Tr}[(\not{p}_2+m)\gamma^\mu(\not{k}_1+\not{k}_2+m)\gamma^\nu(\not{k}_2+m)\gamma_\nu(\not{k}_1+\not{k}_2+m)\gamma_\mu]}{(k_1\cdot k_2)^2}\\
&=\frac{1}{4}\frac{\mathrm{Tr}[(-\not{p}_2+2m)(\not{k}_1+\not{k}_2+m)(-\not{k}_2+2m)(\not{k}_1+\not{k}_2+m)]}{(k_1\cdot k_2)^2}\\
&=\frac{4m^4+m^2(4(k_1+k_2)\cdot(k_1-p_2)+k_2\cdot p_2)+2k_2\cdot(k_1+k_2)p_2\cdot(k_1+k_2)-k_2\cdot p_2(k_1+k_2)^2}{(k_1\cdot k_2)^2}\\
&=\frac{2(k_1\cdot k_2+m^2)^2-2k_2\cdot p_2\,k_1\cdot k_2}{(k_1\cdot k_2)^2}=\frac{2m^4+k_1\cdot k_2(s+t+m^2)}{(k_1\cdot k_2)^2}\\
&=\frac{1}{2}\frac{4m^2+2m^2(s-m^2)-(s-m^2)(u-m^2)}{(k_1\cdot k_2)^2}\,,
\end{aligned}\tag{7.3.10}$$

where along the way in this calculation we used that $k_1+k_2=p_1+p_2$ and $k_1{\cdot}k_2=p_1{\cdot}p_2$, as well as the definitions of the Mandelstam variables

$$\begin{aligned}
s &= (k_1+k_2)^2=(p_1+p_2)^2\\
t &= (k_1-p_1)^2=(k_2-p_2)^2\\
u &= (k_1-p_2)^2=(p_1-k_2)^2
\end{aligned}\tag{7.3.11}$$

and their relation

$$s+t+u=2m^2\,. \tag{7.3.12}$$

By interchanging the s and the u channels it then follows that

$$\frac{1}{2}\sum_{\lambda,\lambda'=1,2}\frac{1}{2}\sum_{r,s}|\overline{u}^r(\vec{p}_2)\not{\epsilon}^{(\lambda)}\frac{i}{\not{k}_2-\not{p}_1-m+i\epsilon}\not{\epsilon}^{(\lambda')}u^s(\vec{k}_1)|^2=\frac{1}{2}\frac{4m^2+2m^2(u-m^2)-(s-m^2)(u-m^2)}{(p_1\cdot k_2)^2}\,. \tag{7.3.13}$$

Finally, after summing and averaging over the polarizations, (7.3.4) reduces to

$$\begin{aligned}
&\frac{1}{2}\sum_{\lambda,\lambda'=1,2}\frac{1}{2}\sum_{r,s}\overline{u}^r(\vec{p}_2)\not{\epsilon}^{(\lambda')}\frac{i}{\not{k}_1+\not{k}_2-m+i\epsilon}\not{\epsilon}^{(\lambda)}u^s(\vec{k}_1)\overline{u}^s(\vec{k}_2)\not{\epsilon}^{(\lambda')*}\frac{-i}{\not{k}_2-\not{p}_1-m+i\epsilon}\not{\epsilon}^{(\lambda)*}u^r(\vec{k}_2)+\ \text{c.c.}\\
&=-\frac{1}{8}\frac{\mathrm{Tr}[(\not{p}_2+m)\gamma^\mu(\not{k}_1+\not{k}_2+m)\gamma^\nu(\not{k}_2+m)\gamma_\mu(\not{k}_2-\not{p}_1+m)\gamma_\nu]}{k_1\cdot k_2\,p_1\cdot k_2}\\
&=-\frac{1}{8}\frac{\mathrm{Tr}[(\not{p}_2+m)(-2m^2\gamma^\nu+4m(k_1^\nu+2k_2^\nu)-2\not{k}_2\gamma^\nu(\not{k}_1+\not{k}_2))(\not{k}_2-\not{p}_1+m)\gamma_\nu]}{k_1\cdot k_2\,p_1\cdot k_2}\\
&=-\frac{1}{8\,k_1\cdot k_2\,p_1\cdot k_2}\mathrm{Tr}\Big[4m^2(\not{p}_2-2m)(\not{k}_2-\not{p}_1+m)+4m(\not{p}_2+m)(\not{k}_2-\not{p}_1+m)(\not{k}_1+2\not{k}_2)\\
&\qquad\qquad -4(2p_2\cdot k_2(\not{k}_1+\not{k}_2)(\not{k}_2-\not{p}_1)-m^2\not{k}_2(\not{k}_1+\not{k}_2)\Big]
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{4\,k_1\cdot k_2\,p_1\cdot k_2}\mathrm{Tr}\Big[4m^2(\not p_2-2m)(\not k_2-\not p_1+m)+4m(\not p_2+m)(\not k_2-\not p_1+m)(\not k_1+2\not k_2)\\
&= -\frac{2}{k_1\cdot k_2\,p_1\cdot k_2}\Big(-2m^4+m^2p_2\cdot(k_2-p_1)+2m^2(2p_2\cdot k_2-k_1\cdot p_1)-m^2(2p_2\cdot k_2-k_1\cdot k_2-m^2)\Big)\\
&= -\frac{m^2(4m^2-t)}{k_1\cdot k_2\,p_1\cdot k_2} = -\frac{m^2(4m^2+(s-m^2)+(u-m^2))}{k_1\cdot k_2\,p_1\cdot k_2}\,. \qquad (7.3.14)
\end{aligned}
$$

Combining the results in (7.3.10), (7.3.13) and (7.3.14), while using $s-m^2=2k_1\cdot k_2$, $u-m^2=-2p_1\cdot k_2$, we reach the final expression for the square of the $\mathcal{M}$-matrix

$$
\frac{1}{2}\sum_{\lambda,\lambda'=1,2}\frac{1}{2}\sum_{r,s}|\mathcal{M}|^2 = 2\,e^4\left(m^4\left(\frac{1}{k_1\cdot k_2}-\frac{1}{p_1\cdot k_2}\right)^2+\frac{2m^2+p_1\cdot k_2}{k_1\cdot k_2}-\frac{2m^2-k_1\cdot k_2}{p_1\cdot k_2}\right). \qquad (7.3.15)
$$

Recall from the fifth chapter of the notes that cross-section is

$$
\sigma_{\gamma e^-\to\gamma e^-} = \frac{1}{(2k_1^0)(2k_2^0)|v_1-v_2|}\frac{1}{2}\sum_{\lambda,\lambda'=1,2}\frac{1}{2}\sum_{r,s}\int\frac{d^3p_1}{(2\pi)^3 2p_1^0}\frac{d^3p_1}{(2\pi)^3 2p_2^0}(2\pi)^4\delta^4(P_{\rm out}-P_{\rm in})\,|\mathcal{M}|^2\,. \qquad (7.3.16)
$$

We first assume that the incoming electron is at rest (the lab frame), which is a common assumption in a Compton scattering experiment. In this case $k_2^0=m$ and $|v_1-v_2|=1$. We will use $E=k_1^0$ and $E'=p_1^0$, while in the lab frame we also have $k_1\cdot k_2=Em$, $k_2\cdot p_1=E'm$. If we let θ be the angle of the outgoing photon, then we can find E' in terms of E and θ from the equation

$$
\begin{aligned}
(k_1+k_2-p_1)^2 &= p_2^2\\
m^2+2k_2\cdot(k_1-p_1)-2k_1\cdot p_1 &= m^2\\
2m(E-E')-2EE'(1-\cos\theta) &= 0 \quad\Rightarrow\quad E'=\frac{mE}{m+E(1-\cos\theta)}\,. \qquad (7.3.17)
\end{aligned}
$$

We can then write the cross-section as

$$
\begin{aligned}
\sigma_{\gamma e^-\to\gamma e^-} &= \frac{e^4}{32\pi^2 mE}\int_0^\infty\frac{E'dE'}{E_e}\int d\Omega\,\delta(E'+E_e-E-m)\\
&\qquad\times\left(m^2\left(\frac{1}{E}-\frac{1}{E'}\right)^2+\frac{2m+E'}{E}-\frac{2m-E}{E'}\right), \qquad (7.3.18)
\end{aligned}
$$

where E_e is the energy of the outgoing electron. Given θ and E', we have that

$$
E_e=\Big(m^2+E^2+E'^2-2EE'\cos\theta\Big)^{1/2}\,, \qquad (7.3.19)
$$

from which it follows that

$$
\delta(E'+E_e-E-m)=\frac{E_eE'}{mE}\,\delta\left(E'-\frac{mE}{m+E(1-\cos\theta)}\right). \qquad (7.3.20)
$$

Therefore, the differential cross-section is

$$\frac{d\sigma_{\gamma e^-\to\gamma e^-}}{d\Omega} = \frac{e^4}{32\pi^2}\Bigg(\left(\frac{1-\cos\theta}{m+E(1-\cos\theta)}\right)^2 - \frac{2(1-\cos\theta)}{(m+E(1-\cos\theta))^2} + \frac{1}{m(m+E(1-\cos\theta))}\left(\frac{m^2}{(m+E(1-\cos\theta))^2}+1\right)\Bigg)\,, \tag{7.3.21}$$

So far this expression is exact at tree-level. If we now take the low energy limit where $E << m$, then the differential cross-section approximates to

$$\frac{d\sigma_{\gamma e^-\to\gamma e^-}}{d\Omega} \approx \frac{\alpha^2}{2m^2}(1+\cos^2\theta)\,, \tag{7.3.22}$$

where $\alpha = \frac{e^2}{4\pi}$ is the fine-structure constant, which at low energies is approximately $\alpha \approx 1/137$. The equation in (7.3.22) is called the Klein-Nishina formula.

At high energies where $E >> m$, the differential cross-section is approximately

$$\frac{d\sigma_{\gamma e^-\to\gamma e^-}}{d\Omega} \approx \frac{\alpha^2}{2mE}\frac{1}{1-\cos\theta}\,, \tag{7.3.23}$$

which peaks in the forward direction. However, this expression is for the lab frame while we are more likely to consider a high energy scattering in the COM frame. At high energies in the COM frame we have that $2k_1{\cdot}k_2 \approx s$ and $2p_1{\cdot}k_2 = m^2 - u \approx s\cos^2\frac{\theta}{2} + m^2$. Inspecting (7.3.15), we see that the dominant term in the $\mathcal{M}$-matrix is

$$\begin{aligned}\frac{1}{2}\sum_{\lambda,\lambda'=1,2}\frac{1}{2}\sum_{r,s}|\mathcal{M}|^2 &\approx 2\,e^4\left(\frac{p_1\cdot k_2}{k_1\cdot k_2}+\frac{k_1\cdot k_2}{p_1\cdot k_2}\right)\\ &= 2\,e^4\,\frac{1+\cos^4\frac{\theta}{2}}{\cos^2\frac{\theta}{2}+m^2/s} = e^4\,\frac{4+(1+\cos\theta)^2}{1+\cos\theta+2m^2/s}\,.\end{aligned} \tag{7.3.24}$$

The cross-section is then approximately given by

$$\sigma_{\gamma e^-\to\gamma e^-} \approx \frac{\alpha^2}{2s}\int d\Omega\,\frac{2+\frac{1}{2}(1+\cos\theta)^2}{1+\cos\theta+2m^2/s}\,, \tag{7.3.25}$$

and with differential cross-section

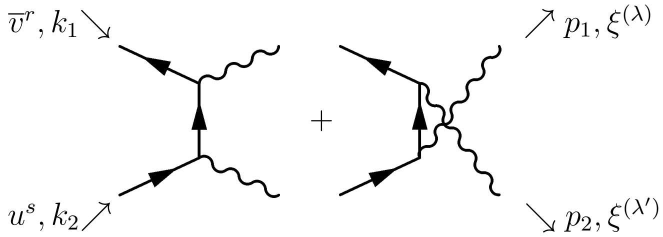
$$\frac{d\sigma_{\gamma e^-\to\gamma e^-}}{d\Omega} \approx \frac{\alpha^2}{2s}\,\frac{2+\frac{1}{2}(1+\cos\theta)^2}{1+\cos\theta+2m^2/s}\,. \tag{7.3.26}$$

We have kept the $m^2/s << 1$ term to prevent a divergence in $\frac{d\sigma_{\gamma e^-\to\gamma e^-}}{d\Omega}$ as $\theta \to \pi$. In fact, the total cross section is completely dominated by the back-scattering at $\theta \approx \pi$, where we find that

$$\sigma_{\gamma e^-\to\gamma e^-} \approx \frac{2\pi\alpha^2}{s}\int_{-1}^{1} d(\cos\theta)\frac{1}{1+\cos\theta+2m^2/s} \approx \frac{2\pi\alpha^2}{s}\log\frac{s}{m^2}\,. \tag{7.3.27}$$

7.3.2 $e^+e^- \to \gamma\gamma$

Our second example with external photons is electron-positron annihilation into two photons. As in the Compton scattering, there are two Feyman diagrams that contribute at tree-level. These are



$$= (2\pi)^4\delta^4(k_1+k_2-p_1-p_2)$$

$$\times(-ie)^2\left[\overline{v}^r(\vec{k}_1)\not{\xi}^{(\lambda)}\frac{i}{\not{k}_2-\not{p}_2-m+i\epsilon}\not{\xi}^{(\lambda')}u^s(\vec{k}_2)+\overline{v}^r(\vec{k}_1)\not{\xi}^{(\lambda)}\frac{i}{\not{k}_2-\not{p}_1-m+i\epsilon}\not{\xi}^{(\lambda')}u^s(\vec{k}_2)\right]. \tag{7.3.28}$$

Using crossing symmetry, we can borrow the result in (7.3.15) to find the square of the $\mathcal{M}$-matrix. Under the crossing we make the replacements $k_1 \to -p_2$, $p_2 \to -k_1$ in (7.3.15), accompanied by an overall change in sign. This last change of sign arises from replacing $\sum_r u^r(\vec{k})\overline{u}^r(\vec{k}) = \not{k}+m$ with $\sum_r v^r(\vec{k})\overline{v}^r(\vec{k}) = -(-\not{k}+m)$.

Assuming unpolarized incoming electrons and positrons and summing over the final photon polarizations, we find

$$\frac{1}{4}\sum_{r,s}\sum_{\lambda,\lambda'=1,2}|\mathcal{M}|^2 = 2\,e^4\left(-m^4\left(\frac{1}{k_2\cdot p_2}+\frac{1}{k_2\cdot p_1}\right)^2+\frac{2m^2+k_2\cdot p_1}{k_2\cdot p_2}+\frac{2m^2+k_2\cdot p_2}{k_2\cdot p_1}\right). \tag{7.3.29}$$

If we consider the ultrarelativistic limit in the COM frame where $s >> 4m^2$, then the differential cross-section is approximately

$$\frac{d\sigma_{e^+e^-\to\gamma\gamma}}{d\Omega} \approx \frac{\alpha^2}{2s}\left(\frac{k_2\cdot p_2}{k_2\cdot p_1}+\frac{k_2\cdot p_1}{k_2\cdot p_2}\right) \approx \frac{\alpha^2}{2s}\left(\frac{1+\cos\theta+2m^2/s}{1-\cos\theta+2m^2/s}+\frac{1-\cos\theta+2m^2/s}{1+\cos\theta-2m^2/s}\right), \tag{7.3.30}$$

where we have kept the $m^2/s << 1$ terms to avoid the singularities at $\theta = 0, \pi$. Away from the singularities we can drop these terms, after which the differential cross-section simplifies to

$$\frac{d\sigma_{e^+e^-\to\gamma\gamma}}{d\Omega} \approx \frac{\alpha^2}{s}\frac{1+\cos^2\theta}{\sin^2\theta}. \tag{7.3.31}$$

The outgoing photons are identical, hence we only integrate over the forward solid angle for the total cross-section. This then gives

$$\begin{aligned}\sigma_{e^+e^-\to\gamma\gamma} &\approx \frac{2\pi\alpha^2}{2s}\int_0^1 d(\cos\theta)\left(\frac{1+\cos\theta+2m^2/s}{1-\cos\theta+2m^2/s}+\frac{1-\cos\theta+2m^2/s}{1+\cos\theta-2m^2/s}\right)\\ &\approx \frac{2\pi\alpha^2}{s}\log\frac{s}{m^2},\end{aligned} \tag{7.3.32}$$

which is the same approximation found for high energy Compton scattering.

Chapter 8

One loop QED

In this final chapter of the notes we discuss one-loop QED. We first cover the generalization of the Ward identity to the Ward-Takahashi identity which applies to correlators. We then discuss three different one-loop calculations: vacuum polarization, fermion self-energy and the vertex function. Finally we discuss applications. These include the running of α, the Landé g-factor and Bremsstrahlung.

8.1 The Ward-Takahashi identity

In the previous chapter of the notes we argued that the $\mathcal{M}$-matrix for a scattering involving an external photon with momentum k_μ and polarization vector $\xi_\mu^{(\lambda)}$ can be written as $\mathcal{M} = \xi_\mu^{(\lambda)}\mathcal{M}^\mu$, and it satisfies $k_\mu\mathcal{M}^\mu = 0$. This result is the Ward identity and is actually a special case of a more general identity called the Ward-Takahashi (W-T) identity which applies to any correlator and not just those that appear in the $\mathcal{M}$-matrix where the momenta are on-shell. We will not derive the general form, but instead consider two special cases.

The W-T identity is derived in the same way that the Schwinger-Dyson equation is derived. Let us write a time ordered correlator in QED as

$$\langle T[\chi(x_1)\dots\chi(x_n)]\rangle = \mathcal{Z}^{-1}\int \mathcal{D}A_\mu\mathcal{D}\overline{\psi}\mathcal{D}\psi\chi(x_1)\dots\chi(x_n)\, e^{iS(A_\mu,\overline{\psi},\psi)}\,, \tag{8.1.1}$$

where $\chi(x_j)$ is one of the fields (A_μ, $\overline{\psi}$ or ψ, perhaps with derivatives). $S(A_\mu,\overline{\psi},\psi)$ is the QED action

$$S = \int d^4x\left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\partial\!\!\!/ - m)\psi - \frac{1}{2\lambda}(\partial_\mu A^\mu)^2\right]\,, \tag{8.1.2}$$

where we have included a gauge-fixing term. We then perform an infinitesimal gauge transformation of the field variables in the path integral,

$$\begin{aligned}
A_\mu(x) &\to A_\mu(x) - \frac{1}{e}\partial_\mu\delta\theta(x)\\
\psi(x) &\to \psi(x) + i\,\psi(x)\delta\theta(x)\\
\overline{\psi}(x) &\to \overline{\psi}(x) - i\,\overline{\psi}(x)\delta\theta(x)\,.
\end{aligned} \tag{8.1.3}$$

But since all fields are being integrated over, this shift leaves the correlator unchanged. Since the only term in the action which is not gauge invariant is the gauge fixing term, we then find

$$\langle T[\frac{i}{e\lambda}\int d^4x(\partial^2\partial_\mu A^\mu)\delta\theta(x)\chi(x_1)\ldots\chi(x_n)]\rangle+\sum_{j=1}^{n}\langle T[\chi(x_1)\ldots\delta\chi(x_j)\ldots\chi(x_n)]\rangle=0 \tag{8.1.4}$$

where $\delta\chi(x_j)$ is one of the gauge transformations in (8.1.3).

The first case we consider is for a single gauge field $A_\nu(y)$. Then (8.1.4) becomes

$$\begin{aligned}\langle T[\frac{i}{e\lambda}\int d^4x(\partial^2\partial_\mu A^\mu)\delta\theta(x)A_\nu(y)]\rangle-\frac{1}{e}\langle T[\partial_\nu\delta\theta(y)]\rangle &= 0\\ \frac{i}{e\lambda}\int d^4x\frac{\partial}{\partial x^\lambda}\frac{\partial}{\partial x_\lambda}\frac{\partial}{\partial x_\mu}G_{\mu\nu}(y-x)\delta\theta(x)-\frac{1}{e}\partial_\nu\delta\theta(y) &= 0\end{aligned} \tag{8.1.5}$$

If we Fourier transform over the y-coordinate we then get

$$\begin{aligned}&\frac{i}{e\lambda}\int d^4ye^{ik\cdot y}\int d^4x\left[\frac{\partial}{\partial x^\lambda}\frac{\partial}{\partial x_\lambda}\frac{\partial}{\partial x_\mu}G_{\mu\nu}(y-x)\right]\delta\theta(x)-\int d^4ye^{ik\cdot y}\frac{1}{e}\partial_\nu\delta\theta(y)=0\\ &\frac{k^\mu k^2}{e\lambda}\int d^4y\,e^{ik\cdot y}G_{\mu\nu}(y)\int d^4x\,e^{ik\cdot x}\delta\theta(x)+\frac{ik_\nu}{e}\int d^4ye^{ik\cdot y}\delta\theta(y)=0\,,\end{aligned} \tag{8.1.6}$$

where we used that $\frac{\partial}{\partial x^\lambda}G_{\mu\nu}(x-y)=-\frac{\partial}{\partial y^\lambda}G_{\mu\nu}(x-y)$. Hence we find

$$k^\mu G_{\mu\nu}(k)=\frac{-i\lambda\,k_\nu}{k^2}\,. \tag{8.1.7}$$

Note that this equation is exact to all orders in perturbation theory. At tree-level we have that

$$G_{\mu\nu}(k)=\frac{-i(\eta_{\mu\nu}-(1-\lambda)k_\mu k_\nu/k^2)}{k^2}\,, \tag{8.1.8}$$

and so $k^\mu G_{\mu\nu}(k)$ is the same expression as in (8.1.7). This indicates that there are no loop corrections to the gauge fixing term. In the next section we will present another argument why this is true.

The second case we consider starts with the fermion propagator $\langle T[\psi(x_1)\overline{\psi}(x_2)]\rangle$. Shifting the variables by a gauge transformation then gives

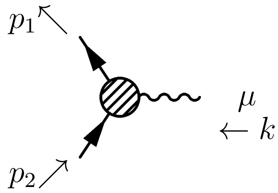
$$\frac{i}{e\lambda}\int d^4x\langle T[(\partial^2\partial_\mu A^\mu(x))\psi(x_1)\overline{\psi}(x_2)]\rangle\delta\theta(x)+i\langle T[\psi(x_1)\overline{\psi}(x_2)]\rangle(\delta\theta(x_1)-\delta\theta(x_2))=0\,. \tag{8.1.9}$$

After Fourier transforming and integrating by parts this becomes

$$\begin{aligned}&\frac{i}{e\lambda}\int\frac{d^4k}{(2\pi)^4}d^4x_1d^4x_2d^4xe^{ip_1\cdot x_1-ip_2\cdot x_2-ik\cdot x}(-ik^\mu k^2)\langle T[A^\mu(x)\psi(x_1)\overline{\psi}(x_2)]\rangle\delta\theta(k)\\ &\qquad+i\int\frac{d^4k}{(2\pi)^4}d^4x_1d^4x_2(e^{ip_1\cdot x_1-ip_2\cdot x_2-ik\cdot x_1}-e^{ip_1\cdot x_1-ip_2\cdot x_2-ik\cdot x_2})\langle T[\psi(x_1)\overline{\psi}(x_2)]\rangle\delta\theta(k)\\ &=0\end{aligned} \tag{8.1.10}$$

We then define the vertex function $V_\mu(p_1, p_2)$ such that

$$(2\pi)^4\delta^4(p_1 - p_2 - k)\, e\, V_\mu(p_1, p_2) = \tag{8.1.11}$$



where the diagram represents the complete truncated three-point correlator. Since the full 3-point correlator ($i.e$ untruncated) is given by

$$\int d^4x_1 d^4x_2 d^4x\, e^{ip_1\cdot x_1 - ip_2\cdot x_2 - ik\cdot x}\langle T[A^\mu(x)\psi(x_1)\overline{\psi}(x_2)]\rangle\,, \tag{8.1.12}$$

we have that

$$-\frac{i\,k^2k^\nu}{\lambda}G_{\mu\nu}(k^2)S(\not{p}_1)V_\mu(p_1,p_2)S(\not{p}_2)\bigg|_{k=p_1-p_2} = -S(\not{p}_2) + S(\not{p}_1)\,, \tag{8.1.13}$$

where $G_{\mu\nu}(k^2)$ and $S(\not{p})$ are the photon and fermion propagators to all orders in perturbation theory. Now we use (8.1.7) and divide by the fermion propagators to finally reach

$$(p_1^\mu - p_2^\mu)V_\mu(p_1, p_2) = S^{-1}(\not{p}_1) - S^{-1}(\not{p}_2) \tag{8.1.14}$$

Even though it is a special case, this equation is often referred to as the Ward-Takahashi identity and it is true to all orders in perturbation theory.

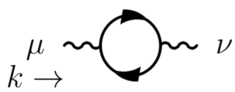
As a quick check, let us verify (8.1.14) at the leading level in perturbation theory. At tree-level we have that $V_\mu = -i\gamma_\mu$ and $S^{-1}(\not{p}) = -i(\not{p} - m)$. Hence (8.1.14) is satisfied at this level of perturbation theory. To higher orders in perturbation theory the W-T identity will put stringent constraints on the counterterms in the theory.

8.2 Three one-loop calculations in QED

8.2.1 Vacuum polarization

The photon self-energy is also called vacuum polarization. We will explain the name after we consider the contributions to the self-energy. The self-energy is described in terms of a tensor $\Pi_{\mu\nu}(k)$, where the leading contribution is the one-loop diagram

$$i\,\Pi_{\mu\nu}(k) = \tag{8.2.1}$$



$+ \ldots +$ counterterms.

To any order of perturbation theory we must have that $k^\mu\Pi_{\mu\nu}(k) = 0$ since the longitudinal component of the propagator has no perturbative corrections. Alternatively we can see this by noticing that the self-energy always has the incoming or outgoing photon coupled to a conserved current. After integrating by parts, the longitudinal component can always be rewritten as a $\partial_\mu j^\mu$ term.

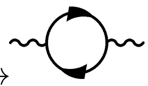
Since the longitudinal component is zero, we can express $\Pi_{\mu\nu}(k)$ as

$$\Pi_{\mu\nu}(k) = (\eta_{\mu\nu} - k_\mu k_\nu/k^2)\Pi(k)\,. \tag{8.2.2}$$

Hence the propagator is given by

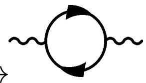
$$G_{\mu\nu}(k^2) = \frac{-i(\eta_{\mu\nu} - k_\mu k_\nu/k^2)}{k^2 - \Pi(k^2) + i\epsilon} + \frac{-i\lambda\, k_\mu k_\nu/k^2}{k^2 + i\epsilon}\,. \tag{8.2.3}$$

Let us now compute the diagram in (8.2.1). Using the Feynman rules we have that

μ  ν, $k \to$

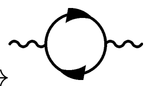
$$\begin{aligned} &= (-1)(-ie)^2 \int \frac{d^4\ell}{(2\pi)^4} \mathrm{Tr}\left[\gamma_\mu \frac{i}{\not\ell - m + i\epsilon}\gamma_\nu \frac{i}{\not\ell - \not k - m + i\epsilon}\right] \\ &= -4e^2 \int \frac{d^4\ell}{(2\pi)^4} \frac{2\ell_\mu\ell_\nu - \ell_\mu k_\nu - \ell_\nu k_\mu - \eta_{\mu\nu}(\ell^2 - \ell\cdot k - m^2)}{(\ell^2 - m^2 + i\epsilon)((\ell-k)^2 - m^2 + i\epsilon)}\,, \end{aligned} \tag{8.2.4}$$

where the (-1) factor is for the fermion loop. If we now Wick rotate and dimensionally regularize this becomes

μ  ν, $k \to$

$$= -4i\, e^2\mu^{4-D} \int \frac{d^D\ell}{(2\pi)^D} \frac{2\bar\ell_\mu\bar\ell_\nu - \bar\ell_\mu k_\nu - \bar\ell_\nu k_\mu + \eta_{\mu\nu}(\ell^2 - \ell\cdot k_E + m^2)}{(\ell^2 + m^2)((\ell - k_E)^2 + m^2)} \tag{8.2.5}$$

where $\bar\ell_\mu = (i\ell_0, -\vec\ell)$ and $k_E^2 = -k^2$. As usual we let $D = 4 - 2\,\varepsilon$. Performing the usual Feynman parameterization trick we then get

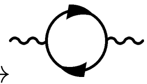
μ  ν, $k \to$

$$\begin{aligned} &= -4\,i\,e^2\mu^{4-D} \int_0^1 dx \int_0^\infty \rho d\rho \int \frac{d^D\ell}{(2\pi)^D} \\ &\qquad \times \left(2\bar\ell_\mu\bar\ell_\nu - \bar\ell_\mu k_\nu - \bar\ell_\nu k_\mu + \eta_{\mu\nu}(\ell^2 - \ell\cdot k_E + m^2)\right) \\ &\qquad \times \exp(-\rho((\ell - (1-x)k_E)^2 + x(1-x)k_E^2 + m^2)) \\ &= -4\,ie^2\mu^{4-D} \int_0^1 dx \int_0^\infty d\rho \int \frac{d^D\ell}{(2\pi)^D} \\ &\times \Bigg(2(\bar\ell + (1-x)k)_\mu(\bar\ell + (1-x)k)_\nu - (\bar\ell + (1-x)k)_\mu k_\nu - (\bar\ell + (1-x)k)_\nu k_\mu \\ &\qquad + \eta_{\mu\nu}\Big((\ell + (1-x)k_E)^2 - (\ell + (1-x)k_E)\cdot k_E + m^2\Big)\Bigg) \exp(-\rho(\ell^2 + x(1-x)k_E^2 + m^2)) \end{aligned} \tag{8.2.6}$$

where in the last step we shifted the integration variable ℓ_μ. We then use the identities

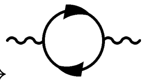
$$\int \frac{d^D\ell}{(2\pi)^D}\bar\ell_\mu f(\ell^2) = 0\,, \qquad \int \frac{d^D\ell}{(2\pi)^D}\bar\ell_\mu\bar\ell_\nu f(\ell^2) = -\frac{\eta_{\mu\nu}}{D}\int \frac{d^D\ell}{(2\pi)^D}\ell^2 f(\ell^2) \tag{8.2.7}$$

where $f(\ell^2)$ is any function of ℓ^2, to write the expression as

μ  ν ($k \rightarrow$)

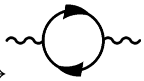
$$= -4\,i\,e^2\mu^{4-D}\int_0^1 dx\int_0^\infty \rho\, d\rho\int\frac{d^D\ell}{(2\pi)^D}$$
$$\times\Big(\eta_{\mu\nu}\big((1-2/D)\ell^2+x(1-x)k^2+m^2\big)-2x(1-x)k_\mu k_\nu\Big)\exp(-\rho(\ell^2-x(1-x)k^2+m^2)) \tag{8.2.8}$$

The gaussian integral over ℓ then gives

μ  ν ($k \rightarrow$)

$$= -4\,i\,e^2\frac{\mu^{4-D}}{(4\pi)^{D/2}}\int_0^1 dx\int_0^\infty \rho^{1-D/2}\, d\rho$$
$$\times\Big(\eta_{\mu\nu}\big((D/2-1)\rho^{-1}+x(1-x)k^2+m^2\big)-2x(1-x)k_\mu k_\nu\Big)\exp(-\rho(m^2-x(1-x)k^2)) \tag{8.2.9}$$

and then the integral over ρ leaves

μ  ν ($k \rightarrow$)

$$= -4\,i\,e^2\frac{\mu^{4-D}}{(4\pi)^{D/2}}\int_0^1 dx(m^2-x(1-x)k^2)^{D/2-2}$$
$$\times\Big(\eta_{\mu\nu}(m^2-x(1-x)k^2)(D/2-1)\Gamma(1-D/2)+\Big(\eta_{\mu\nu}x(1-x)k^2+m^2\Big)-2x(1-x)k_\mu k_\nu\Big)\Gamma(2-D/2)\Big)$$
$$= -8\,i\,e^2\frac{\mu^{4-D}}{(4\pi)^{D/2}}(\eta_{\mu\nu}k^2-k_\mu k_\nu)\Gamma(2-D/2)\int_0^1 dx\,x(1-x)(m^2-x(1-x)k^2))^{D/2-2}\,, \tag{8.2.10}$$

where we used that $(1-D/2)\Gamma(1-D/2)=\Gamma(2-D/2)$. As promised, this contribution to the self-energy has $k^\mu\Pi_{\mu\nu}(k)=0$, and so in terms of $\Pi_0(k^2)$, the self-energy without a counterterm added we find

$$\begin{aligned}\Pi_0(k^2) &= -8\,e^2\frac{\mu^{4-D}}{(4\pi)^{D/2}}k^2\Gamma(2-D/2)\int_0^1 dx\,x(1-x)(m^2-x(1-x)k^2)^{D/2-2}\,.\\ &= -\frac{e^2}{12\pi^2}\left(\frac{1}{\varepsilon}-\gamma_E+\log\frac{\mu^2}{m^2}+\log(4\pi)-R\left(\frac{-k^2}{m^2}\right)\right)\,,\end{aligned} \tag{8.2.11}$$

where

$$R(z)=6\int_0^1 dx\,x(1-x)\log(1+x(1-x)z)\,. \tag{8.2.12}$$

$\Pi_0(k^2)$ is divergent, but if we now add the counterterm,

$$\mu \;\; \text{} \;\; \nu = -i\delta_{Z_3}(k^2\eta_{\mu\nu}-k_\mu k_\nu) \tag{8.2.13}$$

 ($k \rightarrow$)

where

$$\delta_{Z_3}=-\frac{e^2}{12\pi^2}\left(\frac{1}{\varepsilon}-\gamma_E+\log\frac{\mu^2}{m^2}+\log(4\pi)\right) \tag{8.2.14}$$

then the renormalized self-energy $\Pi(k^2)$ is finite and given by

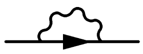
$$\Pi(k^2) = \frac{e^2}{12\pi^2} R\left(\frac{-k^2}{m^2}\right) k^2 \,. \tag{8.2.15}$$

Note that $R(z) = 0$ when $z = 0$, so the residue of the photon propagator at the pole $k^2 = 0$ is 1. The effect of the counterterm in (8.2.13) is to add the term $-\frac{1}{4}\delta_{Z_3} F_{\mu\nu}F^{\mu\nu}$ to the Lagrangian.

8.2.2 Fermion self-energy

The relevant Feynman diagram for the fermion (electron, muon *etc.*) self-energy is

$$-i\Sigma_0(\not k) =$$



$k \rightarrow$

$$\begin{aligned} &= (-ie)^2 \int \frac{d^4\ell}{(2\pi)^4} \frac{-i}{\ell^2 - \bar\mu^2 + i\epsilon} \gamma^\mu \frac{i}{\not k - \not\ell - m + i\epsilon} \gamma_\mu \\ &= -e^2 \int \frac{d^4\ell}{(2\pi)^4} \frac{\gamma^\mu (m + \not k - \not\ell)\gamma_\mu}{(\ell^2 - \bar\mu^2 + i\epsilon)((\ell - k)^2 - m^2 + i\epsilon)} \,. \end{aligned} \tag{8.2.16}$$

In anticipation of a future problem we have included a mass term $\bar\mu$ for the photon. Notice that we have written the self-energy as a function of $\not k$. This is possible, even though some terms do not have explicit $\not k$ terms, because one can write $\not k \not k = k^2$. The upcoming dimensional regularization requires us to modify some of γ-matrix relations. In particular, for a general value of D we have

$$\gamma^\mu\gamma_\mu = D \qquad\qquad \gamma^\mu\gamma^\nu\gamma_\mu = 2\gamma^\nu - \gamma^\nu\gamma^\mu\gamma_\mu = (2 - D)\gamma^\nu \,. \tag{8.2.17}$$

Wick rotating, dimensionally regulating, inserting the Feynman parameters, doing the Gaussian integration and expanding in ε gives

$$\begin{aligned} -i\Sigma_0(\not k) &= -i\,e^2 \frac{\mu^{4-D}}{(4\pi)^{D/2}} \int_0^\infty \rho\, d\rho \int_0^1 dx \int \frac{d^D\ell}{(2\pi)^D} (D\,m + (2 - D)(\not k - \not\ell)) \\ &\qquad\qquad \times \exp(-\rho((\ell - (1-x)k_E)^2 + x(1-x)k_E^2 + (1-x)m^2 + x\,\bar\mu^2) \\ &= -i\,e^2 \frac{\mu^{4-D}}{(4\pi)^{D/2}} \int_0^\infty \rho^{1-D/2}\, d\rho \int_0^1 dx (D\,m + (2-D)x\not k) \exp(-\rho((1-x)(m^2 - xk^2) + x\bar\mu^2)) \\ &= -i\,e^2 \frac{\mu^{4-D}}{(4\pi)^{D/2}} \Gamma(2 - D/2) \int_0^1 dx (D\,m + (2-D)x\not k)((1-x)(m^2 - xk^2) + x\bar\mu^2)^{D/2-2} \,, \\ &= -i\frac{e^2}{16\pi^2} \left[4\left(\frac{1}{\varepsilon} - \gamma_E + \log 4\pi + \log\frac{\mu^2}{m^2} + \frac{3}{2} - R_1\left(\frac{k^2}{m^2}, \frac{\bar\mu^2}{m^2}\right)\right) m \right. \\ &\qquad\qquad \left. - \left(\frac{1}{\varepsilon} - \gamma_E + \log 4\pi + \log\frac{\mu^2}{m^2} + 2 - R_2\left(\frac{k^2}{m^2}, \frac{\bar\mu^2}{m^2}\right)\right) \not k \right] \end{aligned} \tag{8.2.18}$$

where

$$R_1(z, w) = \int_0^1 dx\, \log\left(\frac{1 - x\,z}{1 - x} + \frac{x\,w}{(1-x)^2}\right) \,, \quad R_2(z) = \int_0^1 dx\, 2x\, \log\left(\frac{1 - x\,z}{1 - x} + \frac{x\,w}{(1-x)^2}\right) \,. \tag{8.2.19}$$

Both functions satisfy $R_1(-1,0) = R_2(-1,0) = 0$.

Not surprisingly $\Sigma_0(\not{k})$ is divergent, so we add the counterterms

$$k \rightarrow \qquad = i(\delta_{Z_2}(\not{k} - m) - \delta_m)\,. \tag{8.2.20}$$

Analogously to the real scalar case, in order to leave the pole at m with unit residue the wave-function and mass renormalization counterterms should be given by

$$\begin{aligned}
\delta_{Z_2} = \left.\frac{d\Sigma_0(\not{k})}{d\not{k}}\right|_{\not{k}=m} &= -\frac{e^2}{16\pi^2}\left(\frac{1}{\varepsilon} - \gamma_E + \log 4\pi + \log\frac{\mu^2}{m^2} + 2 - W\left(\frac{\bar\mu^2}{m^2}\right)\right) \\
\delta_m = -\Sigma_0(m) &= -\frac{3e^2}{16\pi^2}\left(\frac{1}{\varepsilon} - \gamma_E + \log 4\pi + \log\frac{\mu^2}{m^2} + \frac{4}{3}\right) m\,,
\end{aligned} \tag{8.2.21}$$

where

$$W\left(\frac{\bar\mu^2}{m^2}\right) = \int_0^1 dx \frac{4(2-x)x}{1 - x + \frac{x\bar\mu^2/m^2}{1-x}} = -2\log\frac{\bar\mu^2}{m^2} - 2 + \mathrm{O}\left(\frac{\bar\mu^2}{m^2}\right)\,. \tag{8.2.22}$$

W is divergent in the limit as $\bar\mu \to 0$. Unlike the divergence that arises as $\varepsilon \to 0$, which is an ultraviolet divergence due to large loop momentum contributions, the divergence in W is an infrared divergence which comes from the small loop momentum contributions. Giving the photon a small mass cuts off this divergence.

After adding the proposed counterterms we are left with a pole at $\not{k} = m$ with unit residue. The renormalized self-energy is then

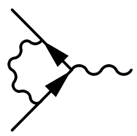
$$\Sigma(\not{k}) = \frac{e^2}{16\pi^2}\left(\left(R_2\left(\frac{k^2}{m^2}, \frac{\bar\mu^2}{m^2}\right) - W\left(\frac{\bar\mu^2}{m^2}\right)\right)\not{k} - \left(4R_1\left(\frac{k^2}{m^2}, \frac{\bar\mu^2}{m^2}\right) - W\left(\frac{\bar\mu^2}{m^2}\right)\right) m\right)\,. \tag{8.2.23}$$

There is something deeply unsettling about this expression. While it gives us a unit residue at the pole, away from the pole it is divergent as $\bar\mu \to 0$. We will delve into this further when we discuss radiation from soft photons.

8.2.3 One-loop vertex function

Our final one-loop computation is the one-loop correction to the vertex function. The diagram is

$k_1 \nwarrow$



$k_2 \nearrow$

$$\mu \atop \leftarrow q \quad = (2\pi)^4\delta^4(k_1 - k_2 + q)\, e\, V^\mu_{1,0}(k_1, k_2)$$

$$= (2\pi)^4\delta^4(k_1 - k_2 - q)(-i\,e)^3 \int \frac{d^4\ell}{(2\pi)^4}\frac{-i}{\ell^2 - \bar\mu^2 + i\epsilon}\,\gamma^\nu \frac{i}{\not{k}_1 - \not{\ell} - m + i\epsilon}\gamma^\mu \frac{i}{\not{k}_2 - \not{\ell} - m + i\epsilon}\gamma_\nu\,, \tag{8.2.24}$$

where $V_{1,0}(k_1,k_2)$ is the bare one-loop correction to the vertex function. Wick rotating and using dimensional regularization we get

$$
\begin{aligned}
i\,V_{1,0}(k_1,k_2) &= -e^2\mu^{4-D}\int_0^\infty \rho^{2-D/2}d\rho\int_0^1 dx_1dx_2dx_3\int\frac{d^D\ell}{(2\pi)^D}\delta(1-x_1-x_2-x_3)\\
&\times\Big(-2m^2\gamma^\mu+4m(k_1^\mu+k_2^\mu-2\bar\ell^\mu)-2(\not k_2-\bar{\not\ell})\gamma^\mu(\not k_1-\bar{\not\ell})+(4-D)\bar{\not\ell}\gamma^\mu\bar{\not\ell}\Big)\\
&\times\exp(-\rho((\ell-x_1k_{1E}-x_2k_{2E})^2+x_1x_3k_{1E}^2+x_2x_3k_{2E}^2+x_1x_2(k_{1E}-k_{2E})^2+(x_1+x_2)m^2+x_3\bar\mu^2)\,.
\end{aligned}
\tag{8.2.25}
$$

Here we have included a dimensionally modified γ-matrix identity,

$$
\gamma^\nu\gamma^\mu\gamma^\lambda\gamma^\sigma\gamma_\nu=-2\,\gamma^\sigma\gamma^\lambda\gamma^\mu+(4-D)\gamma^\mu\gamma^\lambda\gamma^\sigma\,.
\tag{8.2.26}
$$

which leads to a finite contribution because the correction multiplies a divergent piece. The modifications that make no contribution in the $\varepsilon\to 0$ limit have been dropped. Shifting variables, using

$$
\bar{\not\ell}\gamma^\mu\bar{\not\ell}=\bar\ell_\nu\gamma^\nu\gamma^\mu\bar\ell_\lambda\gamma^\lambda\to-\frac{1}{D}\gamma^\nu\gamma^\mu\gamma_\nu\ell^2=\frac{D-2}{D}\gamma^\mu\ell^2\,,
\tag{8.2.27}
$$

and performing the Gaussian integrations leads to

$$
\begin{aligned}
i\,V_{1,0}(k_1,k_2) &= -e^2\frac{\mu^{4-D}}{(4\pi)^{D/2}}\int_0^\infty \rho^{2-D/2}d\rho\int_0^1 dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)\\
&\Big[-2m^2\gamma^\mu+4m(k_1^\mu(1-2x_1)+k_2^\mu(1-2x_2))-2(\not k_2(1-x_2)-x_1\not k_1)\gamma^\mu(\not k_1(1-x_1)-x_2\not k_2)\\
&-\frac{(D-2)^2}{D}\rho^{-1}\gamma^\mu\Big]\exp(-\rho(x_1x_3k_{1E}^2+x_2x_3k_{2E}^2+x_1x_2(k_{1E}-k_{2E})^2+(x_1+x_2)m^2+x_3\bar\mu^2)\,.
\end{aligned}
\tag{8.2.28}
$$

Only the last term inside the square parentheses leads to an ultraviolet divergence. This particular term contributes

$$
\frac{e^2}{16\pi^2}\left(\frac{1}{\varepsilon}-\gamma_E+\log 4\pi+\log\frac{\mu^2}{m^2}-1-R_3\left(\frac{k_1^2}{m^2},\frac{k_2^2}{m^2},\frac{q^2}{m^2}\right)\right)\gamma^\mu
\tag{8.2.29}
$$

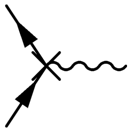
where

$$
\begin{aligned}
R_3(z_1,z_2,y) &= \int dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)\\
&\qquad\times\log\left[1-\frac{(z_1-1)x_1x_3+(z_2-1)x_2x_3+y\,x_1x_2}{(x_1+x_2)^2}\right]\,.
\end{aligned}
\tag{8.2.30}
$$

Note that $R_3(1,1,0) = 0$. We want to add a counterterm to cancel the divergence in (8.2.29). Because of Lorentz invariance we know that the bare one-loop vertex correction $V^\mu_{1,0}$ can be written as a sum of three terms

$$i\,V^\mu_{1,0} = \gamma^\mu\,X_1 + (k_1^\mu + k_2^\mu)X_2 + (k_1^\mu - k_2^\mu)X_3 \tag{8.2.31}$$

where the X_i are Lorentz invariant combinations involving k_1, k_2 and m. These combinations can include $\not{k}_1$ and $\not{k}_2$. The counterterm should have the form



$$= -ie\gamma^\mu(2\pi)^4\delta^4(k_1-k_2-q)\delta_{Z_1}\,, \tag{8.2.32}$$

so combining with the one-loop vertex we get the renormalized one-loop correction to the vertex

$$i\,V^\mu_{1R} = \gamma^\mu(X_1 + \delta_{Z_1}) + (k_1^\mu + k_2^\mu)X_2 + (k_1^\mu - k_2^\mu)X_3 \tag{8.2.33}$$

The counterterm needs to be consistent with the W-T identity. Before any counterterms are added, the W-T identity guarantees that[1]

$$q_\mu V^\mu_{1,0} = i\,(\Sigma_0(\not{k}_1) - \Sigma_0(\not{k}_2))\,. \tag{8.2.34}$$

The counterterms must preserve this such that

$$q_\mu V^\mu_{R} = i\,(\Sigma(\not{k}_1) - \Sigma(\not{k}_2))\,. \tag{8.2.35}$$

Comparing the form of the counterterms in (8.2.20) and (8.2.32), it follows that

$$-\,i\delta_{Z_1}(\not{k}_1 - \not{k}_2) = i(-\delta_{Z_2}(\not{k}_1 - m) + \delta_m + \delta_{Z_2}(\not{k}_1 - m) - \delta_m)\,, \tag{8.2.36}$$

and thus $\delta_{Z_1} = \delta_{Z_2}$, or equivalently, $Z_1 = Z_2$, where $Z_i = 1 + \delta_{Z_i}$. As a check, from (8.2.29) we can see that

$$\delta_{Z_1} = -\frac{e^2}{16\pi^2}\left(\frac{1}{\varepsilon} + \text{finite}\right)\,. \tag{8.2.37}$$

Hence the divergent term in (8.2.37) matches the divergent term for δ_{Z_2}.

8.3 Applications

8.3.1 Form factors

This section is not really an application, but it will be used in the subsequent sections.

[1] Assuming our regularization procedure is consistent with gauge invariance. We won't prove it, but dimensional regularization is consistent.

Suppose the vertex function $V_\mu(k_1,k_2)$ is sandwiched between two on-shell Dirac wave-functions,

$$W^\mu = \overline{u}^r(\vec{k}_1)i\,V^\mu(k_1,k_2)u^s(\vec{k}_2) \tag{8.3.1}$$

The W-T identity gives $q_\mu W^\mu = 0$, where $q^\mu = k_1^\mu - k_2^\mu$. Since

$$\overline{u}^r(\vec{k}_1)(\not{k}_1 - \not{k}_2)u^s(\vec{k}_2) = 0 \quad \text{and} \quad (k_1^\mu + k_2^\mu)q_\mu = 0\,, \tag{8.3.2}$$

it follows that X_3 in (8.2.33) is zero when k_1^μ and k_2^μ are on-shell. Hence we have that

$$W^\mu = \overline{u}^r(\vec{k}_1)(\gamma^\mu(1+X_1+\delta_{Z_1}) + (k_1^\mu+k_2^\mu)X_2)u^s(\vec{k}_2)\,, \tag{8.3.3}$$

where we have included the tree-level and one-loop results, as well as the counterterm. We then use the Gordon identity (ps #5),

$$\overline{u}^r(\vec{k}_1)\gamma^\mu u^s(\vec{k}_2) = \overline{u}^r(\vec{k}_1)\left(\frac{k_1^\mu + k_2^\mu}{2m} + \frac{i\sigma^{\mu\nu}q_\nu}{2m}\right)u^s(\vec{k}_2)\,, \tag{8.3.4}$$

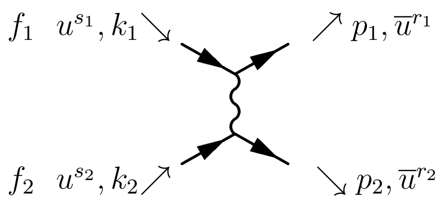
where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$, to write

$$W^\mu = \overline{u}^r(\vec{k}_1)\left(\gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2(q^2)\right)u^s(\vec{k}_2)\,. \tag{8.3.5}$$

The functions $F_1(q^2)$ and $F_2(q^2)$ are called form factors and by Lorentz invariance and momentum conservation can only depend on q^2 and the fermion mass (whose dependence is implied.) We will later show that $F_1(0) = 1$.

8.3.2 The running of the coupling

Suppose we have an electromagnetic scattering between two heavy fermions, f_1 and f_2, each having the same charge as the electron. The masses of the fermions, M_1 and M_2, are much heavier than the electron mass m. The tree-level diagram looks like



$$= (2\pi)^4\delta^4(k_1+k_2-p_1-p_2)$$

$$\times(-ie)^2\frac{-i\,\eta_{\mu\nu}\,\overline{u}^{r_1}(\vec{p}_1)\gamma^\mu u^{s_1}(\vec{k}_1)\overline{u}^{r_2}(\vec{p}_2)\gamma^\nu u^{s_2}(\vec{k}_2)}{(p_1-k_1)^2+i\epsilon} \tag{8.3.6}$$

Including the one-loop effects, the last line of (8.3.6) becomes

$$(e)^2\frac{-i\,\eta_{\mu\nu}\,\overline{u}^{r_1}(\vec{p}_1)V^\mu_{f_1}(p_1,k_1)u^{s_1}(\vec{k}_1)\overline{u}^{r_2}(\vec{p}_2)V^\nu_{f_2}(p_2,k_2)u^{s_2}(\vec{k}_2)}{(p_1-k_1)^2-\Pi(q^2)+i\epsilon}\,, \tag{8.3.7}$$

where $V^\mu_{f_1}$ and $V^\mu_{f_2}$ are the vertex functions for the two fermions and $q^2=(p_1-k_1)^2=t$. There is no contribution from the fermion self-energies since the counterterms were chosen so that the self-energies and their derivatives are zero on-shell. The vacuum polarization is assumed to have contributions from all fermions that are present in the theory, including the electron.

The vertex functions can be replaced with the form factors,

$$\begin{aligned} i\,V^\mu_{f_1}(p_1,k_1) &\rightarrow \gamma^\mu F_1^{(f_1)}(q^2)+\frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2^{(f_1)}(q^2)\\ i\,V^\mu_{f_2}(p_2,k_2) &\rightarrow \gamma^\mu F_1^{(f_2)}(q^2)+\frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2^{(f_2)}(q^2)\,, \end{aligned} \tag{8.3.8}$$

If we now assume that $-q^2<<M_i^2$, then the vertex functions become

$$i\,V^\mu_{f_1}(p_1,k_1)\ \approx\ \gamma^\mu\,,\qquad i\,V^\mu_{f_2}(p_2,k_2)\approx\gamma^\mu \tag{8.3.9}$$

since both sets of form factors have $F_1(q^2)\approx F_1(0)=1$. Furthermore, the heavy fermions make a neglible contribution to $\Pi(q^2)$ when $-q^2<<M_i^2$. Thus, at small q^μ the term in (8.3.6), including the one-loop contribution, is to a very good approximation

$$\begin{aligned} &\frac{(-ie)^2}{1-\frac{e^2}{12\pi^2}R\left(\frac{-q^2}{m^2}\right)}\frac{-i\,\eta_{\mu\nu}\,\overline{u}^{r_1}(\vec p_1)\gamma^\mu u^{s_1}(\vec k_1)\overline{u}^{r_2}(\vec p_2)\gamma^\nu u^{s_2}(\vec k_2)}{-q^2+i\epsilon}\\ &=\frac{-4\pi}{\alpha^{-1}-\frac{1}{3\pi}R\left(\frac{-q^2}{m^2}\right)}\frac{-i\,\eta_{\mu\nu}\,\overline{u}^{r_1}(\vec p_1)\gamma^\mu u^{s_1}(\vec k_1)\overline{u}^{r_2}(\vec p_2)\gamma^\nu u^{s_2}(\vec k_2)}{-q^2+i\epsilon} \end{aligned} \tag{8.3.10}$$

The behavior of $R(z)$ for small and large z is

$$\begin{aligned} R(z) &\approx \frac{z}{5} \qquad z<<1\\ &\approx \log z \qquad z>>1\,. \end{aligned} \tag{8.3.11}$$

Hence, the effective value of the inverse coupling

$$\frac{1}{\alpha_{\rm eff}}=\frac{1}{\alpha}-\frac{1}{3\pi}R\left(\frac{-q^2}{m^2}\right)\,, \tag{8.3.12}$$

satisfies

$$\begin{aligned} q^2\frac{d}{dq^2}\frac{1}{\alpha_{\rm eff}} &\approx 0 \qquad -q^2<<m^2\\ &\approx -\frac{1}{3\pi} \qquad -q^2>>m^2\,. \end{aligned} \tag{8.3.13}$$

Thus the effective fine-structure increases with increasing $-q^2$. To a good approximation, the inverse coupling stays fixed for $-q^2<m^2$ and starts running at a constant rate once $-q^2>m^2$. The effective fine-structure is $\alpha=1/137$ when $-q^2$ is less than the electron mass. Once $-q^2$ is greater than the electron mass, the inverse coupling decreases linearly with the log of the scale, until it reaches the mass of another charged particle (say the muon) where the rate changes again. At some point the inverse coupling is zero, hence the coupling is infinite. This value of $-q^2$ is called the Landau pole.

8.3.3 $g-2$: The anomalous Landé g-factor

In this section we compute the anomalous magnetic moment for a charged fermion (electron or muon, say). The one-loop calculation was originally done by Schwinger in 1948 and is considered one of the seminal triumphs in quantum field theory.

We first show how the magnetic moment relates to the form-factors. Suppose we have an external electromagnetic field $A_\mu(x)$. If the particle is at rest and a constant magnetic field $\vec{B}$ is turned on, then the change in energy is $-\vec{\mu}\cdot\vec{B}$ where $\vec{\mu}$ is the magnetic moment. The magnetic moment is related to the particle's spin by

$$\vec{\mu} = \frac{g\,e}{2m}\,\vec{s}\,, \tag{8.3.14}$$

where g is a dimensionless constant called the Landé g-factor.

The contribution to the Hamiltonian from an external field is

$$H_{\rm int} = e\int d^3x\,\overline{\psi}(x)iV^\mu(x)\psi(x)A_\mu(x) \tag{8.3.15}$$

where $V^\mu(x)$ is the vertex-function in coordinate space. We are treating $\psi(x)$ as a quantized field and A_μ as a classical background. The change in energy due to the magnetic field is then the expectation value of $H_{\rm int}$ for a single particle state at rest. In particular, we have the matrix elements for the spin states

$$(\Delta E)_{rs} = \lim_{q\to 0}\frac{\langle\vec{k}_1,r|H_{\rm int}|\vec{k}_2,s\rangle}{\langle\vec{k}_1,s|\vec{k}_1,s\rangle}\,, \tag{8.3.16}$$

where $q^\mu = k_1^\mu - k_2^\mu$. The denominator is given by $2k_1^0\,(2\pi)^3\delta^3(0)$. The numerator is

$$\begin{aligned}\lim_{q\to 0}\langle\vec{q},r|H_{\rm int}|\vec{0},s\rangle &= \lim_{\vec{q}\to 0} e\int d^3x e^{iq\cdot x}\overline{u}^r(\vec{k}_1)iV^\mu(q)u^s(\vec{k}_2)A_\mu(x)\\ &= \lim_{q\to 0} e\int d^3x e^{iq\cdot x}\overline{u}^r(\vec{k}_1)\left(\gamma^\mu F_1(0)+\frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2(0)\right)u^s(\vec{k}_2)A_\mu(x)\\ &= \lim_{q\to 0} e\int d^3x e^{iq\cdot x}\overline{u}^r(\vec{k}_1)\left(\frac{k_1^\mu+k_2^\mu}{2m}A_\mu F_1(0)-\frac{\sigma^{\mu\nu}}{2m}(1+F_2(0))\partial_\nu A_\mu(x)\right)u^s(\vec{k}_2)\end{aligned} \tag{8.3.17}$$

where in the last step we used the Gordon identity and integrated by parts. Assuming the particle is at rest, $A_0(x)=0$ and $A_i = \frac{1}{2}\varepsilon_{ijk}x^jB^k$, the first term does not contribute while the second gives

$$-\,(2\pi)^3\delta^3(\vec{0})\,e\,\vec{\sigma}_{rs}\cdot\vec{B}\,, \tag{8.3.18}$$

where we used that $\overline{u}^r(\vec{0})\sigma^{ij}u^s(\vec{0}) = 2m\,\varepsilon_{ijk}\sigma^k_{rs}$. Hence we have that

$$(\Delta E)_{rs} = -\frac{2(1+F_2(0))\,e}{2m}\,\tfrac{1}{2}\,\vec{\sigma}_{rs}\cdot\vec{B}\,. \tag{8.3.19}$$

Comparing with (8.3.14), we see that the Landé g-factor is

$$g = 2 + 2F_2(0)\,. \tag{8.3.20}$$

The 2 was originally computed by Dirac. The second part of g is called the anomalous g-factor, or simply $g-2$.

To find $F_2(0)$, we sandwich the vertex function in (8.2.28) between two Dirac wave-functions. Using $\not{k}_2 u^s(\vec{k}_2) = m u^s(\vec{k}_2)$, $\overline{u}^r(\vec{k}_1)\not{k}_1 = \overline{u}^r(\vec{k}_1)m$, the relation

$$\overline{u}^r(\vec{k}_1)\not{k}_2\gamma^\mu\not{k}_1 u^s(\vec{k}_2) = \overline{u}^r(\vec{k}_1)[2m(k_1^\mu + k_2^\mu) - (3m^2 - q^2)\gamma^\mu]u^s(\vec{k}_2)\,, \tag{8.3.21}$$

as well as the symmetry between the Feynman variables x_1 and x_2, we find that

$$\begin{aligned}
W^\mu_{1,0} = \overline{u}^r(\vec{k}_1)i\,V_{1,0}(k_1,k_2)u^s(\vec{k}_2) &= -\,e^2\frac{\mu^{4-D}}{(4\pi)^{D/2}}\int_0^\infty \rho^{2-D/2}d\rho\int_0^1 dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)\\
&\times\overline{u}^r(\vec{k}_1)\Bigg[-2m^2\gamma^\mu+4m(k_1^\mu+k_2^\mu)(1-x_1-x_2)-2(\not{k}_2(1-x_2)-x_1m)\gamma^\mu(\not{k}_1(1-x_1)-x_2m)\\
&\qquad-\frac{(D-2)^2}{D}\rho^{-1}\gamma^\mu\Bigg]u^s(\vec{k}_2)\exp(-\rho((x_1+x_2)^2m^2-x_1x_2q^2+x_3\bar{\mu}^2)\,.\\
&= -\,e^2\frac{\mu^{4-D}}{(4\pi)^{D/2}}\int_0^\infty \rho^{2-D/2}d\rho\int_0^1 dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)\\
&\times\overline{u}^r(\vec{k}_1)\Bigg[\left(2m^2\big(-1+2x_3+x_3^2\big)-2(1-x_1)(1-x_2)q^2-\frac{(D-2)^2}{D}\rho^{-1}\right)\gamma^\mu\\
&\qquad+2m(k_1^\mu+k_2^\mu)x_3(1-x_3)\Bigg]u^s(\vec{k}_2)\exp(-\rho((1-x_3)^2m^2-x_1x_2q^2+x_3\bar{\mu}^2)\,.
\end{aligned} \tag{8.3.22}$$

Only the $k_1^\mu + k_2^\mu$ term contributes to $F_2(q^2)$ and we can see that it is finite. Using the Gordon identity in (8.3.4) and the definition of the form factors in (8.3.5), we find that

$$\begin{aligned}
F_2(0) &= \frac{4m^2e^2}{16\pi^2}\int_0^\infty d\rho\int_0^1 dx_3\int_0^{1-x_3}dx_2\,x_3(1-x_3)\exp(-\rho((1-x_3)^2m^2)\\
&= \frac{\alpha}{\pi}\int_0^1 dx_3\int_0^{1-x_3}dx_2\left(\frac{1}{1-x_3}-1\right) = \frac{\alpha}{2\pi}\,.
\end{aligned} \tag{8.3.23}$$

Since the $k_1^\mu+k_2^\mu$ term is not infrared divergent we have set $\bar{\mu}=0$. Using $\alpha = 1/(137.04)$, we find for $g-2$,

$$g - 2 \approx 0.002323\,. \tag{8.3.24}$$

The actual value is

$$g - 2 \approx 0.0023193043622 \pm 0000000000015\,. \tag{8.3.25}$$

Theoretically, we should only expect our answer to be correct up to order $\alpha^2/\pi^2 \approx 5\times 10^{-6}$, which is the size coming from the next order of perturbation theory. You can see that (8.3.24) differs from (8.3.25) by 4×10^{-6}. The full theoretical calculation has been done up to fifth order in perturbation theory, including effects from outside QED, and the theoretical uncertainty is more or less the same as the experimental uncertainty.

8.3.4 The other form factor

Let us now compute the contribution to $F_1(q^2)$. If we call $\delta F_1(q^2)$ the one-loop correction to the form factor which includes the counterterm, then using (8.3.22), (8.2.29) and the Gordon identity we find

$$\begin{aligned}\delta F_1(q^2) &= \frac{e^2}{16\pi^2}\left(\frac{1}{\varepsilon}-\gamma_E+\log 4\pi+\log\frac{\mu^2}{m^2}-1-R_3\left(1,1,y\right)\right)+\delta_{Z_1}\\ &\quad -\frac{e^2}{16\pi^2}\int_0^1 dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)\\ &\quad \times\left[\frac{4}{U+x_3w}-\frac{4(1-x_3)}{U}-\frac{2(1-x_3)^2}{U}-2\frac{(1-x_1)(1-x_2)y}{U+x_3w}\right],\end{aligned}\tag{8.3.26}$$

where $y=q^2/m^2$, $w=\bar\mu^2/m^2$ and $U=(1-x_3)^2-x_1x_2y$.

Let us consider this result at $q^2=0$. In this case we use the integrals

$$\begin{aligned}\int_0^1 dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)\frac{4}{(1-x_3)^2+x_3\frac{\bar\mu^2}{m^2}} &= \int_0^1 dx_3\int_0^{1-x_3}dx_2\frac{4}{(1-x_3)^2+x_3\frac{\bar\mu^2}{m^2}}\\ &= -2\log\left(\frac{\bar\mu^2}{m^2}\right)+\mathrm{O}\left(\frac{\bar\mu^2}{m^2}\right)\\ \int_0^1 dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)\frac{4}{(1-x_3)} &= 4\\ \int_0^1 dx_1dx_2dx_3\delta(1-x_1-x_2-x_3)2 &= 1\,,\end{aligned}\tag{8.3.27}$$

to find

$$\delta F_1(0) = \frac{e^2}{16\pi^2}\left(\frac{1}{\varepsilon}-\gamma_E+\log 4\pi+\log\frac{\mu^2}{m^2}+4-2\log\left(\frac{\bar\mu^2}{m^2}\right)\right)+\delta_{Z_1}\,. \tag{8.3.28}$$

If we now use that $\delta_{Z_1}=\delta_{Z_2}$ as well as the expressions in (8.2.21) and (8.2.22), we find that

$$\delta F_1(0)=0\,. \tag{8.3.29}$$

This result will hold to all orders in perturbation theory.

8.3.5 Soft photons and infrared divergences

In this section we give a physical interpretation of the infrared divergence found in the vertex function. We will find that in an actual physical measurement the log divergence coming from the zero mass-photon cancels and hence the photon mass can be safely set to zero.

We start this section by computing the correction to the form factor $\delta F_1(q^2)$, but now with $q^2 \neq 0$. For any nonzero value of q^2 the infrared divergence which we had diligently cancelled off for $q^2 = 0$ reappears.

For a general value of q^2 we have that

$$\begin{aligned}\delta F_1(q^2) &= -\frac{e^2}{16\pi^2}\Bigg\{R_3\,(1,1,y) + \int_0^1 dx_1 dx_2 dx_3 \delta(1-x_1-x_2-x_3)\Bigg[-\frac{4-4(1-x_3)-2(1-x_3)^2}{(1-x_3)^2+x_3 w}\\ &\qquad +\frac{4}{U+x_3 w} - \frac{4(1-x_3)}{U} - \frac{2(1-x_3)^2}{U} - 2\frac{(1-x_1)(1-x_2)y}{U+x_3 w}\Bigg]\Bigg\}, \qquad (8.3.30)\end{aligned}$$

We will only explore the limit $-q^2 >> m^2$, where one finds some simplification in the calculation. In this limit we find that the leading behavior comes from the last term in the square brackets, and only in the integration region where $x_3 \to 1$. Thus $\delta F_1(q^2)$ is approximately given by

$$\delta F_1(q^2) \approx -\frac{e^2}{8\pi^2}\int_0^1 dx_1 dx_2 dx_3 \delta(1-x_1-x_2-x_3)\frac{|y|}{(1-x_3)^2 + x_1 x_2 |y| + w}\,, \qquad (8.3.31)$$

where we have assumed that $-y >> 1$ and $w << 1$. Letting $x_1 = (1-x_3)\beta$, the integral becomes

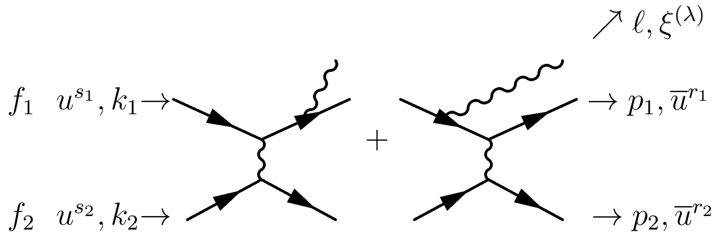
$$\begin{aligned}\delta F_1(q^2) &\approx -\frac{e^2}{8\pi^2}\int_0^1 (1-x_3)dx_3 \int_0^1 d\beta \frac{|y|}{(1-x_3)^2(1+\beta(1-\beta)|y|)+w}\\ &= -\frac{e^2}{16\pi^2}\int_0^1 d\beta \int_0^1 d\tau \frac{|y|}{\tau(1+\beta(1-\beta)|y|)+w}\\ &\approx -\frac{e^2}{16\pi^2}\int_0^1 d\beta \frac{|y|}{1+\beta(1-\beta)|y|}\log\frac{|y|}{w}\\ &\approx -\frac{e^2}{8\pi^2}\log|y|\log\frac{|y|}{w} = -\frac{e^2}{8\pi^2}\log\frac{-q^2}{m^2}\,\log\frac{-q^2}{\bar{\mu}^2}\,. \qquad (8.3.32)\end{aligned}$$

Significantly, the correction to the form factor is negative. Since this form factor multiplies γ^μ in the vertex function, we can interpret it as a correction to the electromagnetic coupling in a scattering amplitude. With this in mind, suppose we consider the tree-level scattering in (8.3.6) where we identify f_1 with the electron and f_2 with some heavy charged particle. Suppose the cross-section for this process is σ_0. If we include the one-loop vertex correction for the electron then the cross-section is modified to

$$\sigma_v \approx \left(1 - \frac{e^2}{8\pi^2}\log\frac{-q^2}{m^2}\,\log\frac{-q^2}{\bar{\mu}^2}\right)^2 \sigma_0 \approx \left(1 - \frac{e^2}{4\pi^2}\log\frac{-q^2}{m^2}\,\log\frac{-q^2}{\bar{\mu}^2}\right)\sigma_0\,, \qquad (8.3.33)$$

where $q^\mu = p_1^\mu - k_1^\mu$ and we assume that this is a "hard" scattering where $-q^2 >> m^2$. There can also be a correction from the f_2 vertex but this can be treated separately and will be briefly discussed at the end of the section. There is also a correction from the running of the coupling, but this does not have an infrared divergence so we ignore it.

Now consider the process in (8.3.6), but in addition to the outgoing f_1 and f_2 a "soft photon" with momentum ℓ^μ is also emitted. By soft, we mean a low energy photon with momentum $|\vec{\ell}| < 1/L$ where L is some large wave-length cutoff. At tree level, the relevant graphs are



(8.3.34)

There are also graphs where the photon attaches to the f_2 line, but these will be discussed at the end of the section along with the f_2 vertex correction.

The extra external photon modifies the fermion sandwich $\overline{u}^{r_1}(\vec{p}_1)\gamma^\mu u^{s_1}(\vec{k}_1)$ to

$$e\left(\overline{u}^{r_1}(\vec{p}_1)\not{\xi}^{(\lambda)}\frac{\not{p}_1+\not{\ell}+m}{(p_1+\ell)^2-m^2+i\,\epsilon}\gamma^\mu u^{s_1}(\vec{k}_1)+\overline{u}^{r_1}(\vec{p}_1)\gamma^\mu\frac{\not{k}_1-\not{\ell}+m}{(k_1-\ell)^2-m^2+i\,\epsilon}\not{\xi}^{(\lambda)}u^{s_1}(\vec{k}_1)\right)\,. \tag{8.3.35}$$

If we assume that $|\vec{\ell}| << m$ then we can approximate this combination as

$$e\,\overline{u}^{r_1}(\vec{p}_1)\gamma^\mu u^{s_1}(\vec{k}_1)\left(\frac{\xi^{(\lambda)}\cdot p_1}{\ell\cdot p_1}-\frac{\xi^{(\lambda)}\cdot k_1}{\ell\cdot k_1}\right)\,, \tag{8.3.36}$$

where we used the Clifford algebra and $\not{k}_1 u^{s_1}(\vec{k}_1) = m u^{s_1}(\vec{k}_1)$, $\overline{u}^{r_1}(\vec{p}_1)\not{p}_1 = \overline{u}^{r_1}(\vec{p}_1)m$. The cross-section for the inclusive process where the external photon has momentum $|\vec{\ell}| < 1/L$ is then

$$\sigma_{\rm soft} \approx \sigma_0\, e^2\sum_\lambda\int_{|\vec{\ell}|<L^{-1}}\frac{d^3\ell}{(2\pi)^3 2\ell^0}\left|\frac{\xi^{(\lambda)}\cdot p_1}{\ell\cdot p_1}-\frac{\xi^{(\lambda)}\cdot k_1}{\ell\cdot k_1}\right|^2\,. \tag{8.3.37}$$

We have assumed that $|\vec{\ell}|^2 << -q^2$ such that there is only a small correction to q^2 in the internal photon propagator, leaving it unchanged to a good approximation. Making the substitution

$$\sum_\lambda \xi_\mu^{(\lambda)}\xi_\nu^{(\lambda)*} \to -\eta_{\mu\nu}\,, \tag{8.3.38}$$

the cross-section simplifies to

$$\begin{aligned}\sigma_{\rm soft} &\approx \sigma_0\, e^2\int_{|\vec{\ell}|<L^{-1}}\frac{d^3\ell}{(2\pi)^3 2\ell^0}\left(\frac{2p_1\cdot k_1}{(\ell\cdot p_1)(\ell\cdot k_1)}-\frac{m^2}{(\ell\cdot p_1)^2}-\frac{m^2}{(\ell\cdot k_1)^2}\right)\\ &\approx \sigma_0\, e^2\int_{|\vec{\ell}|<L^{-1}}\frac{d^3\ell}{(2\pi)^3 2\ell^0}\,\frac{-q^2}{(\ell\cdot p_1)(\ell\cdot k_1)}\,,\end{aligned} \tag{8.3.39}$$

where the last approximation was made assuming that $-q^2 >> m^2$. To make life easier, we assume that $\vec{p}_1 = -\vec{k}_1$ so that $-q^2 \approx 4|\vec{p}_1|^2$. The cross-section can then be written as

$$\begin{aligned}
\sigma_{\rm soft} &\approx& \sigma_0 \frac{e^2}{8\pi^2}\int_{\ell<L^{-1}} \ell\, d\ell \int_{-1}^{1} d(\cos\theta)\, \frac{-q^2}{(\ell\cdot p_1)(\ell\cdot k_1)}\\
&\approx& \sigma_0 \frac{e^2}{8\pi^2}\int_{\ell<L^{-1}} \frac{d\ell}{\ell} \int_{-1}^{1} d(\cos\theta)\, \frac{4}{(1-\cos\theta-\frac{m^2}{q^2})(1+\cos\theta-\frac{m^2}{q^2})}\\
&\approx& \sigma_0 \frac{e^2}{2\pi^2}\log\frac{-q^2}{m^2}\int_{\ell<L^{-1}}\frac{d\ell}{\ell}\,. \qquad (8.3.40)
\end{aligned}$$

The integral over ℓ diverges as $\ell \to 0$. However, if we again assume the photon has a small mass $\bar{\mu}$ then ℓ^0 should be replaced with $\sqrt{\vec{\ell}^2+\bar{\mu}^2}$ and the integral essentially is cut off at $\ell = \bar{\mu}$. Hence we have

$$\sigma_{\rm soft} \approx \sigma_0 \frac{e^2}{4\pi^2}\log\frac{-q^2}{m^2}\log\frac{1}{L^2\bar{\mu}^2}\,. \qquad (8.3.41)$$

If we now combine the contribution of the soft photon with σ_v, we obtain

$$\sigma_v + \sigma_{\rm soft} \approx \left(1 - \frac{e^2}{4\pi^2}\log\frac{-q^2}{m^2}\log(-L^2q^2)\right)\sigma_0\,, \qquad (8.3.42)$$

where the photon mass has dropped out and thus can be set to zero. The interpretation of this combined cross-section is as follows: A scattering experiment involves using a detector to measure the momentum of particles. In order to measure a photon, the detector would have to be larger than the photon wavelength. Hence, if a photon with wavelength greater than the size of the detector should be produced, it will remain undetected. If the detector size is L, then the $f_1f_2 \to f_1f_2$ scattering cross-section will also include the contributions from $f_1f_2 \to f_1f_2 + \gamma$ if the photon wavelength is greater than L. As we make our detector bigger, fewer of the soft photon events are included in the cross-section and hence the cross-section decreases as L increases.

Note that we could play the same game with the other vertex in (8.3.6), where we would again find that the logarithmic divergence with the photon mass in the vertex function cancels with the corresponding log in the soft photon cross-section coming from a soft photon attaching to the f_2 propagator.