PHYSICS 330: Quantum Field Theory I

Lecturer: Professor Bernhard Mistlberger

Notes by: Andrew Lin

Autumn 2022

Introduction

Professor Mistlberger can be reached at bernhard@slac.stanford.edu, and the TA (Kevin Zhou) can be reached at knzhou@stanford.edu. This course will have two lectures a week (on Tuesdays and Thursdays) from 9-10:20am, and we may attend optional section classes Thursdays from 12-1:20pm (for further discussion on homework problems and material) in GESB 150.

Homework will be assigned weekly, going over examples that are related to lecture and helping us keep up to date with the mathematics. They'll be due on Tuesdays at midnight, and the first one will be released this afternoon. (Grades will be based on the best eight of our ten homework grades, as well as a final take-home exam exercise.) For any questions, we can attend Professor Mistlberger's office hours from 5-6pm on Mondays on Zoom.

Since the students here have different academic backgrounds, we'll start off relatively slowly and encourage questions. Our goal is to cover the first seven chapters of Peskin and Schroeder's "An Introduction to Quantum Field Theory," hopefully making us familiar with the ideas and basic structures in the subject. (There is lots of material that we can find in other books too – some good sources are Srednicki's "Quantum Field Theory," Itzykson and Zuber's "Quantum Field Theory," Zee's "Quantum Field Theory in a nutshell," and Weinberg's "The Quantum Theory of Fields." The last of these texts has lots of material and is written by a giant in the field, but is potentially hard to learn from on a first pass.)

1 September 27, 2022

Today's lecture will mostly be a "philosophy lesson," thinking about the general preamble to QFT and the question of "why." When the field first started, the goal was to **unify quantum mechanics and special relativity**, but the problem is that quantum mechanics works with small scales (such as atoms) and special relativity with fast speeds (near the speed of light). The field of (high-energy) particle physics is a natural place where these areas interact, and that's where everything started, but QFT can be applied today to condensed matter physics (crystal excitations and interactions), cosmology (the Big Bang, inflation, correlation), gravitational wave physics.

- One consequence of Einstein's $E = mc^2$ is that if we put a certain amount of energy into a box and shake it hard enough, basically anything can come out of it. Quantum mechanics lacks a way to explain how this process actually occurs, and facilitating the creation of new particles and antiparticles can be answered by QFT.
- The problem of causality (one of the ideas of relativity) will also come into play – if we consider the correlation function $U(t) = \left\langle x \middle| e^{-iHt} \middle| x_0 \right\rangle$ for the free Hamiltonian $H = \frac{\vec{p}^2}{2m}$, we can insert a complete set of states and

calculate

$$U(t) = \int \frac{d^3p}{(2\pi)^3} \left\langle x \middle| e^{-i\frac{\vec{p}^2}{2m}t} \middle| p \right\rangle \langle p|x_0\rangle = \left(\frac{m}{2\pi i t}\right)^{3/2} e^{-im(x-x_0)^2/(2t)},$$

and the point is that we'll always have some probability density to carry x to x_0 in a fixed time t, which is bad because particles are only supposed to be able to travel at (less than or equal to) the speed of light. And even if we use a relativistic version of this with $E = \sqrt{p^2 + m^2}$, that still doesn't solve the problem for us! So a different framework is indeed needed, and QFT does this in a weird but elegant way.

- In QFT, we are postulating that everything in the universe (except gravity) can be described with just a set of fields. So there's a level of reductionism here – we're saying the fields are spanning the universe and the rules are the same everywhere. However, while there's been trouble making observations and measurements to actually probe the relevant scales, there doesn't seem to be any kind of modification needed to the theory like in classical mechanics (except gravity). But that doesn't mean we should be studying subjects like chemistry with QFT – it's not very tractable.

Fact 1

We'll be using natural units in this class – there's only a few actual numbers that the universe throws at us, such as the cosmological constant $\Lambda \approx 10^{-52}$ m^{-2}, Newton's constant $G_N \approx 6.7 \times 10^{-11}$m^2/kg $\cdot$ s, the speed of light $c = 3 \times 10^8$ m/s, Planck's constant $\hbar \approx 10^{-34}$J $\cdot$ s, and the Higgs mass $m_h \approx 125$ GeV. We won't think much about the first two of these, but we will use units where $\hbar = c = 1$ (because for the purposes we're dealing with, this is the natural sense of scale). And we can use dimensional analysis to recover usual units if we need.

We'll use square brackets to denote dimension in mass, so $[m^n] = n$ means the dimension of mass in that quantity is n. So we're setting $[\hbar] = [c] = 0$ and $[m_n] = 1$, and following through the calculations yields $[s] = [L] = -1$ (so 1 second or 1 meter has units of inverse mass in natural units). Our energy unit will be the electron volt (1 eV $\approx 1.6 \times 10^{-19}$ kg $\cdot$ m^2/s^2, which is the same as 1.783 kg c$^2 \cdot 10^{-36} = 5.1 \times 10^6 \frac{\hbar c}{m}$. So we can switch back and forth between kilograms, meters, and seconds by just introducing the necessary $\hbar$s and cs, and we'll drop the additional $\hbar$s and cs from here. And this is convenient because the proton mass is about 938 MeV (so around 1 GeV), while the mass of the electron is 511 keV and the mass of the Higgs boson is about 125 GeV. Energy units should generally be thought of as **corresponding to length scales** – for example, it's useful to keep in mind that $\frac{1}{1\text{ fm}} = 200\text{ MeV}\frac{1}{\hbar c}$.

Remark 2. *Tossing $\hbar$ and c is done because in most problems we discuss, they don't add much to the calculation. But keeping mass as a relevant quantity will be important, and we'll see that moving forward. For example, we find that the charge radius of a proton is $\frac{1}{200MeV}$, which is comparable to its mass. (And we should remember that mass and energy are basically equivalent.) However, we should make sure this correspondence is only thought about in terms of elementary particles rather than composite structures.*

Example 3

As a naive example, light with a wavelength of 600 nm translates to an energy of roughly 2 eV (using $E = h\nu = \frac{\hbar c}{\lambda}$), so we can resolve structures at a 600 nm length scale given a particular microscope of that energy. Electron microscopes have an energy of 511 keV, corresponding to a length scale of 2.5×10^{-12} m – beyond the rest mass, we start getting in regimes where everything becomes quantum mechanical. (The best we can actually do is 50×10^{-12} m with current technology.)

Our next point of discussion is the question "what are fields?" For example, temperature is a scalar field, where there is a temperature value at every point in spacetime. (And the electric field is similar but is instead a vector field.)

The subsequent questions are then where those fields comes from and what they're made of, and we'll answer those in due time. We usually take think of coordinates (x, t) as operators $(\hat{x}, \hat{t})$ in quantum mechanics, but **that won't be done in quantum field theory** – t and x will actually just serve as coordinates, and instead the **fields** $\phi(x, t)$ will be what are promoted to operators $\hat{\phi}$ instead. So everything acting on wavefunctions will be a field operator, and once we learn how to do quantization of fields we'll start to see expressions like $\psi(x, t)\,|\Omega\rangle$. (And the reason that quantizing x and t doesn't work is that when we try it, we run into negative probabilities, lack of ground states, and so on. But we can read Srednicki for more on this.)

Our intuition might be that each point in spacetime influences its "local" neighbors, and to reach other points we must propagate through spacetime. That's what will ultimately be necessary for being compatible with relativity, and we need a way to describe time-evolution. In classical quantum mechanics, we have this idea that there is a differential equation

$$i\partial_0\,|\psi\rangle = H(\phi, \partial\phi)\,|\psi\rangle\,,$$

where we introduce canonical coordinates and their derivatives (momenta). To make sure that our theory is indeed local, it's important that our evolution operator is only being evaluated at a single point, and then Lorentz invariance requires us to have as many time derivatives as spatial derivatives. But we still haven't discussed what these fields are made of – at the end of the day, the idea is that we **put a harmonic oscillator at every spacetime point**, and they interact with their neighbors. We then run into issues with infinities everywhere, and the work that we'll be doing is to remedy this in a clever way. And one way to make sure this is "relativistically fine" is to describe everything with wave equations:

Definition 4

The **Klein-Gordon equation** is the differential equation $\partial_\mu^2\phi + m^2\phi = 0$, where we are using four-indices: we have $\partial_\mu = \frac{\partial}{\partial x^\mu}$ and $\partial_\mu^2 = \frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x^\mu}$, where we use the metric $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$. (If we are using three-indices instead, we'll use the Euclidean metric.)

To understand where the Klein-Gordon equation comes from, we can start by imagining this problem in two dimensions. Suppose we have a string stretched from 0 to x_0 with a vertical displacement $f(x)$ at point x; the force equation due to tension is then

$$F = T\left[\frac{df(x+dx)}{dx} - \frac{df(x)}{dx}\right] = T\frac{d^2f}{dx^2}dx$$

and by Newton's second law this is the same as

$$ma = m\frac{d^2f}{dt^2} = \rho dx\frac{d^2f}{dt^2}.$$

Collecting terms, this gives us the familiar wave equation

$$\frac{d^2f}{dt^2} - c^2\frac{d^2f}{dx^2} = 0, \quad c = \sqrt{\frac{T}{\rho}}.$$

(In other words, we get $\partial_\mu^2 f = 0$.) But if we now embed our string into a sheet of rubber (which restores the string back) and get an additional $F = -Y dx f(x, t)$ force (where Y is Young's modulus), we end up instead with the equation $\partial_\mu^2 f + m^2 f = 0$, where $m = \sqrt{\frac{Y}{\rho}}$. So this equation can indeed describe a classical system that we're used to, but we'll describe relativistic particle mechanics with it going forward and we'll discuss that next time.

2 September 29, 2022

Last lecture, we ended by discussing the relativistic Klein-Gordon wave equation – it'll play an important role in quantum field theory, and we saw a classical example (of a string embedded in rubber) in which that wave equation naturally arises. The equation looks like

$$(\partial_\mu^2 + m^2)\phi = 0, \quad \partial_\mu^2 = \frac{\partial}{\partial t^2} - \frac{\partial}{\partial x_1^2} - \frac{\partial}{\partial x_2^2} - \frac{\partial}{\partial x_3^2}$$

(in which we've already set natural units to suppress the factor of c^2 in front of the spatial derivatives). Today, we'll use a classical system to see what we'll be doing for a good amount of the rest of the course. This should give us some intuition, and it should only be followed somewhat heuristically to give us the general philosophy of the approach.

Example 5

Consider a one-dimensional crystal consisting of a line of atoms each with some vertical displacement, and let ϕ_n be the displacement of the nth atom. There will be some kind of correlative restorative force between the atoms, as well as an overall restoring force (like the "rubber" from last time). This system has Hamiltonian

$$H = E_{\text{kin}} + V = \sum_{n=-\infty}^{\infty} \frac{1}{2}(\partial_0\phi_n)^2 + \frac{1}{2}(\phi_n - \phi_{n+1})^2 + \frac{1}{2}m^2\phi_n^2$$

(Here $\partial_0 = \partial_t$, and the second term will look more like a spatial derivative if we shrink the distance between the atoms and rescale appropriately.)

We'll impose commutation relations as we do in quantum mechanics, specifically

$$[\phi_n, \phi_{n'}] = [\partial_0\phi_n, \partial_0\phi_{n'}] = 0, \quad [\phi_n, \partial_0\phi_{n'}] = i\hbar\delta_{n,n'} = \phi_n\partial_0\phi_{n'} - \partial_0\phi_{n'}\phi_n.$$

So ϕ_n is playing the role of q, and $\partial_0\phi_n$ is playing the role of $\dot{q}$. In other words, ϕ_n "creates" displacement and $\partial_0\phi_{n'}$ "changes" it, and quantum mechanics is saying that these two operations do not commute. Then like in Lagrangian mechanics, we have the canonical momentum

$$\Pi_n = \partial_0\phi_n = \frac{\partial L}{\partial(\partial_0\phi_N)}$$

(we're being sloppy about Lagrangian vs Lagrangian density; this will be clarified later) where

$$L = E_{\text{kin}} - V = \sum_{n=-\infty}^{\infty} \pi_n\partial_0\phi_n - H_n.$$

This Lagrangian has the symmetry $n \mapsto n+1$, so it's useful to do a Fourier decomposition: we write

$$\phi_n = \int_{-\pi}^{\pi} \frac{dk}{(2\pi)} e^{ikn}\tilde{\phi}(k), \quad \Pi = \int_{-\pi}^{\pi} \frac{dk}{(2\pi)} e^{ikn}\tilde{\Pi}(k),$$

where we make the particular choice that $\tilde{\phi}^\dagger(k) = \tilde{\phi}(-k)$ and $\tilde{\Pi}^\dagger(k) = \tilde{\Pi}(-k)$. This then tells us some information about commutation relations:

$$[\tilde{\phi}(k), \tilde{\phi}(k')] = [\tilde{\Pi}(k), \tilde{\Pi}(k')] = 0, \quad [\tilde{\phi}(k), \tilde{\Pi}(k')] = i\sum_{n=-\infty}^{\infty} e^{-i(k-k')n} = 2\pi\delta(k-k'),$$

as long as we make the assumption that $\tilde{\phi}$ and $\tilde{\pi}$ are periodic. Putting this back into our Hamiltonian, we have

$$H = \frac{1}{2}\int_{-\pi}^{\pi}\frac{dk}{2\pi}\left[\tilde{\Pi}^\dagger(k)\tilde{\Pi}(k) + \tilde{\phi}^\dagger(k)[m^2 + 2(1-\cos k)]\tilde{\phi}(k)\right].$$

(We won't worry about the dimensions of m here – just treat it as a parameter.) But now if we set $\omega_k = \sqrt{m^2 + 2(1-\cos(k)}$ and let

$$a_k = \frac{1}{4\pi\omega_k}\left[\omega_k\tilde{\phi}_k + i\tilde{\Pi}(k)\right], \quad a_k^\dagger = \frac{1}{4\pi\omega_k}\left[\omega_k\tilde{\phi}_k - i\tilde{\Pi}(k)\right],$$

we find that

$$[a_k, a_k^\dagger] = \delta(k-k'), \quad [a,a] = [a^\dagger, a^\dagger]] = 0,$$

so that our crystal Hamiltonian simplifies to

$$H = \frac{1}{2}\int_{-\pi}^{\pi} dk\omega_k[a(k)^\dagger a(k) + a(k)a^\dagger(k)]$$

which is the harmonic oscillator Hamiltonian in Fourier space. (This should all look similar to converting from x and p to creation and annihilation operators.) We can then check the commutators of H with a and $a^\dagger$; if we have an energy eigenstate $|E\rangle$ of energy E, then

$$H(a_k|E\rangle) = Ea_k|E\rangle + [H, a_k]|E\rangle = (E-\omega_k)(a_k|E\rangle),$$

and similarly we find that

$$H(a_k^\dagger|E\rangle) = (E+\omega_k)(a_k^\dagger|E\rangle).$$

So a_k and $a_k^\dagger$ lower and raise the energy by some factor, and that allows us to create the spectrum by defining the vacuum state $|0\rangle$ satisfying $a_k|0\rangle = 0$ and then having $|\omega_k\rangle = a_k^\dagger|0\rangle$ as usual. And the way these energy eigenstates look is that we have the kth Fourier mode by collectively oscillating the particles in a sine shape – this gives rise to **phonons** on a crystal. So going away from displacements ϕ and turning to collective "particles" ω_k will give us a sense of what quantum fields are doing – there, we'll promote fields $\phi(\vec{x}, t)$ to act on wavefunctions, where $\vec{x}$ and t will be the points in spacetime that tell us where oscillators are sitting.

To put this all on more systematic footing, we'll now think about **quantization**. The two main ways to do this are **canonical quantization** and **path integrals**, and we'll discuss the latter only in QFT II. The former builds up the concepts in a more pedestrian way, but it'll take more work and be less "natural." Recall that the **least action principle** tells us that the motion taking us from t_0 to t_1 is the one that minimizes the action

$$S = \int_{t_0}^{t} L\,dt = \int d^4x\mathcal{L}(\phi, \partial_\mu\phi),$$

and the difference here is that $\mathcal{L}$ will be the Lagrangian density of fields (which we'll just call Lagrangian). So we want to solve $\delta S = 0$ to get the classical trajectory; using the chain rule yields

$$\begin{aligned}\delta S &= \int d^4x\left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi)\right] \\ &= \int d^4x\left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right)\right].\end{aligned}$$

(Here we're going to assume smoothness so that we can interchange infinitesimal changes and derivatives.) But if we demand that everything falls off to zero at the boundaries of the spacetime, the total derivative term at the end

will always be zero, and what we're left with is the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \phi} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right).$$

For example, with the Klein-Gordon field, we have

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{m^2}{2}\phi^2 \implies \frac{\partial \mathcal{L}}{\partial \phi} = -m^2\phi, \quad \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) = \partial_\mu(\partial^\mu \phi) = \partial_\mu^2 \phi,$$

so we get back $(\partial_\mu^2 - m^2)\phi = 0$. And Lagrangians are useful because they allow us to see symmetries such as Lorentz invariance, and we'll go a bit through that to make sure we're on the same page.

- In ordinary two-dimensional space, we rotate a point by multiplying by a rotation matrix

  $$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

  with the nice properties that $R(\gamma) = R(\theta)R(\delta)$ (the product of rotation matrices is a rotation matrix) and $1 = R(\theta)R^{-1}(\theta)$. And we can generate rotations out of infinitesimal movements – if $r = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then $R(\theta) = e^{\theta r}$.

- In three-dimensional space, we have the space $SO(3)$ of rotations generated by rotations around z, y, x:

  $$\vec{r} = \left\{ \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \right\}$$

  with the generators satisfying the Lie bracket $[r^i, r^j] = \varepsilon^{ijk} r^k$.

- Finally, the Poincare group consists of the transformations which are a Lorentz transformation Λ plus a translation a:

  $$x^\mu \mapsto x^{\mu'} = \Lambda^\mu{}_\nu x^\nu + a^\mu$$

  where the (linear transformation) $\Lambda^\mu{}_\nu$ is given by $\frac{\partial x'^\mu}{\partial x^\nu}$ and where $g^{\rho\sigma} = g^{\mu\nu}\Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu$, so the metric (we're using the $(+,-,-,-)$ metric) is left invariant under Lorentz transformations. Specifically, this includes translations $(1, a)$, homogeneous transformations $(\Lambda, 0)$, rotations $(R, 0)$ (in which only the spatial coordinates are affected), and Lorentz boosts. Examples of the latter two look like

  $$R^\mu{}_\nu = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad L^\mu{}_\nu = \begin{bmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

  respectively. Furthermore, the Poincare group also includes spatial inversion $(P, 0)$ (the diagonal matrix $\mathrm{diag}(1, -1, -1, -1)$), as well as time reversal $(T, 0)$ ($\mathrm{diag}(-1, 1, 1, 1)$) and the composition of those two $(PT, 0)$. We call transformations **proper** if $\det(\Lambda) = 1$ and $a = 0$ and **orthochronous** if $\Lambda^0{}_0 \geq 1$ (so time reversals are not allowed); the only proper orthochronous transformations are the three rotations and the three boosts, and they can always be decomposed as

  $$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu,$$

  where w is "infinitesimal" in the sense that $O(\omega^2) = 0$ is very small. Then plugging into the identity from before

tells us that

$$g^{\mu\nu} = \Lambda^{\mu}{}_{\rho} g^{\rho\sigma} \Lambda^{\nu}{}_{\sigma} = g^{\mu\nu} + \omega^{\mu\nu} + \omega^{\nu\mu} + O(\omega^2),$$

so we must have an antisymmetric tensor $\omega^{\mu\nu} = -\omega^{\nu\mu}$ and thus our tensor looks like

$$\Lambda^{\mu\nu} = \begin{bmatrix} 1 & \omega^{01} & \omega^{02} & \omega^{03} \\ -\omega^{01} & -1 & \omega^{12} & \omega^{13} \\ -\omega^{02} & -\omega^{12} & -1 & \omega^{23} \\ -\omega^{03} & -\omega^{13} & -\omega^{23} & -1 \end{bmatrix}$$

(notice that we have upper indices this time).

Finally, thinking about how our fields transform, imagine we have a scalar field (so for example there is a "blob of temperature" in our field). Then we can notice that $\phi(x) \mapsto \phi'(x') = \phi(\Lambda^{-1}(x'))$ (so it looks like the field in the old coordinates is reached by the **inverse** Lorentz transformation); this is something we should try going through on our own. And if we have a vector field instead, the direction of the field also has to transform – we have

$$A^{\mu} \mapsto A^{\mu'}(x') = \Lambda^{\mu}_{\nu} A^{\nu}(\Lambda^{-1} x).$$

(Things like spin are more complicated, and we'll talk about it later on.)

3 October 4, 2022

Last lecture, we looked at our first field theory (a classical field theory) and talked about it in terms of writing down an action S, leading us to the Euler-Lagrange equations which give us the equation of motion. Importantly, there will be an Euler-Lagrange equation for each field sitting in our space. In particular, the classic Klein-Gordon field Lagrangian (density) led us to the familiar equation $(\partial_\mu^2 + m^2)\phi = 0$.

We mentioned that symmetries will play a big role in our discussion going forward, and one important feature of the equations we're discussing is that they're Lorentz-invariant. So last time, we also looked at the Poincaré group (Lorentz transformations plus a translation) and started thinking about how coordinates transform.

Remark 6. *In general relativity, there is an effort to distinguish between active and passive Lorentz transformations, but we don't need to be so careful here. The main point is that "the field still looks the same."*

It turns out that Lagrangians help us figure out symmetries – for example, suppose we have a Lagrangian for a complex scalar field

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m\phi^* \phi$$

and consider the transformation $\phi \mapsto e^{i\alpha}\phi$ (so that $\phi^* \mapsto e^{-i\alpha}\phi^*$) for some constant α. This keeps the Lagrangian constant (meaning that the overall phase of ϕ doesn't matter), and we'll see less trivial (but useful) examples moving forward – this leads us into **Noether's theorem**, a useful way to establish symmetries in a system:

Proposition 7 (Noether)
Every continuous symmetry of the lagrangian $\mathcal{L}$ gives rise to a conserved current $j^\mu(x)$, such that the equations of motion yield $\partial_\mu j^\mu = 0$.

Proof. Let ϕ_a be a set of fields, and suppose that we have an infinitesimal transformation $\phi_a'(x) = \phi_a + \varepsilon \Delta\phi_a(x)$.

Then $\delta\phi_a = \varepsilon\Delta\phi_a$, so by the chain rule we have

$$\delta\mathcal{L} = \varepsilon\frac{\partial\mathcal{L}}{\partial\phi_a}\Delta\phi_a(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\varepsilon\partial_\mu(\Delta\phi_a).$$

Applying the same trick as last time of separating out a total derivative, we can write this as

$$= \varepsilon\left(\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right)\Delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\varepsilon\Delta\phi_a\right).$$

The first term here is zero by the usual Euler-Lagrange equations, and now if we have a symmetry then the corresponding transformation must satisfy $\delta\mathcal{L} = 0$. This gives us

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\varepsilon\Delta\phi_a\right) = 0 \implies j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\varepsilon\Delta\phi_a.$$

□

Notice that it's also okay for the Lagrangian to change as a total derivative $\partial\mathcal{L} = \partial_\mu\Lambda^\mu$ (this yields $\delta S = 0$), so we can actually have

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\varepsilon\Delta\phi_a - \Lambda^\mu$$

as long as Λ^μ keeps the action constant. For example, translations in the Poincaré group correspond to four symmetries and thus four conserved currents.

We also have corresponding **conserved charge** $Q = \int d^3x j^0(x)$, and notice that this means we have

$$0 = \partial_0 Q = \int d^3x\partial^0 j^0(x) = -\int d^3x\partial_i j^i(x),$$

where in the last equality we've used the fact that $\partial_\mu j\mu = 0$. And this last integral is (by the divergence theorem) $-\int_A d\vec{A}\cdot\vec{j}$ over the boundary A, which can be intuitively thought of as "only caring about the flux of our current within our piece of spacetime."

Example 8

Consider the Noether charges corresponding to the translation $x^{\mu'} = x^\mu + a^\mu$ (where a is infinitesimal).

The corresponding Lagrangian transformation is then

$$\mathcal{L}(x) \mapsto \mathcal{L}(x-a) = \mathcal{L}(x) - a^\mu\partial_\mu\mathcal{L},$$

so that $\delta\mathcal{L} = a^\mu\partial_\nu\mathcal{L} = a_\mu\partial_\nu(\eta^\nu_\mu\mathcal{L})$ (note that we'll regularly switch between η and g here, and $\eta^{\mu\nu}$ is always the "constant" $(+,-,-,-)$ metric). Inserting the metric here is a "common trick:" the variation in a field ϕ is then

$$\delta\phi = -a^\mu\partial_\mu\phi, \quad \delta\partial_\nu\phi = -a^\mu\partial_\mu\partial_\nu\phi.$$

Collecting terms and putting them into the definition of the conserved current, we get

$$j^\mu{}_\nu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \eta^\mu{}_\nu\mathcal{L},$$

where the second spacetime index ν indexes the different currents we get from the four different translations (for a single transformation we would only have j^μ) and where we've removed the arbitrary constant a^μ (we could choose it to pick out a particular component). And this $j^\mu{}_\nu$ that we've calculated is the **energy-momentum tensor** $T^\mu{}_\nu$, which is conserved (though it's not the one we're used to from general relativity, there is some connection). The

corresponding conserved charge is then

$$p_\nu = \int d^3x T^0{}_\nu = \int d^3x \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} \partial_\nu \phi - \eta^0{}_\nu \mathcal{L} \right),$$

and in particular we have

$$p^0 = \int d^3x T^{00} = \int d^3x \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} \partial_0 \phi - \eta^{00} \mathcal{L} \right),$$

and remembering that the canonical momenta are defined as $\pi(x) = \frac{\partial \mathcal{L}(x)}{\partial(\partial_0 \phi)}$, we see that

$$p^0 = \int d^3x \left(\Pi \partial_0 \phi - \mathcal{L} \right)$$

is the usual formula for the Hamiltonian (this $p\dot{q} - \mathcal{L}$ form is called a Legendre transformation). And similarly doing the same thing for three-momentum yields the momentum operator (exercise for us).

Remark 9. *In everything here, we have $\mathcal{L}$ as a function of x, but we shouldn't think of putting indices on it (it's "always contracted" and won't appear on its own).*

Now that we have our Hamiltonian, we can talk more about quantization, and we'll start by thinking about the canonical quantization of scalar fields with our Klein-Gordon Lagrangian

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

with $\Pi = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)}$ (so that we have $\Pi = \partial_0 \phi$ in this case). Then the Hamiltonian is

$$H = \int d^3x \mathcal{H} = \int d^3x \left(\Pi \partial_0 \phi - \mathcal{L} \right) = \int d^3x \left(\frac{1}{2}\Pi^2 + \frac{1}{2}(\partial_i \phi)^2 + \frac{1}{2} m^2 \phi^2 \right),$$

and now (much like we quantized x and p) we think of Π and ϕ as **operators**. If we imagine this to be a classical field theory in the Schrodinger picture, so that Π and ϕ are only spatially dependent (or we're only looking at a particular time slice), then the commutation relations we are imposing (similar to the one for x and p) are that

$$[\phi(\vec{x}), \phi(\vec{y})] = [\Pi(\vec{x}), \Pi(\vec{y})] = 0, \quad [\phi(\vec{x}), \Pi(\vec{y})] = i\hbar \delta^{(3)}(\vec{x} - \vec{y}).$$

Again introducing Fourier modes (this is a transformation that we can do)

$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} (a(p) e^{i\vec{p}\cdot\vec{x}} + a^\dagger(p) e^{-i\vec{p}} \cdot \vec{x}), \quad \Pi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_p}{2}} (a(p) e^{i\vec{p}\cdot\vec{x}} - a^\dagger(p) e^{-i\vec{p}} \cdot \vec{x})$$

where we require the Fourier space relation $\phi^\dagger(-p) = \phi(p)$ (so that ϕ is real) and $\omega_p = \sqrt{\vec{p}^2 + m^2}$, and inverting these relations by adding in a delta function term $\int d^3x e^{i\vec{p}\cdot\vec{x}} \Pi(\vec{x})$, we get

$$a(\vec{p}) = \int d^3x \frac{e^{-i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega_p}} \left(\omega_p \phi(\vec{x}) + i\Pi(\vec{x}) \right), \quad a^\dagger(\vec{p}) = \int d^3x \frac{e^{i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega_p}} \left(\omega_p \Phi^\dagger(\vec{x}) - i\Pi(\vec{x}) \right).$$

The commutation relations are all zero except $[a(\vec{p}, a^\dagger(\vec{p})] = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{p}')$, and plugging back into our Hamiltonian gives us

$$H = \int d^3p \omega_p [a^\dagger(\vec{p}) a(\vec{p}) + a(\vec{p}) a^\dagger(\vec{p}),$$

which by the commutation relations can also be written as

$$= \int d^3p \omega_p \left(a^\dagger(\vec{p}) a(\vec{p}) + \frac{1}{2}[a(\vec{p}, a^\dagger(\vec{p})] \right),$$

which looks a lot like the harmonic oscillator with a particular frequency. So we can do the same thing as last time, seeing how this acts on creation and annihilation operators to create our spectrum and Hilbert space of states. We have

$$[H, a^\dagger(\vec{p})] = \omega_p a^\dagger(\vec{p}), \quad [H, a(\vec{p})] = -\omega_p(\vec{p}),$$

so we do have creation and annihilation operators and we can define a vacuuum state $|0\rangle$ such that $a(\vec{p})\,|0\rangle = 0$ and excited states $|\vec{p}\rangle = \sqrt{2\omega_p}a^\dagger(\vec{p}\,|0\rangle$; we can then define states (which we'll eventually interpret as particle states)

$$|\vec{p}_1, \vec{p}_2, \cdots, \vec{p}_n\rangle = \sqrt{2\omega_{p_1}}\cdots\sqrt{2\omega_{p_n}}a^\dagger(\vec{p}_1)\cdots a^\dagger(\vec{p}_n)\,|0\rangle\,.$$

These $\vec{p}$s can be interpreted as momenta: indeed, $H\,|\vec{p}_1, \cdots, \vec{p}_n\rangle = (\omega_{p_1} + \cdots + \omega_{p_n})\,|\vec{p}_1, \cdots, \vec{p}_n\rangle$, and $a^\dagger(\vec{p})a^\dagger(\vec{q})\,|0\rangle = a^\dagger(\vec{q})a^\dagger(\vec{p})\,|0\rangle$ for any two momenta $\vec{p}, \vec{q}$ (because our commutation relations tell us that the $a^\dagger$ all commute, we have **bosons**). So we can form the states $(a^\dagger(\vec{p}))^n\,|0\rangle$ for any integer n.

Remark 10. *Trying this with "position excitations" instead of "momentum excitations" will be more challenging, but we'll discuss this later.*

Taking the momentum operator $P^\mu = \int d^3x \frac{\partial\mathcal{L}}{\partial(\partial_0\phi)}\partial^\mu\phi - g^{\mu 0}\mathcal{L}$ from our earlier discussion of Noether's theorem, we find that

$$P^i = -\int d^3x \Pi(\vec{x})\partial^i\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3}p^i a^\dagger(\vec{p})a(\vec{p})$$

and so we have a number operator $N = a^\dagger(\vec{p})a(\vec{p})$. And because $[P^i, H] = 0$, what we learn here is that the states $|\vec{p}\rangle$ can be taken to be momentum eigenstates with energy $\omega_{\vec{p}}$. So the point is that we do a Fourier transform and see how the Hamiltonian acts in Fourier space, giving us the usual thing from classical physics.

4 October 6, 2022

Last lecture, we wrote down the proceedings of how to quantize classical field theories. To review the major steps, we started with the classical field theory in four dimensions with Lagrangian density $\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{m^2}{2}\phi^2$ to get the canonical coordinates $\Pi = \frac{\partial\mathcal{L}}{\partial(\partial_0\phi)} = \partial_0\phi$. This allowed us to write down the Hamiltonian density, giving us familiar conserved quantities.

From there, we quantize by promoting our fields ϕ, Π to operator versions (this seems to break Lorentz invariance, but we'll ignore it for now) such that $[\phi(\vec{x}), \phi(\vec{y})] = [\Pi(\vec{x}), \Pi(\vec{y})] = 0$ and $[\phi(\vec{x}), \Pi(\vec{y})] = i\delta^{(3)}(\vec{x} - \vec{y})$. If we then do a Fourier transform (mode expansion), using the frequency $\omega_p = \sqrt{\vec{p}^2 + m^2}$, we have

$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\sqrt{\frac{\omega_p}{2}}\left[ae^{i\vec{p}\cdot\vec{x}} - a^\dagger e^{i\vec{p}\cdot\vec{x}}\right], \quad \Pi(\vec{x}) = -i\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left[e^{i\vec{p}\cdot\vec{x}}a + a^\dagger e^{i\vec{p}\cdot\vec{x}}\right].$$

(More generally, we could have used a and b instead of a and $a^\dagger$, but if we want ϕ to be real we need $b = a^\dagger$.) Requiring that the only nontrivial commutator here is $[a(\vec{p}), a^\dagger(\vec{p}')] = (2\pi)^3\delta^{(3)}(\vec{p} - \vec{p}')$, we recover the Hamiltonian for a harmonic oscillator at **every point** in momentum space. So we can find similar energy eigenstates by starting with a ground state and applying creation $a^\dagger_{p_i}$ operators to them, and $|p\rangle$ are also energy eigenstates of our Hamiltonian with a definite momentum.

Remark 11. *It turns out that the vacuum state is not always unique – it will matter if we have something called a "topological excitation," This comes up in various applications but is far beyond the scope of this class.*

All of the field theories can be quantized in this kind of way, but there's a lot of clean-up that we're going to

have to do first. In particular, we chose a particular time-slice $t = 0$ for all of this argument, which breaks Lorentz invariance.

- Before that, we'll discuss the bizarre **vacuum energy** of this Hamiltonian. We can rewrite this Hamiltonian

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3}\omega_p \left[a^\dagger(\vec{p})a(\vec{p}) + a(\vec{p})a^\dagger(\vec{p})\right] \\ &= \int \frac{d^3p}{(2\pi)^3}\omega_p \left[2a^\dagger(\vec{p})a(\vec{p}) + [a(\vec{p}), a^\dagger(\vec{p})]\right], \end{aligned}$$

  and it's now a problem that this commutator is a delta function, since $\delta^{(3)}(0)$ is "infinite." Furthermore, integrating that delta function over d^3p gives us another infinity. These integrals actually end up telling us about what's "going to haunt us with quantum field theory," and we'll try to talk about them now. The $\int d^3p$ infinity has to do with extremely high energies and momenta, and we call that an **ultraviolet singularity**. On the other hand, usually we have to specify how delta functions come about to get a smooth approximation, but here because we did a mode expansion we know that $\delta^{(3)}(\vec{0}) = \int d^3x e^{i(\vec{p}-\vec{p})\cdot\vec{x}} = \int d^3x$. So that infinity comes from distances that are very far apart (large volume) or momenta $\vec{p} \to 0$, and that's called an **infrared singularity**. Relatedly, there is a **collinear singularity** if we have multiple particles with the same momentum $\vec{p} = \vec{p}'$.

  But the point is that whatever this "infinity term" is, we'll set that as our baseline energy E_0, and we always measure energy values of one state **with respect** to another and look at $H - E_0$. (And there isn't really any way of getting around this – as far as we understand, there aren't ways to measure this E_0, but we're somehow saying that there is an "infinite energy density everywhere.")

  Remark 12. *Notice that in momentum space, our Hamiltonian looks like $H = (\omega_p\phi - i\Pi)(\omega_p\phi + i\Pi)$, and if we impose commutation relations on this Hamiltonian directly we instead get $a^\dagger a$ with no additional infinity term. So we can think of this as the true Hamiltonian of nature, but there's no real physical difference in the two cases because we can always add a constant to H. So this is not something we can really resolve, and it's not something we should worry about much.*

  One tool we do have to help deal with infinities is the concept of **operator ordering** (which is denoted by putting : before and after the operators). Specifically, we always put annihilation operators to the right and creation operators to the left:

$$: a_p a_p^\dagger : = : a_p^\dagger a_p : = a_p^\dagger a_p$$

  In particular, this means our normal ordered Hamiltonian is given by

$$: H : = \int \frac{d^3p}{(2\pi)^3}\omega_p a_{\vec{p}}^\dagger a_{\vec{p}}.$$

  (We're not saying that we can ignore commutators in general – it's just that in this particular case it gets rid of the zero-point energy.)

- Our energy doesn't seem like it should be relativistically invariant, and indeed if we look at our eigenstates

$$|\vec{p}\rangle = \sqrt{2E_p}a^\dagger(\vec{p})\,|0\rangle\,,$$

  we have

$$\langle\vec{q},\vec{p}|=\rangle\,\sqrt{2E_p}\sqrt{2E_q}\,\langle 0|a(\vec{q})a^\dagger(\vec{p}|0\rangle = \sqrt{2E_p}\sqrt{2E_q}\delta^{(3)}(\vec{p}-\vec{q}) = 2E(\vec{p})\delta^{(3)}(\vec{p}-\vec{q}).$$

But now if we do a Lorentz boost in the x-direction with coordinates

$$p'_x = \gamma(p_x + \beta E), \quad E' = \gamma(E + \beta \frac{p}{x}) \implies E = \gamma(E' - p'\beta), \quad p_x = \gamma(p' - E'\beta)$$

(still no time-dependence here), we find that the inner product above can be simplified via

$$\begin{aligned} E(\vec{p})\delta^{(3)}(\vec{p} - \vec{q}) &= E\delta(p_x - q_x)\delta^{(2)}(p_{y,z} - q_{y,z}) \\ &= \gamma(E' - p'_x\beta)\delta(\gamma(p'_x - E'_x\beta - q'_x + E'_y\beta))\delta^{(2)}(p_{y,z} - q_{y,z}) \\ &= \frac{\gamma(E' - p'_x\beta)}{(\gamma(E' - p'_x\beta))/E'}\delta^{(3)}(\vec{p}' - \vec{q}'), \end{aligned}$$

where in the last step we've used that $\delta(cx) = \frac{1}{|c|}\delta(x)$ for any constant c (filling in the details for the Jacobian is a good exercise). So the point is that we end up with $E'\delta^(3)(\vec{p}' - \vec{q}')$ again, and Lorentz invariance is still present.

So if we think about the process of "creating particles," we can have $\phi(x)$ act on the vacuum to get

$$\phi(\vec{x})\,|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}a^\dagger(\vec{p})e^{i\vec{p}\cdot\vec{x}}\,|0\rangle = \int \frac{d^3p}{(2\pi)^2}\frac{1}{2E(\vec{p})}e^{i\vec{p}\cdot\vec{x}}\,|\vec{p}\rangle\,.$$

We can then rewrite this all as

$$= \int \frac{d^4p}{(2\pi)^4}(2\pi)\delta_+(p^2 - m^2)e^{i\vec{p}\cdot\vec{x}}\,|p\rangle\,,$$

where we define $\delta_+(p^2 - m^2) = \delta(p^2 - m^2)\Theta(p^0)$ where Θ is the Heaviside function. (Basically, we're picking out the p_0 component that gives us the energy, because there's a positive and negative solution for p_0 that is picked out by the delta function in $\int dp^0\delta(p_0^2 - p_i^2 - m^2)\Theta(p^0)$, and we only want to take the positive one.) This will be interpreted physically as "creating a state of $\vec{x}$:" to make analogies to quantum mechanics, we can calculate that

$$\langle 0|\phi(\vec{x})|\vec{p}\rangle = e^{i\vec{p}\cdot\vec{x}},$$

which is the position space wavefunction (since we're still doing everything at $t = 0$ here), and similarly

$$\langle 0|\,\phi(x) = \int \frac{d^4p}{(2\pi)^4}\delta_+(p^2 - m^2)e^{i\vec{p}\cdot\vec{x}}\,\langle\vec{p}|$$

can be thought of as "annihilating a particle at $\vec{x}$." And because everything is contracted nicely, there's no problems with Lorentz invariance here. (And in general, we often think of things as momentum eigenstates, so we shouldn't think too much about the particles actually being localized at $\vec{x}$ – that's just a label.)

But we still haven't talked about time-dependence up to this point, and now we need to figure out how to bring it back into the picture. Recall that in the **Schrodinger picture**, our operators are static, while in the **Heisenberg picture**, operators involve in time. And for our purposes, it makes sense to take the latter and define our operators to evolve in time as

$$O(\vec{x}) = O(\vec{x}, t) = e^{iHt}O(\vec{x}, 0)e^{-iHt}, \quad \Pi(\vec{x}) = e^{iHt}\Pi(\vec{x}, 0)e^{-iHt}.$$

In momentum space, this looks like

$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(e^{-ipx}a(p) + a^\dagger(p)e^{ipx}\right)$$

where we now have that px are contractions of momentum and position **four-vectors** (rather than three-vectors), and

the time-dependence is encoded in

$$e^{iHt}a(p)e^{-iHt} = e^{-iE(p)(t)}a(\vec{p}).$$

We then can make sense of the expression for

$$\Pi = \partial_0\phi = -\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}p^0\left(e^{-ipx}a - e^{ipx}a^\dagger\right).$$

We then find that (taking another time-derivative)

$$-\partial_0^2\phi = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}(p^0)^2\left(e^{-ipx}a + e^{ipx}a^\dagger\right),$$

where $(p^0)^2 = \vec{p}^2 + m^2 = -\partial_i^2 + m^2$. Pulling that operator out of the integral, we thus have

$$-\partial_0^2\phi = (-\partial_0^2 + m^2)\phi \implies (\partial_\mu^2 + m^2)\phi(x) = 0,$$

which is our original Klein-Gordon equation. And notice that what we've managed to do is to create negative frequency states (which will turn out to correspond to antiparticles) which still have positive energies. This is not very cool for scalar fields because they will end up "being their own antiparticle," but we'll see soon (in exercises) that this does manifest in a cool way.

5 October 11, 2022

Last lecture, we followed classical quantization steps and introduced operators in the Heisenberg picture. Specifically, we defined (here x is a four-vector)

$$\phi(x) = \phi(\vec{x}, t) = e^{iHt}\phi(\vec{x}, 0)e^{-iHt},$$

which (for the Klein-Gordon Hamiltonian) yields

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega(\vec{p})}(e^{ipx}a(p) + e^{-ipx}a^\dagger(p)),$$

Then the definition of conjugate momentum $\Pi(x)$ is really the time-derivative of our field $\phi(x)$, and we can calculate that $\partial_0\Pi = \partial_0^2\phi = \partial_i\phi - m^2\phi$, getting back to the Klein-Gordon equation. We then found our spectrum and found physical interpretations of "creating particle states." Defining $|\vec{p}\rangle = \sqrt{2\omega_p}a^\dagger(p)|0\rangle$, we can write

$$\phi(x)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}e^{-ipx}|p\rangle,$$

which allows us to create and annihilate particles "at a particular state" x.

One thing we didn't clean up last time was the problem of **causality**, and in order to do that we'll study the vacuum correlation function

$$\langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\frac{1}{\sqrt{2\omega_q}}e^{-ipx+iqy}\left\langle 0\middle|a(p)a^\dagger(q)\middle|0\right\rangle,$$

where we've used the fact that the other term in $\phi(x)$ goes away because applying a to the vacuum gives us 0. Applying the commutator $\langle p|q\rangle = \sqrt{4\omega_p\omega_q}\left\langle 0\middle|a(p)a^\dagger(q)\middle|0\right\rangle = 2\omega_p(2\pi)^3\delta^{(3)}(p-q)$, this simplifies to

$$\langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{d^3q}{(2\pi)^3}\frac{1}{2\omega_p}e^{-ip(x-y)} = \int \frac{d^4p}{(2\pi)^4}2\pi\delta_+(p^2-m^2)e^{ip(x-y)} = D(x-y),$$

where D is what we call a **Green's function**. But now if $(x-y)^2 < 0$, meaning that our distance between x and y is outside the light-cone (we have spacelike separation), then we can calculate this function $D(x-y)$ by using Lorentz invariance and noticing that we can do a transformation so that x and y are both chosen to be at the same time-slice. So we'll set $x^0 = y^0 = 0$ and get some radial vector $\vec{r} = \vec{x} - \vec{y}$, and (exercise) we can check that $D(x-y) = -\frac{i}{2(2\pi)^2 r}\int_{-\infty}^{\infty} d|\vec{p}|\frac{e^{i|\vec{p}||\vec{r}|}|\vec{p}|}{\sqrt{|\vec{p}|^2+m^2}}$, and in particular this falls off exponentially as $r \to \infty$. So the correlation is nonvanishing, but that doesn't mean we have any issues with causality since we're not communicating any signals. Instead, the question is to ask whether we can impact measurements at x using measurements at y, and we do this by looking at commutators (because they tell us whether the order of measurement matters)

$$\langle 0|[\phi(x),\phi(y)]|0\rangle = [\phi(x),\phi(y)] = \int \frac{d^4p}{(2\pi)^4} 2\pi\delta_+(p^2-m^2)\left[e^{-ip(x-y)} - e^{-ip(y-x)}\right]$$

(as an important note, we set up **equal-time** commutation relations $[\phi(x),\phi(y)] = 0$, but the quantization is different if we're not doing things at a fixed time. So we have to actually go through the calculation here). We'll now undo the Lorentz invariant notation and rewrite in a more complicated way to get

$$\int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega(p)}\left[e^{-ip(x-y)} - e^{-(-i\omega(p))(x^0-y^0)-i\vec{p}(\vec{x}-\vec{y})}\right],$$

and then doing a substitution $\vec{p} \mapsto -\vec{p}$ and also "picking out the particular value of p^0" lets us rewrite this as

$$\langle 0|[\phi(x),\phi(y)]|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{dp_0}{2p_0}\left[e^{-ip(x-y)}\delta(p^0-\omega(p)) - e^{-ip(x-y)\delta(p^0+\omega(p))}\right]$$

(so we're now integrating with respect to d^4p again). Now we'll write these delta functions as contour integrals: fix some infinitesimal $\varepsilon > 0$ and define the integral

$$I = \oint \frac{d^4p}{(2\pi^4)} e^{-ip(x-y)}\frac{i}{2p^0}\left(\frac{1}{p^0-\omega(p)+i\varepsilon} - \frac{1}{p^0+\omega(p)+i\varepsilon}\right),$$

where this expression now requires us to have $x^0 > y^0$ as a constraint. Combining these terms yields

$$\boxed{I = \int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)}\frac{i}{(p^0+i\varepsilon)^2 - |\vec{p}|^2 - m^2}}.$$

Cauchy's residue theorem tells us that for any contour C enclosing some set of poles $\{z_i\}$ of a function f (meaning that f looks locally like $\frac{\tilde{f}(z_i)}{z-z_i}$) we have

$$\oint_C f(z)dz = 2\pi i\sum_{\{z_i\}}\tilde{f}(z_i).$$

We're actually only doing this contour integral in the p^0 plane, and the introduction of ε will be explained now: we have a pole at $p^0 = -\omega_p - i\varepsilon$ and at $p^0 = \omega_p - i\varepsilon$, while the ordinary d^4p integral travels along the real axis for p^0. So now we can "close the contour" and complete our contour C by basically traveling through a semicircle in the negative imaginary axis, taking its radius to infinity:

$$p^0 = r\cos\theta - ir\sin\theta, \theta \in [0,\pi].$$

In particular, we see that as $r \to \infty$, $e^{i(x^0-y^0)p^0} = e^{-r\sin\theta(x^0-y^0)}$ decays exponentially as long as $x^0 > y^0$, and the length of the arc is only linear in r. Thus the contour integral that we wrote only gets contributions along the real axis, giving us the boxed integral I, and using the residue theorem gets us back to $D(x-y)$ (which is basically integrating

$e^{-ip(x-y)}$. So if we then define the **retarded propagator**

$$D_R(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{(p^0+i\varepsilon)^2 - |\vec{p}|^2 - m^2} e^{ip(x-y)} = \langle 0|[\phi(x), \phi(y)]|0\rangle \, \theta(x^0 - y^0)$$

and analogously the **advanced propagator**

$$D_A(x-y) = \langle 0|[\phi(x), \phi(y)]|0\rangle \, \theta(y^0 - x^0),$$

we can check that $(\partial_\mu^2 + m^2)D_A(x-y) = -i\delta^{(4)}(x-y)$ (and the same for D_R). These Green's functions are useful because they relate to equations of motion if we have our Lagrangian $\mathcal{L}_I$ and we add in an additional interaction term ϕj (where j is some current), then our equations of motion become $(\partial^2 + m^2)\phi = j$. And to get a solution for j, we can use the Green's function in a convolution

$$\phi(y) = \int d^4x D_{A,R}(x-y) j(y)$$

(where we use the advanced or retarded propagator depending on the situation); indeed, we see that $(\partial^2 + m^2)\phi = i\int d^4y - i\delta^{(4)}(x-y)j(y) = j(x)$, so integrating Green's function (specifically, convolving it with the source) is useful for recovering ϕ.

Definition 13

The **Feynman propagator** is defined as

$$D_F(x-y) = \theta(x^0 - y^0)D(x-y) = \theta(y^0 - x^0)D(y-x) = \langle 0|\phi(x)\phi(y)|0\rangle + \langle 0|\phi(y)\phi(x)|0\rangle \, .$$

(In other words, if $x^0 > y^0$, we perform $\phi(y)$ first, and vice versa.)

In operator notation, we'll write this as

$$D_F(x-y) = \langle 0|T\{\phi(x)\phi(y)\}|0\rangle \, .$$

We can give this propagator an integral representation as well: indeed, we see that

$$D_F(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} \frac{i}{p^2 - m^2 + i\varepsilon},$$

where this time one pole will be below the real axis and the other will be above, because (using that ε is infinitesimal)

$$\frac{i}{p^2 - m^2 + i\varepsilon} = \frac{i}{(p^0 - \sqrt{\vec{p}^2 + m^2 - i\varepsilon})(p^0 + \sqrt{\vec{p}^2 + m^2 - i\varepsilon})} = \frac{1}{(p^0 - \sqrt{p^2 + m^2} - i\varepsilon)(p^0 + \sqrt{p^2 + m^2} - i\varepsilon)},$$

and we see that this indeed gives us $\frac{i}{2\omega_p}\left(\frac{1}{p^0 - (\omega_p - i\varepsilon)} - \frac{1}{p^0 + (\omega_p - i\varepsilon)}\right)$. So now the contour integral picks up either one residue or the other, but it depends on whether our semicircle is in the top half or bottom half of the imaginary plane: this is related to whether we have $x^0 > y^0$ or $x^0 < y^0$. So for the advanced propagator we close above the real axis, and for the retarded propagator we close below it, based on which side allows $e^{\pm ip(y-x)}$ to vanish along the circular arc.

With the way we've set everything up, it's a bit difficult to get rid of time-ordering – we'll continue along this way for a while, but we'll see soon that time will disappear from the picture. Everything so far has been free field theory, but now we'll talk a bit about interactions: it turns out that what we want to add into our Lagrangian looks like

$$\mathcal{L} = \mathcal{L}_0 - \frac{\lambda}{4!}\phi^4(x),$$

where λ is some arbitrary number. In general, we'll see products of fields (local field operators) like this appearing in our interaction term, and we'll see that having more than four powers of fields will not be relevant in most interaction theories. (Heuristically, that has to do with energy dimension, but there's a whole discussion of the Wilsonian renormalization group for that.) Similarly, we can then split our Hamiltonian into a free Hamiltonian and an interaction term

$$H = H_0 + H_{\text{int}}, H_{\text{int}} = \int \frac{\lambda}{4!}\phi^4 d^4x.$$

We can write out a mode expansion of the same ansatz

$$\phi(\vec{x}, t_0) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left[a(p)e^{i\vec{p}\cdot\vec{x}} + a^\dagger(p)e^{i\vec{p}\cdot\vec{x}}\right],$$

so that our initial fields will now take the form (this is now the **interaction picture**)

$$\phi_I(\vec{x}, t) = e^{iH_0(t-t_0)}\phi(\vec{x}, t_0)e^{-iH_0)(t-t_0)}.$$

In the Heisenberg picture, this means

$$\phi_H(x) = e^{iH(t-t_0)}e^{-iH_0(t-t_0)}\phi_I(\vec{x}, t)e^{iH_0(t-t_0)}e^{-iH(t-t-0} = U^\dagger(t, t_0)\phi_I(t, t_0)U(t, t_0)$$

for some unitary operator $U(t, t_0) = e^{iH_0(t-t_0)}e^{-iH(t-t_0)}$ involving the interaction term. We thus have $U(t_0, t_0) = 1$ and $i\partial_0 U(t, t_0) = e^{iH_0(t-t_0)}(H_0 - H)e^{-iH(t-t_0)}$, where $H - H_0$ is the interacting Hamiltonian $H_{\text{int, s}}$ in the Schrodinger picture. This then simplifies to

$$e^{iH_0(t-t_0)}H_{\text{int, S}}e^{-iH_0(t-t_0)}e^{iH_0(t-t_0)}e^{iH(t-t_0)} = H_{\text{int, I}}U(t, t_0) = i\partial_0 U(A, A_0).$$

6 October 13, 2022

Remark 14. *When we quantize a classical field theory, sometimes we have products of ϕs and πs which we could originally put in any order, but then the order ends up making a difference when we turn them into operators. It turns out tat this just adds a constant to the Hamiltonian (which may be infinite, but that's not really a problem), which never changes the physics and thus yields the same field theory.*

Remark 15. *The charge associated with the symmetry $\phi \to e^{i\alpha}\phi$ and $\phi^* \to e^{-i\alpha}\phi^*$ is a difference of number operators $N_a - N_b$; it turns out this will represent the number of particles and antiparticles, and it will turn out to show that particles are always created in pairs.*

Last lecture, we calculated the Green's function $D(x, y) = \langle 0|\phi(x)\phi(y)|0\rangle$, finding that $\langle 0|[\phi(x), \phi(y)]|0\rangle$ is zero for spacelike separated particles (which means causality is not violated). We then used it to write down the Feynman propagator

$$D_F(x-y) = \langle 0|T\{\phi(x), \phi(y)\}|0\rangle = \langle 0|\phi(x)\phi(y)|0\rangle\,\theta(x^0 - y^0) + \langle 0|\phi(y)\phi(x)|0\rangle\,\theta(y^0 - x^0)$$

$$= \lim_{\varepsilon\to 0^+}\int \frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}\frac{i}{p^2 - m^2 + i\varepsilon}$$

which is a manifestly Lorentz-invariant quantity which we obtain by thinking about contour integrals in the complex plane. The point is that by solving the free scalar field theory completely, we are now able to calculate correlation functions explicitly, and it will turn out that arbitrary correlation functions can be written in terms of this quantity $D_F(x_i - x_j)$.

But what we've been discussing recently is that we "turn on interactions" by making our Lagrangian more complicated, varying the original Lagrangian $\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2$ by adding a term $-\frac{\lambda}{4!}\phi^4$ for some very small λ. We did the same quantization procedure: doing a mode expansion and thinking of a and $a^\dagger$ as creation and annihilation operators, then converting to the **interaction** picture $\phi_I(\vec{x}, t) = e^{iH_0(t-t_0)}\phi(\vec{x}, t_0)e^{-iH_0)(t-t_0)}$, and we find that for the full time-evolution we want the **Heisenberg** picture with the time-evolution operator

$$U(t, t_0) = e^{iH_0(t-t_0)}e^{-iH(t-t_0)},$$

where H_0 is the original Hamiltonian and H is the full one. What we're going to do in the next two lectures is find a mathematically useful way to calculate correlation functions, arriving at a framework (Feynman rules) that lets us do this mathematics quickly. The point is to relate correlation functions like $\langle 0|\phi_H(x)\phi_H(y)|0\rangle$ to things that we can already calculate.

Last time, we found that we could evolve that time-evolution operator U via

$$i\partial_0 U(t, t_0) = e^{iH_0(t-t_0)}(H_0 - H)e^{-iH(t-t_0)}$$

(where we've canceled out an $e^{-iH_0(t-t_0)}e^{iH_0(t-t_0)}$ term, and this right-hand side is basically acting with the interaction Hamiltonian **in the interaction picture** $H_{\text{int},I}U(t, t_0)$ (where $H_{\text{int}} = H - H_0$). This differential equation is a bit tricky, but we'll rely on Dyson series for this and do perturbative calculations in λ (which we're assuming to be small). Let H_I be the interaction picture Hamiltonian for the interaction picture, so that we can make the ansatz

$$U(t, t_0) = C - i\int_{t_0}^{t} dt' H_I(t')U(t', t_0)$$

where $C = I$ because $U(t_0, t_0)$ is the identity operator. Plugging this back into itself, we find that

$$U(t, t_0) = I - i\int_{t_0}^{t} dt' H_I(t')I + (-i)^2\int_{t_0}^{t} dt' H_I(t')\int_{t_0}^{t'} dt'' H_I(t'') + O(\lambda^3)$$

and then we can keep repeating this process. But we can also make use of the fact that

$$\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2 H_I(t_1)H_I(t_2) = \int_{t_0}^{t} dt_2\int_{t_1}^{t} dt_1 H_I(t_2)H_I(t_1) = \frac{1}{2}T\left\{\int_{t_0}^{t} dt_1 dt_2 H_I(t_1)H_I(t_2)\right\},$$

where T denotes the time-ordered product (here we can imagine that to integrate along the square $[t_0, t]\times[t_0, t]$, we can integrate along the regions where $t_1 > t_2$ and $t_2 > t_1$ and add them up). So plugging this in infinitely, we find that the time-evolution operator is given by (here the $\frac{1}{2}$ goes into the series expansion for the exponential)

$$U(t, t_0) = T\left\{e^{-i\int_{t_0}^{t} dt H_I(t)}\right\}.$$

We can check that $U(t_1, t_2)U(t_2, t_3) = U(t_1, t_3)$ and that $U(t_1, t_3)U^\dagger(t_2, t_3) = U(t_1, t_2)$, so that $U(t_1, t_2)U^\dagger(t_1, t_2) = I$ (so that we have unitary time-evolution). We can rewrite U also as

$$U(t, t_0) = T\left\{e^{i\int_{t_0}^{t} dt L_I}\right\} = T\left\{e^{i\int_{t_0}^{t} d^4x\mathcal{L}_I}\right\}$$

where L_I is the interaction picture Lagrangian because $\mathcal{H}_I = -\mathcal{L}_I$ (the free stuff is gone, so the usual kinematics m^2 potential is gone). But the problem here is that we don't have a vacuum state (because we don't know how to solve the theory exactly). To find the ground state $|\Omega\rangle$, we again use the fact that λ is small, so that we should expect the new ground state and the original one to have significantly overlap. Thus we'll assume that $\langle 0|\Omega\rangle \neq 0$ and that

$E_i > E_0$ for all other i. We find that acting on the original vacuum state (and for some time T)

$$e^{-iHT}\,|0\rangle = \sum_{i=0}^{\infty} e^{-iE_iT}\,|i\rangle\,\langle i|0\rangle$$

which we can also write in terms of the new ground state as

$$= e^{-iE_0T}\,|\Omega\rangle\,\langle\Omega|0\rangle + \sum_{i=1}^{\infty} e^{-iE_iT}\,|i\rangle\,\langle i|0\rangle\,.$$

Taking the limit as T gets large, specifically setting $T \to \infty(1-i\varepsilon)$ (this imaginary part is like preparing a quantum dot in a particular state and having it interact with a noisy environment, so that we get exponential decay at long times), we only get the most stable state living the longest, namely the lowest-energy one. Thus we can throw away all contributions in this last expression to get

$$|\Omega\rangle = \lim_{T\to\infty(1-\varepsilon)} \frac{e^{iE_0T}}{\langle\Omega|0\rangle} e^{-iHT}\,|0\rangle\,.$$

Doing a transformation $T \to T + t_0$, we can now rewrite this ground state as

$$|\Omega\rangle = \lim_{T\to\infty(1-\varepsilon)} (\langle\Omega|0\rangle\, e^{-i(T+t_0)E_0})^{-1} e^{-iH(T+t_0)} e^{iH_0(T+t_0)}\,|0\rangle$$

where we're inserting a unit because H_0 is the free-field Hamiltonian and $e^{iH_0(T+t_0)}\,|0\rangle = 1\,|0\rangle$. But this means we actually have a time-evolution operator here, and we find that

$$|\Omega\rangle = \lim_{T\to\infty(1-\varepsilon)} (\langle\Omega|0\rangle\, e^{-i(T+t_0)E_0})^{-1} U(t_0,-T)\,|0\rangle$$

so that complex conjugating and replacing T with $-T$ yields

$$\langle\Omega| = \lim_{T\to\infty(1-\varepsilon)} (\langle 0|\Omega\rangle\, e^{-i(T+t_0)E_0})^{-1}\,\langle 0|\,U(T,t_0)$$

(Everything here with the vacuum states is in the Heisenberg picture.) And what we want to calculate now is correlation functions of the new ground state: we have (converting to the interaction picture and plugging in our expression for Ω)

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle = \Bigg[\lim_{T\to\infty(1-\varepsilon)} \left(e^{-iE_0(T-t_0)}\,\langle 0|\Omega\rangle\right)^{-1} \big\langle 0\big|U(T,t_0)U(t_0,x^0)\phi_I(x)U(x^0,y^0)\phi_I(y)U(y^0,t^0)U(t_0,-T)\big|0\big\rangle$$

$$\left(e^{-iE_0(t_0+T)}\,\langle\Omega|0\rangle\right)^{-1}\Bigg].$$

To make this look nicer, we'll normalize our vacuum state, which means that

$$\langle\Omega|\Omega\rangle = 1 = \lim_{T\to\infty} \left(|\,\langle 0|\Omega\rangle\,|^2 e^{-iE_0(2T)}\right)^{-1} \big\langle 0\big|U(T,x^0)U(x^0,y^0)U(y^0,t_0)U(t_0,-T)\big|0\big\rangle = \langle 0|U(T,-T)|0\rangle\,,$$

which means that

$$\frac{\langle\Omega|\phi(x)\phi(y)|\Omega\rangle}{1} = \lim_{T\to\infty(1-\varepsilon)} \frac{\big\langle 0\big|U(T,x^0)\phi_I(x)U(x^0,y^0)\phi_I(y)U(y^0,-T)\big|0\big\rangle}{\langle 0|U(T,-T)|0\rangle},$$

which is nice because the overlap now only involves time-evolution operators instead of the exponential terms. And

this can in fact further be written in terms of a time-ordered product as

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle = \lim_{T\to\infty(1-\varepsilon)} \frac{\left\langle 0\middle|T\left\{\phi_I(x)\phi_I(y)e^{i\int d^4x\mathcal{L}_I}\right\}\middle|0\right\rangle}{\left\langle 0\middle|T\left\{e^{i\int d^4x\mathcal{L}_I}\right\}\middle|0\right\rangle}.$$

But the point is that in the interaction picture, we can calculate everything with respect to the free Hamiltonian, so we can calculate the right-hand side. Similarly if we want to calculate an arbitrary correlation function, we have

$$\boxed{\langle\Omega|\phi(x_1)\cdots\phi(x_n)|\Omega\rangle = \lim_{T\to\infty(1-\varepsilon)} \frac{\left\langle 0\middle|T\left\{\phi_I(x_1)\cdots\phi_I(x_n)e^{i\int d^4x\mathcal{L}_I}\right\}\middle|0\right\rangle}{\left\langle 0\middle|T\left\{e^{i\int d^4x\mathcal{L}_I}\right\}\middle|0\right\rangle}}.$$

We'll see next time how to calculate this and write something more concrete down.

7 October 18, 2022

Last lecture, we analyzed interacting theories and discussed how to calculate correlation functions in the interacting Hamiltonian vacuum state $|\Omega\rangle$. Basically, we found that if we assume the overlap $\langle 0|\Omega\rangle$ between the free vacuum and interacting vacuum ground states, then we could write the correlation function

$$\langle\Omega|T\{\phi(x_1)\cdots\phi(x_n)\}|\Omega\rangle = \lim_{T\to\infty(1-\varepsilon)} \frac{\left\langle 0\middle|T\left\{\phi_I(x_1)\cdots\phi_I(x_n)e^{i\int d^4x\mathcal{L}_I}\right\}\middle|0\right\rangle}{\left\langle 0\middle|T\left\{e^{i\int d^4x\mathcal{L}_I}\right\}\middle|0\right\rangle}$$

(where we may add a time-ordering on the left compared to what we derived last time because the expression is independent of the ordering of the x_is). The point is that we can now calculate in terms of the free Hamiltonian, and we're going to see today how to make use of Feynman diagrams to avoid doing too many calculations, starting with **Wick's theorem**. We need a way to evaluate something like $\langle 0|T\{\phi(x_1)\cdots\phi(x_n)\}|0\rangle$, and we should recall that we defined

$$\phi = \int \frac{d^3p}{(2\pi)}\frac{1}{\sqrt{2\omega_p}}(ae^{-ipx} + a^\dagger e^{ipx}),$$

which we can decompose into two terms ϕ_- and ϕ_+, corresponding to the two exponentials. Then we can define the fields $A = \phi(x) = A^+ + A^-$ and $B = \phi(y) = B^+ + B^-$, and the time-ordered product, if we assume $x^0 > y^0$, is

$$T\{AB\}|_{x_0>y_0} = AB = A^+B^+ + A^+B^- + A^-B^+ + A^-B^-.$$

Having this operator act on the vacuum state means many of these terms will drop out, and this is almost a normal operator except for the A^-B^+ term. But if we introduce the normal ordering (recall that this is where creation operators go to the left and annihilation operators to the right) $:B^-A^+: = A^+B^-$ and observe that $A^-B^+ = [A^-, B^+] + B^+A^-$, then we will find that

$$T\{AB\}|_{x_0>y_0} = :AB: + [A^-, B^+]|_{x_0>y_0}.$$

In the opposite ordering we find that

$$T\{AB\}|_{y_0>x_0} = :AB: + [A^+, B^-]|_{x_0>y_0}.$$

But now because we can directly

$$\begin{aligned}\langle 0|[A,B]|0\rangle &= \left\langle 0\middle|(A^+ + A^-)(B^+ + B^-) - (B^+ + B^-)(A^+ + A^-)\middle|0\right\rangle \\ &= \left\langle 0\middle|A^-B^+ - B^-A^+\middle|0\right\rangle,\end{aligned}$$

we can put things together to find that $\boxed{T\{AB\} =: AB : +D_F(x-y)}$, where D_F is the Feynman propagator, so we get a relationship between the time-ordered product and the normal ordered product. (Remember that we like normal operators because they have expectation zero in the vacuum.) More generally, we get following identity:

Theorem 16 (Wick)
For operators $A = \phi(x_A), B = \phi(x_B), \cdots, Z = \phi(x_Z)$ (or generically any number of operators),

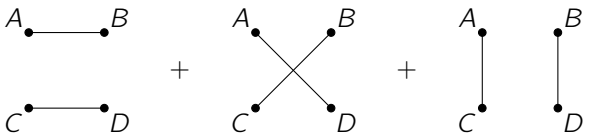
$$T\{ABC\cdots XYZ\} =: ABC\cdots XYZ : + D_F(x_A - x_B) : C\cdots XYZ : + D_F(x_A - x_C) : BD\cdots XYZ : + \cdots$$

$$+ D_F(x_A - x_B)D_F(x_C - x_D) : EF\cdots XYZ : + \cdots + D_F(x_A - x_B)\cdots D_F(x_Y - x_Z) + \cdots .$$

In other words, we contract pairs of variables in all possible ways into Feynman propagators and sum over everything, and we don't necessarily have to contract adjacent variables – all possibilities allowed by the combinatorics contribute. And now notice that if we take the expectation of this operator in the vacuum state, everything drops out except when all of these are Feynman propagators (since normal operators annihilate the vacuum) – for example, that means we just have

$$\langle 0|T\{ABCD\}|0\rangle = D_F(x_A - x_B)D_F(x_C - x_D) + D_F(x_A - x_C)D_F(x_B - x_D) + D_F(x_A - x_D)D_F(x_B - x_C)$$

for any operators $A = \phi(x_A), B = \phi(x_B), C = \phi(x_C), D = \phi(x_D)$ (and note here that the Feynman propagator is symmetric). Graphically, we can imagine drawing out the possibilities as shown below:



Remark 17. *Notice that if we have an odd number of operators, these expectations will always be zero, because we will always have a leftover normal operator which will annihilate the vacuum state.*

We can now return to the correlation function and calculate expectations, and we'll do this with perturbation theory. Recall that our interaction Lagrangian $\mathcal{L}_I = -\frac{\lambda}{4!}\phi^4(x)$ is small in λ, we want to calculate the denominator by expanding out the exponential

$$\left\langle 0\middle|T\left\{e^{-i\frac{\lambda}{4!}\int d^4x\phi^4}\right\}\middle|0\right\rangle = 1 - \frac{i\lambda}{4!}T\left\{\left\langle 0\middle|\int d^4x\phi^4(x)\middle|0\right\rangle\right\}$$

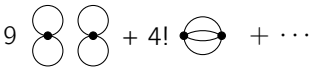
$$+\frac{-(i\lambda)^2}{2(4!)^2}T\left\{\int d^4x\int d^4y\,\langle 0|\phi^4(x)\phi^4(y)|0\rangle\right\} + O(\lambda^3).$$

Remark 18. *We're now going to assume that taking integrals and computing expectations can be interchanged, so we can move the integrals outside the time-ordering.*

Applying Wick's theorem to $\phi\phi\phi\phi$, we see that $\langle 0|T\{\phi^4(x)|0\rangle = 3D_F(x-x)D_F(x-x)$ (we can imagine three times the diagram  in which we contract the point x to itself twice). The next term is then

$$\langle 0|T\{\phi^4(x)\phi^4(y)\}|0\rangle = \langle 0|T\{\phi_x\phi_x\phi_x\phi_x\phi_y\phi_y\phi_y\phi_y|0\rangle\,,$$

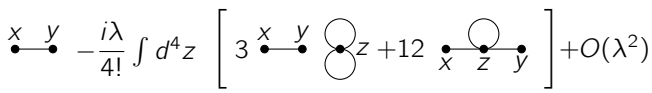
and doing the combinatorics turns out to give us the following picture (where the dots represent x and y), where it depends on how many times we contract xs to ys versus within each variable:



(We can work out the details on our own.) On the other hand, the numerator can also be expanded out in powers of λ. The $O(\lambda^0)$ term is $\langle 0|T\{\phi(x)\phi(y)\}|0\rangle = D_F(x-y)$, the $O(\lambda^1)$ term is

$$\left\langle 0\middle|T\left\{\phi(x)\phi(y)\left(-\frac{i\lambda}{4!}\right)\int d^4z\phi^4(x)\right\}\middle|0\right\rangle,$$

and we can apply Wick's theorem again. This time the combinatorics works out to 3 times $D_F(x-y)D_F(z-z)D_F(z-z)$ and 12 times $D_F(x-z)D_F(y-z)D_F(z-z)$, and then integrating that over d^4z and multiplying by $-\frac{i\lambda}{4!}$. So putting everything together, our numerator looks as shown below:



Thus we've expanded both the numerator and denominator in small λs, so we can do a geometric series expansion and keep the leading order term. Putting everything together, we find that the correlation function we wanted is (here the 12 cancels out mostly with the 4! in the denominator)

$$\boxed{\langle\Omega|T\{\phi_H(x)\phi_H(y)\}|\Omega\rangle = D_F(x-y) - \frac{i\lambda}{2}\int d^4z D_F(x-z)D_F(y-z)D_F(z-z) + O(\lambda^2)}.$$

(It turns out that the **disconnected diagrams** (or **disconnected graphs**) in which z is not connected to the endpoints x and y are going to be problematic, because we cannot integrate $D_F^2(0)$ over d^4z. So consider the set of all **disconnected contributions** and call them V, enumerating the different graphs as $V_1, V_2, \cdots$. (For example, V_1 could be the diagram  .) It will turn out that in general we always get total contributions of the form

$$(\text{connected part})\prod_i V_i^{n_i}\frac{1}{n_i!},$$

so adding up all such contributions

$$\sum_{\text{connected graph}}\sum_{\{n_i\}}(\text{connected contribution})\prod_i V_i^{n_i}\frac{1}{n_i!} = \sum_{\text{connected graph}}(\text{connected contribution})\prod_i\left(\sum_{n_i}V_i^{n_i}\frac{1}{n_i!}\right),$$

which simplifies to

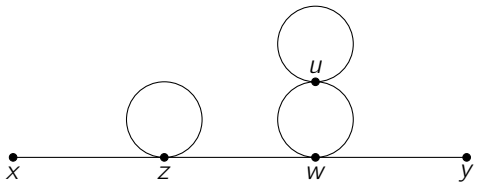
$$= \sum_{\text{connected}}(\text{connected contribution})e^{\sum_i V_i}.$$

So if we apply this to our time-ordered product and want to evaluate something like the numerator

$$\left\langle 0\middle|T\{\phi_I(x)\phi_I(y)e^{i\int d^4x\mathcal{L}_I}\middle|0\right\rangle,$$

then the expectation will be the sum of contributions over all connected diagrams (the two graphs we saw for $D_F(x-y)$ and $D_F(x-z)D_F(y-z)D_F(z-z)$, as well as other terms), times the exponential of the sum of the disconnected diagrams like $D_F(z-z)$, $D_F(z-z)D_F(z-z)$, and so on. And the denominator is particularly simple because there are no connected diagrams, so we just have the exponential of the sum of all disconnected diagrams. But that means that when we divide, the exponential "cancels out," and thus **if we wanted to calculate something like** $\langle\Omega|T\{\phi_H(x_1)\cdots\phi_H(x_n)\}|\Omega\rangle$, **we just sum over all connected diagrams with** n **external points** $x_1,\cdots,x_n$, with as many zs as we want but with each z coming with a power of λ (and for a theory of the form $\frac{i\lambda}{4!}\int d^4z\phi^4(z)$ we may

have four lines to connect from z). So we can now just draw diagrams like the one below and get a λ^3 term, and we want to understand what the prefactor is corresponding to it. For example, consider the following diagram:



This term corresponds to trying to calculate one of the possible types of contractions in the term

$$\int d^4z d^4u d^4w \frac{1}{3!}\left(-\frac{i\lambda}{4!}\right)^3 \langle 0|\phi(x)\phi(y)\phi(z)^4\phi(w)^4\phi(u)^4|,\rangle$$

in which we need to figure out the number of ways to contract terms to get all of the edges that we drew above. But we can just calculate the combinatorics from our diagram – for z we get $4\cdot 3$, for w we get $4!$, for u we get $\frac{4!}{4}$, and then we can exchange z, u and w and get the same shape diagram up to relabeling so we get an overall factor of $3!$, and multiplying everything together yields $\frac{3!(4!)^3}{8}$, which mostly nicely cancels out with the prefactor before the integral. So we see that the overall prefactor for this Feynman diagram is $\frac{(-i\lambda)^3}{8}$. But we can make this process more systematic, deriving easy rules for **symmetry factors** (which are different for every interacting theory):

- Every propagator connecting a vertex to itself gives a $\frac{1}{2}$ factor.
- Every n propagators that connect the same two vertices give us a factor of $\frac{1}{n!}$.
- Exchanging vertices without changing the diagram yields a factor of $\frac{1}{n!}$.

(And there are programs like QGRAF that help us do these kinds of calculations as well, giving correct symmetry factors and statistics.)

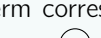
8 October 20, 2022

We'll start by summarizing what we've been doing in the last few lectures: we're working with an interacting system with a Lagrangian

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

for some small λ. We can write the expectation of the time-ordered operator $T\{\phi_I(x)\phi_I(y)\}$ in the original vacuum $|0\rangle$ as an integral which yields $D_F(x-y)$, and we can then use Dyson's formula for the unitary time-evolution $U(t,t_0)$ to get a formula for the two-point (and in fact n-point) correlation function in the new vacuum state $|\Omega\rangle$ in terms of correlation functions in the original vacuum. We then used Wick's theorem to write $T\{\phi_I(x_1)\cdots\phi_I(x_n)\}$ in terms of normal orderings and contractions, and it turns out we end up getting a sum over all connected Feynman diagrams (since the expectation of any normal operator is zero in $|0\rangle$, we must contract all coordinates in pairs).

Example 19

If we want to calculate the two-point correlation function $\langle\Omega|T\{\phi_H(x)\phi_H(y)\}|\Omega\rangle$ in our scalar ϕ^4 theory, we have a zeroth order term corresponding to the diagram  , and then at order λ we get $\frac{1}{2}$ (a symmetry factor) times the diagram  . Next, at order λ^2 we get $\frac{1}{4}$ times the diagram  , $\frac{1}{6}$ times the diagram  (there's a distinction between "bubble diagrams" and "tadpole diagrams" based on how these kinds of graphs look), and $\frac{1}{4}$ times the diagram  .

Remark 20. *Physically, we should be thinking of z as a "virtual particle" at which the field may oscillate but which we cannot be observed. And later, we'll see that this has implications (telling us about resolutions and energy scales).*

Recall that to get to a mathematical expression from something like this, we correspond the line segment between two points with the Feynman propagator between those points, associate to each other point z besides x and y the integral $-i\lambda\int d^4z$, and then account for symmetry factors. So the diagram $\frac{1}{2}$ times  really means we have an integral of the form

$$-\frac{\lambda}{2}\int d^4z D_F(x-y)D_F(z-z)D_F(z-y)$$
$$=-\frac{i\lambda}{2}\int d^4z\frac{d^4p}{(2\pi)^4}e^{-ip(x-z)}\frac{i}{p^2-m^2+i0}\int\frac{d^4q}{(2\pi)^4}e^{-iq(y-z)}\frac{i}{q^2-m^2+i0}\int\frac{d^4k}{(2\pi)^4}\frac{ie^{-ik(z-z)}}{k^2-m^2+i0},$$

where $+i0$ is basically the same as $+i\varepsilon$ and all of the integrals can be written down in any order. We can collect a few z-terms, writing that $\int d^4z e^{i(p+q)z}=(2\pi)^4\delta^{(4)}(q+p)$, to simplify the above expression to

$$-\frac{\lambda}{2}\int\frac{d^4p}{(2\pi)^4}e^{-ipx}\int\frac{d^4q}{(2\pi)^4}e^{-iqy}(2\pi)^4\delta^{(4)}(p+q)\int\frac{d^4k}{(2\pi)^4}\frac{1}{k^2-m^2+i0}\frac{1}{p^2-m^2+i0}\frac{1}{q^2-m^2+i0}.$$

And now this is in a form that motivates Fourier transforming: if we switch out by integrating $\int d^4x e^{ip'x}\int d^4y e^{iq'y}$, we see that in momentum space this integral becomes, after carrying out the d^4p and d^4q integrals,

$$-\frac{\lambda}{2}\int\frac{d^4k}{(2\pi)^4}\frac{(2\pi)^4\delta^{(4)}(p'+q')}{(k^2-m^2+i0)(p'^2-m^2+i0)(q'^2-m^2+i0}$$

We can then read off the Feynman rules in momentum space from this expression: since the propagator in momentum space looks like

$$D_F=\int\frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}\frac{i}{p^2-m^2+i0},$$

we see that a line segment with momentum p contributes a propagator term $\frac{i}{p^2-m^2+i0}$ to our expression, any additional point included in our diagram again contributing $-i\lambda$, an outward connection (external line coming into our diagram) corresponding to an e^{-ipx} term, and we get a loop integration $\int\frac{d^4p}{(2\pi)^4}$ over all unconstrained momenta. (And we keep the same symmetry factors as before.) For example, with just the vacuum diagram with incoming momenta p and q from the two sides, we don't have any unconstrained momenta so we don't have any integrals, just $e^{-ipx}e^{-iqy}\frac{i}{p^2-m^2+i0}(2\pi)^4\delta^{(4)}(p+q)$ (with the propagator term only in one direction, and the delta term corresponding to momentum conservation).

But now we can specify what we want to calculate, which is the **S-matrix** (or **scattering matrix**) – the probability

transition from the incoming state to the outgoing state. Specifically, we consider the initial and final states

$$|i\rangle = \lim_{t\to-\infty} |\psi(t)\rangle\,, \quad |f\rangle = \lim_{t\to\infty} |\phi(t)\rangle$$

(we can imagine ψ as being some set of particles, and ϕ as some potentially different set of particles), and we define

$$S_{fi} = \langle f|S|i\rangle = \lim_{T\to\infty} \langle f|U(T,-T)|i\rangle\,,$$

which is the probability amplitude for a transition $|i\rangle \to |f\rangle$. Here, remember that U is a time-ordered exponential which can be written as

$$U(\infty,-\infty) = \sum_{n=0}^{\infty} \frac{1}{n!}(-i)^n \int_{-\infty}^{\infty} dt_1 \cdots \int_{-\infty}^{\infty} dt_n T\{H_I(t_1)\cdots H_I(t_n)\},$$

where because U is unitary we aren't gaining or losing probability. We can write the scattering matrix element as

$$S_{fi} = \delta_{fi} + i(2\pi)^4\delta^{(4)}(p_f - p_i)T_{fi} = \mathbf{1} + iT,$$

where we often write the δ function as $\mathbf{1}$ (it's in some sense a "unit") and T is called the **transfer matrix**. Then it turns out that (we'll prove this later) our computations are made easier by the fact that

$$S_{fi} = \langle f|S|i\rangle = {}_0\langle f|S|i\rangle|_0\,,$$

but where on the right-hand side we only sum over connected diagrams that are "amputated" (see below):

Example 21

For $2\to 2$ scattering, in which we have $p_1 + p_2 \to p_3 + p_4$, we have the matrix element

$$S_{fi} = \sqrt{2E_1}\sqrt{2E_2}\sqrt{2E_3}\sqrt{2E_4}\,\langle 0|a(p_3)a(p_4)Sa^\dagger(p_1)a^\dagger(p_2)|0\rangle\,.$$

Here the $\sqrt{2E}$ normalization gives us Lorentz invariance, and the idea is that we prepare momentum eigenstates for p_1 and p_2 and also for p_3 and p_4 and see if they are related using the S-matrix. And we modify our Feynman rules slightly, saying that we won't connect propagators coming from the outside in an amputated diagram, so that instead of a Feynamn propagator term of the form e^{-ipx} we do not have any propagator from p_1 and p_2 in our scattering. And now the probability density of a process $i\to f$ occurring is $P_{i\to f} = |\langle f|S|i\rangle|^2$.

But to get from these plane waves of particular momenta to an actual situation where we have particles colliding, we'll prepare wavepackets which we want to scatter into each other. Specifically, our initial state can be

$$|f_1, f_2; i\rangle = \int \frac{d^3p_1}{(2\pi)^3}\frac{d^3p_2}{(2\pi)^3}\frac{1}{2\omega_{p_1}\omega_{p_2}} f_1(p_1)f_2(p_2)\,|p_1,p_2\rangle\,,$$

where we can write down (for example) a "molded Gaussian wavepacket" with momentum-space shape given by the function $\frac{f(k)}{2\omega_k} = \frac{1}{(\pi\Delta p)^{3/2}}e^{-i(\vec{p}-\vec{k})^2/\Delta p^2}$ with $\Delta p \ll |\vec{p}|$, and we normalize so that $\int \frac{d^3k}{2\omega_k}|f_k|^2 = 1$. Then the probability amplitude looks like (since we're integrating over the square of the S-matrix element, we have the matrix element times its conjugate)

$$P_{i\to f} = |\langle f|S|f_1,f_2;i\rangle|^2 = \int \frac{d^3p_1}{(2\pi)^3}\frac{d^3p_2}{(2\pi)^3}\frac{d^3p_1'}{(2\pi)^3}\frac{d^3p_2'}{(2\pi)^3}\frac{1}{(2\omega_{p_1})(2\omega_{p_2})(2\omega'_{p_1})(2\omega'_{p_2})} f_1(p_1)f_2(p_2)f_1^*(p_1')f_2^*(p_2')$$
$$(2\pi)^4\delta^{(4)}(p_1+p_2-p_f)(2\pi)^4\delta^{(4)}(p_1'+p_2'-p_f)\,\langle p_1',p_2'|T|f\rangle\,\langle f|T|p_1,p_2\rangle$$

(where the $*$ denotes complex conjugate).We can then convert by Fourier transform again, writing $2\pi\delta^{(4)}(p_1' +$

$p_2' - p_1 - p_2$) as $\int d^4x e^{i(p_1+p_2-p_1'-p_2')x}$ and for each momentum variable we can convert with an integral $f_i(x) = \int \frac{d^3p_i}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{p_i}}} f(p_i)e^{-ip_i x}$, where (here is a physics statement) we are saying we can write $p_i = \overline{p}_i + O(\Delta p_i)$ because we have narrow wavepackets. (The point is that the wavepacket will peak at a particular momentum, so the contributions with something like $\langle f|T|p_1p_2\rangle$ will only come primarily from $\langle f|T|\overline{p}_1\overline{p_2}\rangle$.) We thus find that

$$P_{f\to i} = \int d^4x \frac{|f_1(x)|^2|f_2(x)|^2}{2\omega_{\overline{p}_1}2\omega_{\overline{p}_2}}(2\pi)^4\delta^{(4)}(p_f - \overline{p}_1 - \overline{p}_2)|\langle f|T|\overline{p}_1\overline{p}_2\rangle|^2 + O(\Delta p_i).$$

(And we're only looking at a particular space-time coordinate x, corresponding to the local interactions of the scattering process, so we only need one space coordinate. But then we integrate over all of space because interaction can happen anywhere.) Notice that we've switched from S to T during this process, and that's saying that we don't care about the "unit $\mathbf{1}$" contributions to the scattering amplitudes because that is the case where nothing happens. So we can write this final expression we get as

$$P_{f\to i} = \int d^4x \frac{dP_{i\to f}}{d^4x}$$

and think of the integrand as a scattering probability density. We'll see next time how to relate this probability density to something that we can actually measure!

9 October 25, 2022

Last time, we defined the S-matrix for scattering, which is one of the main tools for doing calculations and understanding observables. Specifically, we found that if our initial state is $|i\rangle = \lim_{T\to-\infty}|\psi(T)\rangle$ and our final state is $|f\rangle = \lim_{T\to\infty}|\psi(T)\rangle$, then the matrix element we care about is

$$S_{fi} = \langle f|S|i\rangle = \langle f|U(\infty,-\infty)|i\rangle$$

which we claimed is just ${}_0\langle f|U(\infty,-\infty)|i\rangle_0$ (the expectation when we look only at the connected and amputated diagrams.

Example 22

Consider four-particle ($2\to 2$) scattering, in which particles of momenta p_1, p_2 become particles of momenta p_3, p_4.

We can write down the matrix element in terms of our momentum eigenstates as

$$S_{fi} = \prod_{i=1}^{4}\sqrt{2E_i}\left\langle 0\middle|a(p_3)a(p_4)Sa^\dagger(p_1)a^\dagger(p_2)\middle|0\right\rangle,$$

where if we have a small coupling constant in our interacting Lagrangian term $-\frac{1}{4!}\lambda\phi^4$, we can do perturbation theory

$$S = U(\infty,-\infty) + T\left\{e^{i\int d^4\mathcal{L}_I}\right\} = \mathbf{1} + \int d^4x\mathcal{L}_I + \cdots$$

and find that this simplifies at zeroth order to

$$S_{fi}|_{O(\lambda^0)} = \prod_{i=1}^{4}\sqrt{2E_i}\left\langle 0\middle|a_1a_2a_3^\dagger a_4^\dagger\middle|0\right\rangle = (2\pi)^6(2E_1)2(E_2)\left[\delta^{(3)}(p_1-p_3)\delta^{(3)}(p_2-p_4) + \delta^{(3)}(p_1-p_4)\delta^{(3)}(p_2-p_3)\right],$$

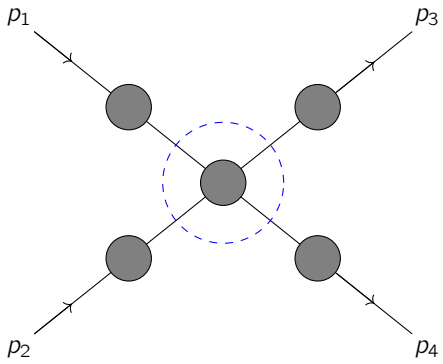
where the idea is that a_1 can either pair up with a_3 or a_4 and a_2 pairs up with the other one, and here because $E_i = \sqrt{\vec{p}_i^2 + m^2}$ conservation of momentum also means conservation of energy. But we care much more about the

terms where we do see something happening (scattering), so we're not going to care about that leading order term and now want to study the next order. We have

$$S_{fi}|_{O(\lambda^1)} = \prod_{i=1}^{4} \sqrt{2E_i} \left\langle 0 \middle| a_3 a_4 \int d^4x \left(-\frac{i\lambda}{4!}\right) \phi_x \phi_x \phi_x \phi_x a_1^\dagger a_2^\dagger \middle| 0 \right\rangle .$$

Again we need to pair up our creation and annihilation operators until we get to normal ordering. We have a few calculations that we need to do for this:

- Letting $a_3 = a(p_3)$, we have $[a_3, \phi_x] = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left(e^{-ipx}[a_3, a_p] + e^{ipx}[a_3, a_p^\dagger] \right)$, where the commutators are 0 and $(2\pi)^3 \delta^{(3)}(\vec{p_3} - \vec{p})$, so doing the integral gives us $\frac{e^{ip_3 x}}{\sqrt{2E_3}}$.
- There are other Feynman diagrams we can draw too – basically, we should imagine that we have $1, 2, 3, 4$ connected with two lines (either $1- > 3, 2- > 4$ or $1- > 4$, $2- > 3$), plus either an additional loop attached to one of those lines (each of those has $\frac{1}{2}$ symmetry factor), or an additional disconnected loop (each of those has a $\frac{1}{4}$ symmetry factor), or the two lines can intersect. But for the S-matrix element that we care about, we only want the connected, amputated diagrams, which can be described as shown here:



  And the point is that we only care about the contributions within the blue part when we say "amputated." (For more, we should look up LSZ reduction, which we'll come back to in a few weeks.) Here, remember that "connected" means "connected to one of p_1, p_2, p_3, p_4."

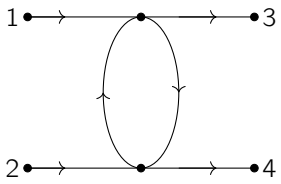
So it turns out that contributions to S_{fi} of order λ only come from the case  , where the center point represents the spacetime point x (and then the contributions of order λ are where we have two external points instead of just one, and so on). This diagram contributes

$$(-i\lambda) \prod_{i=1}^{4} \sqrt{2E_i} \int d^4x \frac{e^{ix(p_3+p_4-p_1-p_2}}{\sqrt{16E_1E_2E_4E_4}} = -i\lambda(2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4),$$

so that we really have

$$S_{fi} = \mathbf{1} + i(2\pi)^4 \delta^{(4)}(p_i - p_f) T_{fi},$$

where T_{fi} is just $-\lambda$. And we could have used **Feynman rules** for matrix elements to get here faster: internal lines give $\frac{i}{k^2 - m^2 + i0}$, external lines give us 1, every external vertex with four lines coming out of it gives us a $-i\lambda$, we get an $\int \frac{d^4p}{(2\pi)^4}$ for every unconstained momentum, but we must have momentum conservation at every vertex, and we need to think about symmetry factors. So imagine now that we have the diagram as shown below (where $1, 2, 3, 4$ just mean p_1, p_2, p_3, p_4, and there is some momentum k going up and momentum k' going down in the loop):



Then the Feynman rules give us $p_4 = p_1 + p_2 - p_3$ and $k' = k + p_2 - p_4 = k + p_1 - p_3$, so the contribution to S_{fi} will be (the complex number)

$$\frac{(-i\lambda)^2}{2}\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i0}\frac{i}{(k+p_2-p_4)^2-m^2+i0}.$$

And remember that our Feynman rules need to look different for different interaction Lagrangians: if we had something like $-\frac{\lambda}{3!}\phi^3$, then each external vertex would need to have three outgoing vertices instead of four. The point is that we're abstracting away calculations just by reading off what $\mathcal{L}_I$ looks like. (And we can take a look at the FeynRules package for assistance too – this is meant to be systematic.)

Last time, we went a bit further and looked at probabilities $P_{i\to f} = |\langle f|S|i\rangle|^2$ of actually going through this scattering process. (This is a real physics question we can ask – we want to find the probability that particles of some given momentum will scatter and give the desired output momentum.) If we have an incoming wavepacket

$$|f_1, f_2, \text{in}\rangle = \int \frac{d^3p_1}{(2\pi)^3}\frac{1}{2E_{p_1}}\frac{d^3p_2}{(2\pi)^3}\frac{1}{2E_{p_2}} f_1(p_1)f_2(p_2)\,|p_1,p_2\rangle\,,$$

where for example $f(p)$ could basically be a peaked Gaussian with some width Δp, then doing some algebra and plugging into our scattering matrix shows that the probability of scattering is given by

$$\int d^4x \frac{|f_1(x)|^2}{2\omega_{p_1}}\frac{|f_2(x)|^2}{2\omega_{p_2}}(2\pi)^4\delta^{(4)}(p_f-p_1-p_2)\,|\langle f|T|p_1,p_2\rangle|^2 + O(\Delta p/p),$$

and taking this integrand we can think of it as a probability **density** (per spacetime point / time slice) of scattering an initial state into a final state.

Example 23

For example (to draw an analog), if we have two "screens" of particles approaching each other of velocities v_1, v_2, this density looks like $\frac{dP}{d^4x} = |v_1 - v_2|\rho_1\rho_2\sigma$, where $v_1 - v_2$ tells us about the rate of potential collision and ρ_1, ρ_2 the density of particles under consideration, so that the scattering cross section σ is

$$\sigma = \frac{1}{|v_1-v_2|\rho_1\rho_2}\frac{|f_1|^2}{2\omega_{p_1}}\frac{|f_2|^2}{2\omega_{p_2}}(2\pi)^4\delta^{(4)}(p_f-p_i)\,|\langle f|T|p_1,p_2\rangle|^2$$

where f_1 is the probability distribution of finding a particle at a particular spacetime point, so $|f_1|^2 = \rho_1$.

We can then write that

$$E_1E_2|v_1-v_2| = E_1E_2\left|\frac{\vec{p}_1}{E_1}+\frac{\vec{p_2}}{E_2}\right| = |\vec{p}_1||E_1+E_2| = \sqrt{2(p_1p_2)^2 - m_1^2m_2^2}$$

(this is just kinematics and using that $\vec{p}_1 = -\vec{p}_2$ in the center of mass), so that our cross-section looks like

$$\sigma = \frac{(2\pi)^4\delta^{(4)}(p_1+p_2-p_f)}{\sqrt{(2p_1p_2)^2-m_1^2m_2^2}}\,|\langle f|T|p_1,p_2\rangle|^2\,.$$

So this number basically gives a sense of how difficult it is to hit a scattering event. For physical situations like at the

LHC, we might see that the number of events is given by $N = L\sigma$, where σ can be calculated from first principles and L is the **luminosity** (an experimental parameter) given by $\frac{N_a N_b f}{(4\pi)A_{\text{beam}}}$, where A_{beam} is the cross-sectional area of the intersection beam of particles and f is the frequency.

But the point now is that even when we have an intersection event, many different things could happen. From a probabilistic point of view, what we can write down is that we have

$$\sigma = \frac{1}{2\sqrt{(2p_1p_2)^2 - m_1^2m_2^2}} \sum_{\text{states } s} \int \prod_{i=1}^{N_s} \frac{d^4q_i}{(2\pi)^4} 2\pi\delta_+(q_i^2 - m_i^2)(2\pi)^4\delta^{(4)}\left(p_1 + p_2 - \sum_{j=1}^{N_s} q_j\right) |\langle q_1, \cdots, q_{N_s}|T|p_1, p_2\rangle|^2$$

where N_s is the number of particles in our final state s (so we just integrate over all possible momenta that the outgoing particles have, and we look at the cross-section that would be produced from such a scattering), though we need to double-count if we have identical outgoing particles and add a Bose factor $\frac{1}{n!}$ whenever we have n-fold identical particles. (This is called an **inclusive cross section**.) But it's interesting for us now to think about differential cross sections – if we want to pick out a particular value of the observable of a particular particle and find $\delta(X - \hat{X}(\{q_i\}, p_1, p_2)$ (where $\hat{X}$ might be a particular energy of the 15th particle or something), we need to consider the equation

$$\frac{d\sigma}{dX} = \frac{1}{2\sqrt{(2p_1p_2)^2 - m_1m_2^2}} \sum_{\text{states } s} \int \prod_{i=1}^{N_s} \frac{d^4q_i}{(2\pi)^4}\delta_+(q_i^2 - m_i^2)(2\pi)^4$$

$$\cdot\delta^{(4)}\left(p_1 + p_2 - \sum_{i=1}^{N_s} q_i\right) |M_{p_1+p_2\to q_1+\cdots+q_{N_s}}|^2\delta(X - \hat{X}(\{q_i\}, p_1, p_2).$$

And if we have just a single particle and want to think about its decay rate, then the expression for probability that we want to write down is

$$P_{p_1\to f} = \int d^4x \frac{|f_1(x)|^2}{2\omega_{p_1}}(2\pi)^4\delta^{(4)}(p_1 - p_f)|\langle p_i|T|p_f\rangle|^2$$

where $\int d^4x|f_i(x)|^2 = 1$. Choosing a frame so that the particle is at rest and $\omega_p = m$, we find that

$$\Gamma = \frac{1}{2m} \sum_{\text{final states } s} \prod_{i=1}^{N_s} \frac{d^4q_i}{(2\pi)^4}\delta_+(q_i^2 - m_i^2)(2\pi)^4 \cdot 2\pi\delta^{(4)}\left(p_1 - \sum_{q=1}^{N_s} q_i\right) |\langle q_1, \cdots, q_{N_s}|T|p_1\rangle|^2.$$

And in particular, we get the expected lifetime of the particle by calculating $\tau = \frac{1}{\Gamma}$ if we know the decay rules for different kinds of particles.

10 October 27, 2022

We've been dealing with scalar quantum field theories up until now (and gotten to a point where we've related our formalism to something that we can measure in a scattering experiment), but we'll now change topics dramatically – we'll be discussing **fermions** today, starting with the Dirac equation. Dirac did have something to work with – he knew that the Schrodinger equation $i\partial_0\psi = -\frac{\Delta}{2m}\psi + V(\vec{x})\psi$ was non-relativistic but definitely useful (for studying systems like the hydrogen atom), and so to get a relativistic equation we should try to have an equal number of time- and space-derivatives. We know that in the Schrodinger equation we basically have $E = \frac{p^2}{2m} + V$, so it makes sense to try to work with a linearized version of Einstein's equation. But we see that

$$E^2 - p^2 - m^2 = 0 \implies E = \sqrt{\vec{p} + m^2} = m^2 - \frac{1}{2m}\vec{p}^2 - \frac{1}{8m^2}(\vec{p}^2)^2,$$

and the right-hand side has various numbers of derivatives so it's not going to work as we want – instead, we'll take the square root of our Klein-Gordon equation and write

$$-(i\partial_\mu - m)(i\partial_\mu + m) = \partial_\mu^2 + m^2,$$

which doesn't quite make any sense because the indices don't line up, but the point is to use this to motivate an ansatz for the first term. We're going to try to construct an equation of the form

$$(i\gamma^0\partial_0 + i\gamma^i\partial_i + m)\psi = 0 \implies (E\gamma_0 + \gamma^i p_i + m)\psi = 0.$$

So to figure out what the γs should be, we multiply by the inverse of γ_0 to get

$$\boxed{(i\partial_0 + i\gamma_0^{-1}\gamma^i\partial_0 + m\gamma_0^{-1})\psi = 0},$$

and then multiplying by $i\partial_0$ (taking another time-derivative) yields

$$-\partial_0^2\psi = (-i\gamma_0^{-1}\gamma^i\partial_i + m\gamma_0^{-1})i\partial_0\psi = (-\gamma_0^{-1}\gamma^i\gamma_0^{-1}\gamma^j\partial_i\partial_j + im\gamma_0^{-1}\gamma_0^{-1}\gamma^i\partial_i + im\gamma_0^{-1}\gamma^i\gamma_0^{-1}\partial_i - m^2(\gamma_0^{-1})^2)\psi.$$

where in the last step we've substituted in the expression for $i\partial_0\psi$ from the boxed equation. So now matching terms back in (to try to get Klein-Gordon), the blue part tells us we must have $(\gamma_0^{-1})^2 = I$, meaning that $\gamma_0 = \gamma_0^{-1}$, and the red part must go away because it's not in the Klein-Gordon equation, so we must have $\{\gamma_0^{-1}, \gamma_i\} = 0 \implies \{\gamma_0, \gamma_i\} = 0$ (where $\{A, B\} = AB + BA$ is the anticommutator). For the green part, we can first rewrite as $-\gamma^0\gamma^i\gamma^0\gamma^j\partial_i\partial_j = \gamma^)\gamma^0\gamma^i\gamma^j\partial_i\partial_j = \gamma^i\gamma^j\partial_i\partial_j$. Since we have two indices (and we can interchange them), this last expression is the same as $\frac{1}{2}(\gamma^i\gamma^j + \gamma^j\gamma^i)\partial_i\partial_j$, and now we want this to be equal to $-\partial_i^2$ to match Klein-Gordon. Thus $\{\gamma^i, \gamma^j\} = -2\delta_{ij}$.

We can now collect all of these different relations together, known as the **Clifford algebra**,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}.$$

We claim that these γ^μs are matrices, and we can find out a little more about them: we have the trace

$$\mathrm{Tr}[\gamma^i] = \mathrm{Tr}[\gamma^0\gamma^0\gamma^i] = \mathrm{Tr}[-\gamma^0\gamma^i\gamma^0] = \mathrm{Tr}[-\gamma^i\gamma^0\gamma^0] = -\mathrm{Tr}[\gamma^i]$$

where we've used our relations in the first two steps and the cyclicity of the trace in the last. Thus γ^i is traceless, and similarly we can check that γ^0 is traceless. Furthermore, because $\gamma^0\gamma^0 = I$ all eigenvalues of γ^0 must be ± 1, and because $\{\gamma^i, \gamma^j\} = -2\delta_{ij}$ we must have $(\gamma^i)^2 = -1$ and thus all eigenvalues of γ_i must be $\pm i$. Finally, we want our Hamiltonian H (whatever is on the right-hand side of our equation) to be a Hermitian operator. Thus we must have

$$-i\gamma^0\gamma^i\partial_j + m\gamma^0 = i\gamma_i^\dagger\gamma_0^\dagger\partial_i + m\gamma_0^\dagger,$$

and all of the eigenvalues of γ_0 are real so $\gamma_0^\dagger = \gamma_0$ (this is like saying the mass term shouldn't change). On the other hand, matching the other term tells us that $\gamma_i^\dagger\gamma_0 = -\gamma^0\gamma^i = \gamma^i\gamma^0$ (because γ^0 and γ^i anticommute), and the point is that we get $\gamma^0\gamma^\mu\gamma^0 = (\gamma^\mu)^\dagger$. Additionally, because the trace of γ^0 is zero, γ^0 must be even-dimensional (since each eigenvalue is ± 1). If we try two dimensions, we need four complex 2×2 matrices, and it makes sense to use the Pauli matrices to form a basis $\{I, \sigma^i\}$ – unfortunately, this doesn't work because the trace of I is 2. So we must try to make the γ^μs four-dimensional matrices. Recall that the Pauli matrices explicitly look like

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and from here we can write down a basis (the first expression in block form)

$$\gamma^0 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = I \otimes \sigma_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

along with

$$\gamma^i = \begin{bmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{bmatrix} = \sigma^i \otimes i\sigma_2;$$

this is known as the **Dirac representation**. An alternative representation we could also use is

$$\gamma^0 = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = I \otimes \sigma_1, \quad \gamma^i = \begin{bmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{bmatrix} = \sigma^i \otimes i\sigma_2;$$

this is known as the **Weyl** or **chiral representation**. The point is that either one of these will satisfy the Clifford algebra (as well as many other examples), but we can just pick any representation that works and we'll get the same physics. But summarizing our discussion, we have an equation

$$\left(i\gamma^0\partial_0 + i\gamma^i\partial_0 - m^i\right)\psi(\vec{x}, t) = 0,$$

which we can further simplify as

$$(i\partial_\mu\gamma^\mu - m)\psi(\vec{x}, t) = \boxed{(i\not{\partial} - m)\psi(\vec{x}, t) = 0},$$

where we use the "slash" notation $\not{A} = \gamma^\mu A_\mu$. We can now test and see if what we've done makes any sense – we take $0 = (i\not{\partial} - m)\psi$ and multiply it on the left by $(-i\not{\partial} + m)$, and we'll find indeed that we get $(\not{\partial}^2 + m^2)\psi = (\partial_\mu^2 + m^2)\psi$, so we get the Klein-Gordon equation back again. The point is that we've managed to introduce new quantities in a way that looks like the Schrodinger equation, which is what we wanted! Furthermore, we have

$$\not{\partial}\not{\partial} = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\partial_\mu\partial_\nu = g^{\mu\nu}\partial_\mu\partial_\nu = \partial_\mu^2.$$

We now want to write down a Lagrangian, because once we have that we'll be able to apply what we've been doing in the class so far. We may first try to write

$$\mathcal{L} = \psi^\dagger(i\not{\partial} - m)\psi$$

so that we have a scalar Lagrangian, but we find that $\mathcal{L}^\dagger = \psi^\dagger(-i\gamma^0\overleftarrow{\not{\partial}}\gamma^0 - m)\psi$, where the left arrow means $\not{\partial}$ acts on the left instead of the right, which isn't quite what we want. So we'll have to try something slightly different – we can define $\overline{\psi} = \psi^\dagger\gamma^0$ and write down

$$\mathcal{L} = \overline{\psi}(i\not{\partial} - m)\psi,$$

so that

$$\mathcal{L}^\dagger = \psi^\dagger(-i\gamma^0\overleftarrow{\not{\partial}}\gamma^0 - m)\overline{\psi}^\dagger;$$

here we've used that $\overline{\psi}^\dagger = (\psi^\dagger\gamma^0)^\dagger = (\gamma^0)^\dagger\psi = \gamma^0\psi$. So then by using integration by parts because we can add a total derivative to the Lagrangian, we end up with

$$\mathcal{L}^\dagger = -\overline{\psi}(i\overleftarrow{\not{\partial}} + m)\psi = \overline{\psi}(i\not{\partial} - m)\psi,$$

so this is a good Lagrangian to use. And for Lorentz invariance, if we replace $\psi(x)$ with $U(\Lambda)\psi(\Lambda^{-1}x)$ (the same

transformation as for scalar fields), $\psi^\dagger$ will be sent to $\psi^\dagger(\Lambda^{-1}x)U^\dagger(\Lambda)$. The mass term of our Dirac Lagrangian then tells us that we must have

$$U^\dagger(\Lambda)\gamma^0 U(\Lambda) = \gamma^0,$$

and the derivative term says that

$$U^\dagger(\Lambda)\gamma^0\Lambda^{-1}_{\mu\nu}\gamma^\mu\partial^\nu U(\Lambda) = \gamma^0\not{\partial}.$$

This means that (inserting a unit $U(\Lambda)U^{-1}(\Lambda)$)

$$U^\dagger(\Lambda)\gamma^0 U(\Lambda)U^{-1}(\Lambda)\gamma^\nu\lambda^{-1}_{\nu\mu}U(\Lambda) = \gamma^0\gamma^\mu,$$

but now the first three terms form γ^0 so what we really have is

$$U^{-1}(\Lambda)\gamma^\mu U(\Lambda) = \Lambda^{\mu\nu}\gamma_\nu \implies U^{-1} = \gamma^0 U^\dagger\gamma^0.$$

(Note that U is different from the time-evolution operator that we derive before!) The point is that an arbitary infinitesimal Lorentz transformation looks like

$$\Lambda^{\mu\nu} = g^{\mu\nu} + \omega^{\mu\nu}, \quad \Lambda^\mu{}_\rho g^{\rho\sigma}\Lambda_\sigma{}^\nu$$

which tells us that we must have $\omega^{\mu\nu} = -\omega^{\nu\mu}$. A general transformation of a field $\phi_a(x)$ is then $U_{ab}\phi_b(\Lambda^{-1}x)$. Since we must have $U(\Lambda)U(\Lambda') = U(\Lambda\Lambda')$ and in particular $U(\Lambda)U(\Lambda^{-1}) = -I$, we must have $U = I + \tau$, where $\tau = \omega^{\mu\nu}M_{\mu\nu}$. But if we write $M_{\mu\nu} = a\gamma_\mu\gamma_\nu + b\gamma_\nu\gamma_\mu$, ω being antisymmetric shows that we must have $b = -a$. If we plug in our ansatz and use the Clifford algebra wherever we can, we find that

$$a = \frac{1}{8} \implies M_{\mu\nu} = \frac{1}{8}[\gamma_\mu, \gamma_\nu].$$

Thus if we define $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$, then for any infinitesimal transformation we have

$$U(\Lambda) = I - \frac{i}{4}\omega^{\mu\nu}\sigma_{\mu\nu},$$

meaning that for finite Lorentz transformations we get

$$U(\Lambda) = e^{-\frac{i}{4}\omega^{\mu\nu}\sigma_{\mu\nu}}.$$

Example 24

Let's see what a rotation around the z-axis does to everything here.

Recall that a generation of the rotation is $\omega = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \omega_6 I_6^{\mu\nu}$; in general, we then get

$$U(\vec{\omega}) = e^{-\frac{i}{4}\omega_i I_i^{\mu\nu}\sigma_{\mu\nu}}.$$

In the Dirac basis, we have $\sigma^{12} = \frac{i}{2}[\gamma^1, \gamma^2] = \begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix}$, so that our infinitesiaml transformation is

$$\tau = \frac{u}{2}\Delta\phi\sigma_{12} = \frac{i}{2}\Delta\psi\begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix},$$

so that our field transforms as

$$\psi \to e^{\frac{i}{2}\phi \begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix}} \psi,$$

which is indeed $\left(\cos\frac{\phi}{2} + i\sin\frac{\phi}{2}\right)\psi$. In particular, we see that ψ flips to $-\psi$ for $\phi = 2\pi$ and flips back to ψ for $\phi = 4\pi$, which is exactly what we have in spin one-half. So we have indeed arrived at fermions!

11 November 1, 2022

Last time, we introduced the Dirac equation and Clifford algebra to set up a framework for describing spin one-half particles. We started with the Klein-Gordon equation and manipulated it to get a linear equation that describes relativistic wave mechanics, using the operator $\not{\partial} = \gamma^\mu \partial_\mu$. Those matrices γ^μ need to satisfy a certain anticommutation relation $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbf{1}$, and then we saw that our wave fields transform infinitesimally as $\psi \to e^{-\frac{i}{4}\omega^{\mu\nu}\sigma_{\mu\nu}}\psi$ under $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$.

Reviewing some properties of our γ matrices, we want to solve the equation $(i\not{\partial} - m)\psi = 0$, expanding out the indices gives us the equation

$$i\partial_0\psi = \left((-\gamma^0)^{-1}\gamma^i\partial_i + (m\gamma^0)^{-1}\right)\psi,$$

and plugging in a plane wave $\gamma = e^{-ipx}$ gives us the equation

$$E\psi = \left((\gamma^0)^{-1}\vec{p}\cdot\vec{\gamma} + m(\gamma^0)^{-1}\right)\psi.$$

Since we want E to be real, we saw that $(\gamma^0)^{-1} = \gamma^0$ and that $\gamma^0\gamma^i\gamma^0 = (\gamma^i)^\dagger$. Now we can get an explicit basis for the (4×4) matrices γ – we're going to use the specific choices

$$\Gamma^s = \mathbf{1}, \quad \Gamma^\nu_\nu = \gamma_\mu, \quad \Gamma^T_{\mu\nu} = \sigma_{\mu\nu}$$

(where s, v, T stand for "scalar," "vector," and "tensor"), as well as

$$T^P = i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma^5 = \frac{i}{4!}\varepsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma$$

(calling this matrix γ^5 is because it serves as a fifth basis vector, and P stands for "parity"). Additionally, we have

$$T^{AV}_\mu = \gamma_5\gamma_\mu$$

((here AV stands for "axial vector"). There are $1, 4, 6, 1, 4$ matrices for the scalar, vector, tensor, parity, and axial vector basis elements (6 for tensor because we need them to be antisymmetric), and we can check they are all linearly independent so this gives us a way to write down any 4×4 matrix as a linear combination of these terms, and each of them will correspond to some physics. We'll write down a few more properties that we can check now on our own:

- $\gamma_5^2 = \mathbf{1}$.
- $\{\gamma_5, \gamma_\mu\} = 0$ for all μ.
- $(\gamma_5)^\dagger = \gamma_5$.
- $\mathrm{Tr}[\gamma_5] = 0$.
- In the Weyl representation, we have $\gamma_5 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$ in block form.

But a Lagrangian is never a vector, so we need a way to form Lorentz scalars out of all of this, and we'll do so by creating bilinear forms with these different matrices in between. We now see the reason for our naming – we have

$$\overline{\psi}\psi = \overline{\psi}\Gamma^s\psi,$$

which is a scalar under Lorentz transformations,

$$\overline{\psi}\Gamma^V_\mu\psi = \overline{\psi}\gamma_\mu\psi,$$

which Lorentz transforms as a vector, and

$$\overline{\psi}\Gamma^T_{\mu\nu}\psi,$$

which Lorentz transforms by multiplying by $\Gamma_\mu{}^\nu$. Then we also have

$$\overline{\psi}\Gamma^P\psi = \overline{\psi}\gamma_5\psi$$

transforming as a pseudo-scalar (changing signs under parity inversion), and

$$\overline{\psi}\Gamma^{AV}_\mu\psi = \overline{\psi}\gamma_5\gamma_\mu\psi$$

transforming as an axial vector. So now if we have any $\psi = \begin{bmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{bmatrix}$, we can decompose ψ into $\begin{bmatrix}\psi_R\\ \psi_L\end{bmatrix}$ (so ψ_R is the upper two components and ψ_L is the lower two) and define the projectors

$$P_R = \frac{1}{2}(\mathbf{1}+\gamma_5) = \begin{bmatrix} I & 0\\ 0 & 0\end{bmatrix}, \quad P_L = \frac{1}{2}(\mathbf{1}-\gamma_5) = \begin{bmatrix} 0 & 0\\ 0 & I\end{bmatrix}.$$

Indeed they are projectors because $P_R^2 = P_R$ and $P_L^2 = P_L$, they are complete because $P_R + P_L = \mathbf{1}$, and they are orthogonal because $P_R \cdot P_L = P_L \cdot P_R = 0$. Additionally, they have commutation relations

$$\gamma^\mu P_R + \gamma^\mu\frac{1}{2}(\mathbf{1}+\gamma_5) = \frac{1}{2}(\mathbf{1}-\gamma_5)\gamma^\mu = P_L\gamma^\mu$$

by the anticommuting relation of γ_5 with γ_ν.

Remark 25. *All of these statements are basis-independent, but using the Weyl representation makes it easier to see how some of the computations work.*

Turning back to the Dirac equation, we see that

$$\overline{\psi}(i\partial\!\!\!/ - m)\psi = \overline{\psi}(P_R+P_L)(i\partial\!\!\!/ - m)(P_R+P_L)\psi = \overline{\psi}P_R i\partial\!\!\!/ P_L\psi + \overline{\psi}P_L i\partial\!\!\!/ P_R\psi - m\overline{\psi}(P_RP_R + P_LP_L)\psi,$$

and now if we define $\overline{\psi}P_R = \overline{\psi}_L$ and $\overline{\psi}P_L = \overline{\psi}_R$, so that $P_R\psi = \psi_R$ and $P_L\psi = \psi_L$, our Dirac Lagrangian becomes

$$\mathcal{L} = \overline{\psi}_R i\partial\!\!\!/\psi_R + \overline{\psi}_L i\partial\!\!\!/\psi_L - \overline{\psi}_R\psi_L m - \overline{\psi}_L\psi_R m.$$

So if the mass is zero, we have two copies of the same field here, and thus having two different particles sitting here gives us an enhanced degree of symmetry. We can next see the property

$$\gamma_5\psi_R = \psi_R, \quad \gamma_5\psi_L = -\psi_L,$$

so this is a **chirality** property – fermions with a definite χ are called **chiral fermions**, and if we commute this operator with the time-evolution operator we find

$$[H, \gamma_5] = H\gamma_5 + i\gamma_5\gamma^0\gamma^i\partial_i - \gamma_5\gamma^0 m = 2\gamma^0\gamma_5 m.$$

So if the mass is zero, then chirality will be unchanged under time-evolution, and so **for massless fermions this is a good quantum number**.

Keep in mind that everything here is still for the free Dirac equation. But nothing that we've done is quantum yet, so like for the classical field theory we'll want to quantize everything here. But first we should find classical solutions and study their properties. We start again with the Lagrangian

$$\mathcal{L} = \overline{\psi}(i\partial\!\!\!/ - m)\psi$$

and solve the Euler-Lagrange equations $\frac{\partial\mathcal{L}}{\partial\overline{\psi}} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\right) = 0$ and similar for $\overline{\psi}$. We end up with

$$-\overline{\psi}m + \partial_\mu\left(\overline{\psi}i\gamma^\mu\right) = 0, \quad \left(i\partial\!\!\!/ - m\right)\psi = 0$$

so this is the Dirac equation from the Dirac Lagrangian (as we expect). If we now define

$$\psi_{1,2} = u_\pm(p) = e^{-ipx}, \quad \psi_{3,4} = v_\mp(p)e^{ipx},$$

we want to have

$$(i\partial\!\!\!/ - m)\psi_{1,2} = (p\!\!\!/ - m)\psi_{1,2} = 0, \quad (i\partial\!\!\!/ - m)\psi_{3,4} = -(p\!\!\!/ + m)\psi_{3,4} = 0$$

where $p\!\!\!/ = p^\mu\gamma_\mu$. We can make our life easy by defining the **spinors**

$$u_\pm(p) = (p\!\!\!/ + m)A_\pm(p)u_{0\pm}, \quad v_\mp(p) = -(p\!\!\!/ - m)B_\mp(p)v_{0\mp},$$

where $A_\pm(p)$ and $B_\pm(p)$ are some scalars and

$$u_{0+} = \sqrt{2m}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \quad u_{0-} = \sqrt{2m}\begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \quad v_{0-} = \sqrt{2m}\begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \quad v_{0+} = \sqrt{2m}\begin{bmatrix}0\\0\\0\\1\end{bmatrix}.$$

Plugging into the Dirac equation and using some matrix algebra, we find that

$$(p\!\!\!/ + m)(p\!\!\!/ - m) = p\!\!\!/\,p\!\!\!/ - m^2 = \frac{1}{2}p^\mu p^\nu\{\gamma_\mu, \gamma_\nu\} - m^2 = p^\mu p^\nu\eta_{\mu\nu} - m^2 = p^2 - m^2 = 0,$$

assuming that we are **on shell** (meaning that we have a particle living on its mass shell). We can then also define

$$\chi_+ = \begin{bmatrix}1\\0\end{bmatrix}, \quad \chi_- = \begin{bmatrix}0\\1\end{bmatrix},$$

so that $u_{0\pm} = \sqrt{2m}\begin{bmatrix}\chi_\pm\\0\end{bmatrix}$ – this will be useful later.

Next, it turns out these spinors are properly normalized: we want to have

$$\overline{u}_+(p)u_+(p) = 2m,$$

a Lorentz scalar, where $\overline{U} = U^\dagger \gamma^0$. We then have to define our constants $A_\pm$ so that

$$\begin{aligned} 2m = \overline{U}_\pm(p)U_\pm(p) &= |A_\pm|^2 u_{0\pm}^\dagger (\not{p} + m)^\dagger \gamma_0 (\not{p} + m) u_{0\pm} \\ &= |A_\pm|^2 \overline{u}_{0\pm}(\not{p} + m)(\not{p} + m)u_{0\pm} \\ &= |A_\pm|^2 \overline{u}_{0\pm}(p^2 + m^2 + 2\not{p}m)u_{0\pm}. \end{aligned}$$

Picking a particular basis, we have the block form

$$\not{p} = \begin{bmatrix} p^0 I & \vec{p}\cdot\vec{\sigma} \\ -\vec{\sigma}\cdot\vec{p} & -p^0 I \end{bmatrix}$$

where $\vec{\sigma}$ is the Pauli matrices, and plugging everything in yields

$$2m = |A_\pm|^2 (2m) \left((2m^2) + 2 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} p^0 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \right)$$

(we only care about the first coordinate because of the dot product), which means

$$A_\pm = \frac{1}{\sqrt{2m(p^0 + m)}}.$$

Similarly, we will want the normalization

$$\overline{v}_\pm v_\pm = -2m \implies B_\pm = A_\pm.$$

So collecting terms, we find that our spinors are of the form

$$u_\pm = \frac{\sqrt{2m}}{\sqrt{2m(p^0+m)}}(\not{p} + m) \begin{bmatrix} \chi_\pm \\ 0 \end{bmatrix} = \frac{1}{\sqrt{m+p^0}} \begin{bmatrix} (p^0+m)\chi_\pm \\ -\sigma_i p_i \chi_\pm \end{bmatrix},$$

and similarly

$$v_\pm = \frac{1}{\sqrt{m+p_0}} \begin{bmatrix} \sigma_i p_i \chi_\mp \\ (p^0+m)\chi_{mp} \end{bmatrix}.$$

So we have explicit solutions to our Dirac equation now, and more generally

$$\overline{u}_r u_s = 2m\delta_{rs}, \quad \overline{v}_r v_s = 2m\delta_{vs}, \quad \overline{u}_r v_s = \overline{v}_r u_s = 0,$$

where r, s basically label $\pm$. If we define $\Lambda_\pm = \frac{\pm\not{p}+m}{2m}$, we find that

$$\Lambda_+ u_\pm = u_\pm, \quad \Lambda_- v_\pm = v_\pm, \quad \Lambda_+ v_\pm = 0, \quad \Lambda_- u_\pm = 0.$$

Thus we find that $\Lambda_+ + \Lambda_- = \mathbf{1}$, and we have

$$2m\Lambda_+ = \boxed{\sum_{s=\pm} u_s(p)\overline{u}_s(p) = \not{p} + m}$$

and similarly

$$2m\Lambda_- = \boxed{\sum_{s=\pm} v_s(p)\overline{v}_s(p) = -\not{p} + m}.$$

These can be directly computed, or we can use the rules above to see that $(\sum u_s \overline{u}_s)u_+ = 2mu_+$ and note that Λ_+ should be the identity on u_+.

Remark 26. *It will turn out that Λ_+ projects out a particle with spin $\pm\frac{1}{2}$, and Λ_- projects out an antiparticle with spin $\mp\frac{1}{2}$, but we'll see this more concretely later.*

So we want to use these classical solutions $u_\pm, v_\pm$ for quantization later, and the boxed equations are really what's important in all of these derivations. We'll now move to talking about **discrete symmetries** of the Dirac Lagrangian:

- Parity: we know that $\Lambda^\mu{}_\nu = \text{diag}(1,-1,-1,-1)$ is a valid transformation, and we have

$$U^{-1}\gamma^\mu U = \Lambda^\mu{}_\nu \gamma^\nu, \quad U = \gamma^0 e^{i\phi} \implies U^{-1} = \gamma^0 e^{-i\phi}.$$

  (So under parity, ψ goes to $\gamma^0 e^{i\phi}\psi$.)

- Time reversal: we can check that a Lorentz transformation $\text{diag}(-1,1,1,1)$ corresponds to $U = i\gamma_1\gamma_3$.

- Charge conjugation: this is useful for quantum electrodynamics, in which we want invariance between particles and antiparticles. Here if we have

$$\mathcal{L}_{\text{QED}} = -e\overline{\psi}A^\mu\gamma_\mu\psi = -e\overline{\psi}\not{A}\psi,$$

  and we do the charge conjugation transformation

$$\overline{\psi}(i\not{\partial} - e\not{A} - m)\psi \to \overline{\psi}_c(i\not{\partial} + e\not{A} - m)\psi_c,$$

  then we want to figure out how the fields should correspondingly transform. Taking complex conjugates flips the sign of the $i\not{\partial}$ term, so the transformation gives us

$$(psi^\dagger)^*(\gamma^0)^*(-i\not{\partial}^* - e\not{A}^* - m)\psi^*,$$

  and now if we introduce $\psi_c = C\psi^*$ for some invertible matrix $C^{-1}C = \mathbf{1}$, this expression is the same as

$$((C^{-1}C\psi)^\dagger)^*\gamma^0(-i(\gamma^\mu)^*\partial_\mu - e\not{A}^* - m)C^{-1}C\psi^* = \psi_c^\dagger(C^{-1})^\dagger\gamma^0\left(-i(\gamma^\mu)^*\partial_\mu - i\not{A}^* - m\right)C^{-1}\psi_c.$$

  Now matching terms with our ansatz, we find that $(C^{-1})^\dagger C^{-1} = \mathbf{1}$ so $C^\dagger = C^{-1}$, and $C(\gamma^\mu)^*C^{-1} = -\gamma^\mu$. Furthermore, the imaginary part of $\gamma^0, \gamma^1, \gamma^3$ should be zero, $\{C,\gamma^0\} = \{C,\gamma^1\} = \{C,\gamma^3\} = 0$, and then we should have $\gamma_2^* = -\gamma_2 \implies [C,\gamma_2] = 0$ and $\gamma_2^2 = -\mathbf{1}$. So flipping the charge of the electron is a symmetry: we have

$$C = i\gamma_2 \implies \psi_c = i\gamma_2\psi^*$$

  in this particular representation that we've chosen.

Next time, we'll finish up discussing properties of the Dirac Lagrangian and move on to electrodynamics!

12 November 3, 2022

Last time, we analyzed the Dirac Lagrangian $\mathcal{L} = \overline{\psi}(i\not{\partial} - m)\psi$ and found some classical solutions to the Dirac equation of motion: letting $\chi_+ = \begin{bmatrix}1\\0\end{bmatrix}$ and $\chi_- = \begin{bmatrix}0\\1\end{bmatrix}$ (here we're looking at the **spinor space**, since ψ is of the form $\begin{bmatrix}\psi_1\\\psi_2\\\psi_3\\\psi_4\end{bmatrix}$),

we wrote down plane wave solutions

$$\psi = u_+(p)e^{-ipx} + u_-(p)e^{-ipx} + v_+e^{ipx} + v_-e^{ipx},$$

where $p^2 = m^2$ and in the Dirac representation we have

$$u_\pm = \frac{1}{\sqrt{p^0+m}}\begin{bmatrix}(p^0+m)\chi_\pm \\ -\vec{\sigma}\cdot\vec{p}\chi_\pm\end{bmatrix}, \quad v_\pm = \frac{1}{\sqrt{p^0+m}}\begin{bmatrix}\vec{\sigma}\cdot\vec{p}\chi_\mp \\ (p)+m)\chi_\mp\end{bmatrix}.$$

In particular, we get unit vectors in the spinor space if we take $\vec{p}\to 0$ and $p^0\to m$. We then looked at symmetry transformations (parity, time reversal, and so on) and tried to figure out how the spinors change. Specifically, charge conjugation symmetry $\psi\mapsto\psi_c = i\gamma_2\psi^*$, the Lagrangian stays constant and we have $-e\overline{\psi}\not{A}\psi$ sent to $e\overline{\psi}\not{A}\psi$. So now if we look at the u_+ and u_- terms of our plane wave solution, we have under charge conjugation that

$$\psi_{c,u} = i\gamma_2\psi_u^* = \frac{1}{\sqrt{p^0+m}}\begin{bmatrix}0 & i\sigma_2 \\ -i\sigma_2 & 0\end{bmatrix}\begin{bmatrix}(p^0+m)\chi_\pm \\ -\vec{\sigma}\cdot\vec{p}\chi_\pm\end{bmatrix}e^{ipx}$$

(recall that the first matrix above is $i\gamma_2$ in block form), and now we can perform the multiplication: because $i\sigma_2\chi_\pm = \chi_\mp$ by definition of the Pauli matrix, and $i\sigma_2(\vec{\sigma}\cdot\vec{p}) = -\vec{\sigma}\cdot\vec{p}\chi_\mp$,

$$\psi_{c,u} = \frac{1}{\sqrt{p^0+m}}\begin{bmatrix}\vec{\sigma}\cdot\vec{p}\chi_\mp \\ (p^0+m)\chi_\mp\end{bmatrix} = v_\pm e^{ipx} = \psi_v$$

so something we associate with particles ends up being associated with antiparticles. So this is explaining why when we change the sign of the charge (dealing with positrons instead of electrons), everything else will be the same.

We'll now introduce certain linear combinations of spinors

$$\psi_1 = u_+e^{-ipx} - v_-e^{ipx}, \quad \psi_2 = u_-e^{-ipx} + v_+e^{ipx},$$

where $\psi_{1,c} = \psi_1$ and $\psi_{2,c} = \psi_2$, so these particles are their own antiparticles (we call them **Mayorama particles**) – for example, neutrinos could satisfy these properties, but we presently don't know whether they are.

Example 27

One ongoing idea for checking whether this property holds for neutrinos is to observe the **neutrinoless double β-decay** process ($0\nu\beta\beta$). Consider the particle decay

$$N\to P + e^- + \overline{\nu}_e.$$

(In particular, the maximum energy of the electron would be $E_{\max}(e^-) = M_N - M_p - m_\nu$, and repeating this can give us an estimate of the neutrino mass. This is ongoing work – we know that $m_{\nu_e} < 0.7$ eV, and the best-fit value is around 0.26 ± 0.34 eV.) This process has to do with Mayorama fermions, because (when doing this decay process) we could imagine having two neutrons coming in, both undergo beta decay, but with a neutrino absorbed from one decay vertex to the other if neutrinos and antineutrinos are the same. So such a coincident decay could be observed, and that would imply that the neutrino must indeed be a Mayorama particle.

There's one more transformation that we care about, called **helicity**. Recall that we previously introduced

$$\psi\to e^{-\frac{i}{4}\omega^{\mu\nu}\sigma_{\mu\nu}}\psi, \quad \sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu,\gamma_\nu]$$

and we saw that we have spin one-half by considering the case σ_{12}. We can now write down a more general **spin**

matrix

$$\vec{\Sigma} = \frac{i}{2}\left([\gamma^2,\gamma^3],[\gamma_1,\gamma_3],[\gamma_1,\gamma_2]\right) = \begin{bmatrix}\vec{\sigma} & 0\\ 0 & \vec{\sigma}\end{bmatrix}$$

and we define the **helicity** to be the projection of this matrix in the direction of our momentum:

$$h(p) = \frac{1}{2}\vec{\Sigma}\cdot\frac{\vec{p}}{|\vec{p}|}.$$

Helicity is then a valid quantum number, because recalling that we have the equation of motion

$$i\partial_0\psi = H\psi = \left(-i\gamma^0\gamma^i\partial_i + \gamma^0 m\right)\psi,$$

we can write the right-hand side as $\begin{bmatrix} m & \vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -im\end{bmatrix}\psi$ in the Dirac representation or $\begin{bmatrix}-\vec{\sigma}\cdot\vec{p} & m \\ -m & \vec{\sigma}\cdot\vec{p}\end{bmatrix}\psi$ in the Weyl representation, and then we have $[h(p), H] = 0$. So we can find some states of definite helicity ψ_h, but if we want to stick with the Dirac representation, we can imagine we have a particle flying in the z-direction $\vec{p} = \begin{bmatrix}0\\0\\p_z\end{bmatrix}$. We then have

$$h(e_z) = \frac{1}{2}\begin{bmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{bmatrix} \implies h(e_z)\chi_\pm = \pm\frac{1}{2}\chi_\pm,$$

meaning that $h(e_z)u_\pm = \pm\frac{1}{2}u_\pm(p)$ and thus a $+1/2$ state (a **right-handed particle**) if the spin and direction of motion are aligned and $-1/2$ (a **left-handed particle**) if they are antialigned. We'll use λ for helicity. We can similarly find that for the left-moving plane wave, $h(e_z)v_\pm = \mp\frac{1}{2}v_\pm(p)$, and we have $\lambda = \frac{1}{2}$ if the spin and particle are aligned again (but this time we have a **right-handed antiparticle**) and $\lambda = -\frac{1}{2}$ if we have opposite alignment (a **left-handed antiparticle**). But we **don't have Lorentz invariance** for the helicity – for a particle with mass, we can imagine boosting into a frame faster than the momentum $\vec{p}$, which would flip the sign. However, it's still useful to measure in a particular frame in the lab (for example polarization), and it's still a valid quantum number.

So in summary, Lorentz invariant combinations of spinors take the form $\overline{\psi}\Gamma_a\psi$ (where the Γs live in the space of 16 matrices that transform in various ways that we've previously discussed), we have four fields in ψ, corresponding to particles and antiparticles with helicity-spin in or opposite the direction of motion. We're now ready to quantize our fields here, and the anticommutation turns out be a little weird. We want to think about observables, so we must first figure out what quantities are actually observable – the issue is that we don't know what to do with the indices of ψ. So observables are really given by **local observables** at a single spacetime point

$$O(\vec{x},t) = \overline{\psi}(\vec{x},t)\Gamma_a\psi(\vec{x},t),$$

for which we know how these transform (some as scalars, some as vectors, etc. under Lorentz transformations) and all of the ψ indices are now contracted. We then see that we want causality, so we require

$$[O(\vec{x},t), O(\vec{y},t)] = 0 \text{ if } (x-y)^2 < 0 \implies [\overline{\psi}(x)\Gamma_1\psi(x), \overline{\psi}(y)\Gamma_2\psi(y)] = 0$$

for spacelike separation. One such observable is the momentum p^μ, which is the conserved quantity associated with the Noether current of spatial translation:

$$\begin{aligned} p^\mu &= \int d^3x \frac{\partial\mathcal{L}}{\partial(\partial_0\phi_a)}\partial^\mu\phi_a - \eta^\mu{}_0\mathcal{L} \\ &= \int d^3x \left[i\overline{\psi}\gamma^0\partial^\mu\psi - \eta^\mu{}_0\left(\overline{\psi}i\not{\partial}\psi - \overline{\psi}m\psi\right)\right]. \end{aligned}$$

We do see that everything written here is in terms of "bilinear forms" of ψ, and now we want to know how to work with this. Much like our procedure for the scalar field, we postulate that we have a mode expansion

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega(\vec{p})}} = \sum_{s=\pm\frac{1}{2}} \left[a_s u_s(p)e^{-ipx} + b_s^\dagger v_s(p)e^{ipx}\right]$$

where because we have a complex field we have two sets of creation and annihilation operators, a and b. We then have

$$\overline{\psi}(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega(\vec{p})}} = \sum_{s=\pm\frac{1}{2}} \left[a_s^\dagger \overline{u}_s(p)e^{ipx} + b_s \overline{v}_s(p)e^{-ipx}\right],$$

where remember that $\overline{u}_s = u_s^\dagger\gamma^0$. We now want to interpret $a_s^\dagger(p)$ as creating a spin s (s is one of $\pm 1/2$) fermion field / particle and similar for $b_s^\dagger$, so we **want** the commutation relations

$$[p^\mu, a_s(p)] = -p^\mu a_s(p), \quad [p^\mu, b_s(p)] = -p^\mu b_s(p), \quad [p^\mu, a_s^\dagger(\vec{p})] = p^\mu a_s^\dagger(p), \quad [p^\mu, b_s^\dagger(\vec{p})] = p^\mu b_s^\dagger(p).$$

We can then see that (combining these requirements with the definition of ψ and $\overline{\psi}$) that fermion particles and antiparticles must satisfy the anticommutation relations

$$\{a_r(p), a_r^\dagger(q)\} = (2\pi)^3\delta_{rs}\delta^{(3)}(\vec{p}-\vec{q}), \quad \{b_r(p), b_s^\dagger(q)\} = (2\pi)^3\delta_{rs}\delta^{(3)}(\vec{p}-\vec{q})$$

and all other commutation relations $\{a_r^\dagger(p), a_s^\dagger(q)\}, \{a_r(p), a_s(q)\}$ equal to zero (and all as and bs also anticommute). We then also see that we must have

$$\{\psi(\vec{x},t), \overline{\psi}(\vec{y},t)\} = \gamma^0\delta^{(3)}(\vec{x}-\vec{y}), \quad \{\psi(\vec{x},t), \psi(\vec{y},t)\} = \{\overline{\psi}(\vec{x},t), \overline{\psi}(\vec{y},t)\} = 0.$$

Example 28

We'll see one example of this in action to demonstrate the calculations involved.

Plugging in the mode expansions and looking at the $t = 0$ timeslice to make things easier, we have

$$\{\psi(\vec{x},0), \overline{\psi}(\vec{x},0)\} = \int \frac{d^3p}{(2\pi)^3}\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega(p)}\sqrt{2\omega(\vec{q})}} \sum_{r,s=\pm\frac{1}{2}} \Big[e^{i\vec{p}\cdot\vec{x}}e^{-i\vec{q}\cdot\vec{y}} v_r(\vec{p})\overline{v}_s(q)\{b_r^\dagger(\vec{p}, b_s(\vec{q})\}$$

$$+e^{-i\vec{p}\cdot\vec{x}}e^{i\vec{q}\cdot\vec{y}} u_r(\vec{p})\overline{u}_s(\vec{q})\{a_r(\vec{p}), a_s^\dagger(\vec{q})\}\Big]$$

(here the other two terms instantly vanished because the anticommutators were zero). Integrating over q, this simplifies to

$$= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega(p)} \sum_{s=\pm\frac{1}{2}} \left[e^{i\vec{p}(\vec{x}-\vec{y})} v_s\overline{v}_s + e^{-i\vec{p}(\vec{x}-\vec{y})}(u_s\overline{u}_s)\right]$$

and now we must remember (from the explicit representations of our Dirac spinors) that $\sum_{s=\pm\frac{1}{2}} v_s(p)\overline{v}_s(p) = \not{p} + m$ and $\sum_{s=\pm\frac{1}{2}} u_s(p)\overline{u}_s(p) = \not{p} - m$, so this simplifies further to

$$\int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega(p)} \left[e^{-ip(\vec{x}-\vec{y})}(\not{p} - m) + e^{-ip(\vec{x}-\vec{y})}(\not{p} + m)\right]$$

and if we substitute $\vec{p} \to -\vec{p}$, turning $\not{p}$ into $\gamma^0 p^0 + \gamma^i p^i$, this becomes

$$\int \frac{d^3p}{(2\pi)^3} e^{-ip(\vec{x}-\vec{y})}\left[\gamma^0 p^0 - \gamma^i p_i - m + \gamma^0 p^0 + \gamma^i p^i + m\right] = \int \frac{d^3p}{(2\pi)^3}\int e^{-i\vec{p}(\vec{x}-\vec{y})} = \gamma^0\delta^{(3)}(\vec{x}-\vec{y}),$$

as desired. So the key difference is that we'll have to start doing spin sums and take combinations of our slash operators so that they cancel out, but in the end it's very similar to what we did with scalar fields. For observables, this means that

$$[\overline{\psi}(\vec{x})\Gamma_1\psi(\vec{x})\overline{\psi}(\vec{y})\Gamma_2\psi(\vec{y})]$$

and use that $[AB, CD] = A\{B, C\}D - AC\{B, D\} - C\{A, D\}B + \{C, A\}DB$ to write this as

$$= \overline{\psi}(\vec{x})\left(\Gamma_1\gamma^0\Gamma_2 - \Gamma_2\gamma^0\Gamma_1\right)\psi(\vec{x})\delta^{(3)}(\vec{x}-\vec{y}),$$

so **equal-time commutation relations are zero whenever we live not on the same spacetime curve**, and we can always find a Lorentz boost back to this whenever we have spacelike separation, so we do have the desired zero commutators. That means that $\psi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega(p)}}\sum_{s=\pm\frac{1}{2}}\left[e^{-ipx}a_s(p)u_s(p) + e^{ipx}b_s^\dagger(p)v_s(p)\right]$ annihilates particles of spin s and momentum $\vec{p}$ at x and creates anti-particles with that spin and momentum.

13 November 10, 2022

In the last few lectures, we've thought about how to study fermions in quantum field theory. We start with classical (Dirac) theory and wrote down a Lagrangian, giving us the Dirac equation of motion which can be written in terms of certain γ matrices (acting on spinors ψ) that satisfy the Clifford algebra.

Last time, we studied quantum observables in this formalism and wanted commutation relations to make sense locally (in particular having causality, so that we have commutation relation). We also wrote down mode expansions for our solutions in terms of creation and annihilation operators a_s and $b_s^\dagger$, seeing that we get the expected complex scalar field commutation relations.

We'll now finish the quantum mechanical discussion of fermions and see what their properties look like – the thing that's still missing at that point is a discussion of photons. We'll start by thinking about the spectrum, which we've already hinted at: we want to create momentum eigenstates, where we look at the momentum operator P^μ associated with the Noether current. We find that

$$[P^\mu, a_s^\dagger(p)]\left|0\right\rangle = P^\mu a_s^\dagger\left|0\right\rangle = p^\mu a_s^\dagger\left|0\right\rangle$$

(that is, we have an eigenvector of P^μ), so we'll define the electron state

$$\left|e^-(p,s)\right\rangle = \sqrt{2E_p}a_s^\dagger(p)\left|0\right\rangle$$

and the positron state

$$\left|e^+(p,s)\right\rangle = \sqrt{2E_p}b_s^\dagger(p)\left|0\right\rangle.$$

We then have the correct normalization

$$\left\langle e^-(p,s)\middle|e^-(p',s')\right\rangle = \left\langle 0\middle|a_s(p)a_{s'}^\dagger(p')\middle|0\right\rangle\sqrt{2E_p 2E_{p'}} = \sqrt{2E_p 2E_{p'}}\left\langle 0\middle|a_s(p), a_{s'}^\dagger(p')\middle|0\right\rangle = (2\pi)^3 2E_p\delta^{(3)}(\vec{p}-\vec{p}')\delta_{s,s'},$$

where we've used the anticommutation relations for a_ps. And now if we create two particles, we see that we have the antisymmetry

$$a_r^\dagger(p_1)a_s^\dagger(p_2)\left|0\right\rangle = -a_s^\dagger(p_2)a_r^\dagger(p_1)\left|0\right\rangle$$

because the anticommutator is zero, so $\left|e^-(p_1,r), e^-(p_2,s)\right\rangle = -\left|e^-(p_2,s), e^-(p_1,r)\right\rangle$ and we have Dirac particles, specifically fermions (antisymmetry of the wavefunction under exchange). If we now consider the generic one-particle

wavefunction

$$|f\rangle = \int \frac{d^3p}{(2\pi)^3} \sum_{s=\pm\frac{1}{2}} f_s(p) a_s^\dagger(p) |0\rangle$$

(which could be a wavepacket or a state at a particular frequency) and we consider "two of them in a single wavefunction," we have

$$|2f\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} \sum_{s,r} f_s(p) f_r(q) a_s^\dagger(p) a_r^\dagger(p) |0\rangle$$

and applying antisymmetry, we also see that

$$|2f\rangle = -\int \frac{d^3p}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} \sum_{s,r} f_s(p) f_r(q) a_r^\dagger(p) a_s^\dagger(p) |0\rangle$$

and then swapping the roles of p and q gives us back the original expression, so $|2f\rangle = -|2f\rangle$ and thus $|2f\rangle = 0$ – this is the **Pauli exclusion principle**.

Remark 29. *Remember that the whole reason we have anticommutation relations for fermions is that we required* $[\overline{\psi}(x)\Gamma_1\psi(x), \overline{\psi}(\vec{y})\Gamma_2\psi(\vec{y})]$ *must be some number times* $\delta^{(3)}(\vec{x}-\vec{y})$ *if we want causality. And commutation relations for such observables requires anticommutation of fields.*

We'll now turn to the conserved quantities in Dirac theory – we'll follow a similar path as what we've done before. We have

$$\mathcal{L} = \overline{\psi}(i\not{\partial} - m)\psi \implies \partial_\mu j^\mu = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\mu)} \delta\phi_\mu \right) - \delta\mathcal{L}.$$

Consider an infinitesimal translation by a^μ, so that we have

$$\delta\psi = a^\mu \partial_\mu \psi, \quad \delta\overline{\psi} = a^\mu \partial_\mu \overline{\psi} \implies \delta\mathcal{L} = a^\mu \partial^\nu (\eta_{\mu\nu}\mathcal{L}),$$

and for translation invariance, we see that

$$j^{\mu\nu} = \overline{\psi} i j^\mu \partial^\nu \psi.$$

Plugging in our mode expansion, we then find that going into normal ordered form,

$$P_1^\mu = \int \frac{d^3p}{(2\pi)^3} p^\mu \sum_s (a_s^\dagger a_s - b_s b_s^\dagger) = \int \frac{d^3p}{(2\pi)^3} p^\mu \sum (a_s^\dagger a_s + b_s^\dagger b_s - (2\pi)^3 \delta^{(3)}(0)\delta_{s,s})$$

but much like before this rightmost infinity term is a vacuum energy which we can ignore (since we only care about energy differences). Similarly, if we consider multiplication by a phase $\psi \to e^{iq}\psi$ and $\overline{\psi} \to e^{-iq}\overline{\psi}$, then we have a current $j^\mu = q\overline{\psi}\gamma^\mu\psi$, so that

$$Q = \int d^3x j^0 = q \int \frac{d^3p}{(2\pi)^3} \sum_s (a_s^\dagger a_s - b_s^\dagger b_s),$$

and notice that this is a count of particles minus antiparticles. So if we take $q = -e$, we see that this gives us **charge conservation** from symmetry of the Dirac Lagrangian.

We can list a few useful facts: the Hamiltonian $H = p^0$ is positive definite, the charge operator Q is indefinite, $[Q, P^\mu] = 0$, ψ is the field operator for fermions, and momentum eigenstates are $|p,s\rangle = \sqrt{2E} a_s^\dagger(p)|0\rangle$ for particles and $|\overline{p},s\rangle = \sqrt{2E} b_s^\dagger(p)|0\rangle$ for antiparticles (the $\overline{p}$ is just notation for having an antiparticle). So now we can solve our free Dirac theory in the same way as before: we have non-equal-time anticommutation relations

$$\{\psi(x), \overline{\psi}(y)\} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \frac{d^3q}{(2\pi)^3} \frac{1}{\sqrt{2E_q}} \sum_{s,r} \left[e^{ipx-iqy} v_r \overline{v}_s(q) \{b_v^\dagger(p), b_s(q)\} + e^{-ipx+iqy} u_v(p) \overline{u}_s(q) \{a_r(p), a_s^\dagger(q)\} \right]$$

(here we've only kept the nonzero anticommutators); plugging in our anticommutation relations gives us a delta

function that cancels out one of the integrals and one of the spin sums, simplifying to

$$= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_p}\sum_s [v_s\overline{v}_s e^{ip(x-y)} + e^{-ip(x-y)}u_s\overline{u}_s].$$

But noting that $\sum_s v_s\overline{v}_s = \not{p} - m$ and $\sum_s u_s\overline{u}_s = \not{p} + m$ (from exercises), plugging back in and writing in terms of derivatives yields

$$\{\psi(x), \overline{\psi}(y)\} = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_p}(i\not{\partial} - m)(e^{-ip(x-y)} - e^{ip(x-y)}),$$

and now taking the differential operator out of the integration, this is

$$= (i\not{\partial} + m)(D(x-y) - D(y-x)),$$

so the two-point correlation function should look familiar to us. We can now define the Feynman propagator (vacuum expectation of the time-ordered product of two scalar fields) in exactly the same fashion: since we had

$$D_F(x-y) = \langle 0|T\{\phi(x), \phi(y)\{0|0\rangle = \Theta(x^0 - y^0)D(x-y) + \Theta(y^0 - x^0)D(y-x),$$

we now have something similar for fermions: the **Feynman propagator for fermions** is

$$S_F(x-y) = (i\not{\partial} + m)D_F(x-y) = \int \frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}\frac{i(\not{p} + m)}{p^2 - m^2 + i0}.$$

We then find that

$$(i\not{\partial} - m)S_F(x-y) = \int \frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}\frac{i(\not{p} + m)(\not{p} - m)}{p^2 - m^2 + i0} = i\delta^{(4)}(x-y)\cdot I_4.$$

So we've solved the Dirac theory as long as there is no coupling – the solutions to the vacuum two-point correlation function give us the entire theory, and we found the time-ordered product and generic two-point function for fermions. So this is basically all we could do for the free Dirac theory, and we've found charge conservation, antisymmetric wavefunctions, and the Pauli principle.

We'll now move on to photons, starting with a review: we study the electromagnetic field

$$E_i = -\partial_i\phi - \partial_0 A_i, \quad B_i = \varepsilon_{ijk}\partial^j A^k,$$

and where we can frame electromagnetism in terms of a field tensor

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu,$$

and writing $\phi = A^0$ and $F^{\mu\nu} = \begin{bmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & -B_1 & 0 \end{bmatrix}$, we have the Lagrangian $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$. The equations of motion are then

$$\partial_\mu F^{\mu\nu} = 0 \implies \partial^\mu\partial_\mu A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0,$$

where this is basically telling us that $\begin{bmatrix} -\partial E^i \\ -\partial_0 E_1 + \varepsilon_{ijk}B_k \\ -\partial_0 E_2 + \varepsilon_{ijk}B_k \\ -\partial_0 E_3 + \varepsilon_{ijk}B_k \end{bmatrix}$ (so we need $\vec{\nabla}\cdot\vec{E} = 0$ and $\frac{\partial\vec{E}}{\partial t} = \vec{\nabla}\times\vec{B}$). One important property of electromagnetism is that we have a gauge transformation under which Maxwell's equations stay the same:

replacing the vector potential A_i with $A_i - \partial_i \Lambda$ and replacing ϕ with $\phi + \partial_0 \Lambda$ (in other words, writing $A^\mu \to A^\mu + \partial^\mu \Lambda$), and the equations of motion stay the same so the physics should do so as well. This can be seen easily by looking at the Lagrangian: we have (here note $\partial^2 = \partial_\mu \partial^\mu$, so ∂^2 refers to applying two derivatives, not the 2 component)

$$\partial^2 A^\nu - \partial^\nu(\partial A) \to \partial^2 A^\nu + \partial^\nu \partial^2 \Lambda - \partial^\nu(\partial A) - \partial^\nu(\partial^2 \Lambda)$$

and the second and fourth term cancel out. Since the physics is not changed, we can therefore choose a convenient Λ – there are a few **gauge choices** that work well. For example, we can use the **Lorentz gauge** $\partial_\mu A^\mu \to \partial_\mu A^\mu + \partial^2 \Lambda = \partial_\mu A'^\mu$, so that $\partial^2 \Lambda = -\partial_\mu A^\mu$ and therefore $\partial_\mu A'^\mu = 0$. On the other hand, there is the **Coulomb gauge** where $\partial_0 \lambda = -\phi$, giving us a shifted potential $\phi' = 0$ and where $\partial_i A_i = 0$. So we should count degrees of freedom in our A^μs – a classical photon has two degrees of freedom (polarization states), and if we want A^μ to be a field that we put in our Lagrangian and eventually quantize, we need to account for the fact that A^μ has four degrees of freedom. That's what our symmetries do – the Lorentz gauge is actually a family of gauges because A^μ still has three degrees of freedom in that case, and we'll fix that other degree of freedom soon. But the point is that the two remaining degrees of freedom will be the correct ones.

Photon quantization is easier than fermion quantization because we have a vector of four scalar fields: we know that A^μ transforms as $A^\mu{}_\nu A^\nu$ under Lorentz transformations, so

$$[A^\mu, \Pi^\mu] = i g^{\mu\nu} \delta^{(3)}(\vec{x} - \vec{y}).$$

So we should do something completely analogous as for scalar particles, but we run into a bit of a snag: the 0th component of Π in this case is $\Pi^0 = \frac{\partial \mathcal{L}}{\partial(\partial_0 A^0)} = 0$, so we can't have the desired commutation relations. So we're not going to fix the field values from the start, and instead we'll modify our Lagrangian slightly, setting

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{\lambda}{2} (\partial_\mu A^\mu)^2$$

and use Lagrange multipliers. We find that we must have $\partial^2 A^\mu - (1 - \lambda)\partial^\mu(\partial_\nu A^\nu) = 0$; setting $\lambda = 1$ gives us the **Feynman gauge**, $\lambda = \infty$ yields the **Landau gauge**, and $\lambda = 0$ gives us the **unitary gauge**. (We'll mostly be using the Feynman gauge.) This then enables us to write down

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = \partial^\mu A^0 - \partial^0 A^\mu - \lambda \partial_0 A^0 g^{0\mu} \implies \Pi^0 = -\lambda \partial_0 A^0.$$

So if the Lorentz gauge is not manifest at the operator level, we'll instead require that on whatever physical set of states we define it on, we have the expectation $\langle \psi_{\text{phys}} | \partial_\mu A^\mu | \psi_{\text{phys}} \rangle = 0$. (This is called the Gupta-Bleuler quantization – it may look like a crutch, but we'll see how it comes up usefully for quantization procedure.)

14 November 15, 2022

Last time, we started trying to frame electromagnetism in terms of quantum field theory. We know that solving the equations of motion for the Lagrangian $\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ leads us to quantizing the vector potential $A^\mu = (\phi, A^i)$, and we find that the usual canonical momentum $\pi^\mu = \partial^\mu A^0 - \partial^0 A^\mu$ leads us to $\Pi^0 = 0$, which is bad. So instead, we need to appeal to gauge freedom and consider the local symmetry $A'_\mu = A_\mu + \partial_\mu \Lambda$. We chose the Lorentz gauge (which is relativistically invariant) in which we can guarantee $\partial_\mu A'^\mu = 0$ (classically this is adding a gauge-fixing Lagrangian $-\frac{1}{2}\xi(\partial_0 A^\mu)^2$ and then using Lagrange multipliers), and then we find that $\Pi^\mu = \partial^\mu A^0 - \partial^0 A^\mu - \xi \eta^{0\mu} \partial^0 A^0$. But on the quantum mechanical level we'll do a different modification, splitting our set of total space of states into a physical and a nonphysical part, where the physical states are those where we have $\langle \psi_{\text{phys}} | \partial_\mu A^\mu | \psi_{\text{phys}} \rangle = 0$. (The point is that

we need to include this extra symmetry in our theory somehow, and in future lectures we'll see how to quantize field theories in the presence of local symmetries without needing this kind of crutch.)

From here on, we'll set $\xi = 1$ (this is known as the Feynman gauge), so that we have

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial_\mu A^\mu)^2 = -\frac{1}{2}\left[\partial_\mu A_\nu \partial^\mu A^\nu - \partial_\nu A_\mu \partial^\mu A^\nu - \partial_\mu A^\mu \partial_\nu A^\nu\right],$$

and now adding total derivative terms to make these terms look more similar leads us to

$$= \frac{1}{2}A_\nu\left(\partial_\mu\partial^\mu A^\nu - \partial_\mu\partial^\nu A^\mu + \partial^\nu\partial_\mu A^\mu\right) = -\frac{1}{2}A_\nu\partial_\mu\partial^\mu A^\nu = -\frac{1}{2}(\partial_\mu A^\nu)(\partial^\mu A_\nu),$$

so that the associated momentum is

$$\Pi^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\mu)} = -\partial^0 A^\mu.$$

Doing a mode expansion, we can write

$$A^\mu = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\sum_{\lambda=0}^{3}\left[e^{ipx}\varepsilon^\mu_{(\lambda)}(p)a^{(\lambda)}(p) + e^{-ipx}\varepsilon^{\mu *}_{(\lambda)}(p)a^{\dagger(\lambda)}(p)\right]$$

where the $\varepsilon^{(\lambda)}_\mu$s (we can put the (λ) label in the top or bottom, with no difference) are basis vectors in the directions $0, 1, 2, 3$ – these are called **polarization vectors** and we require that

$$p^\mu \cdot \varepsilon^{(0)}_\mu = p^0, \quad p^\mu \cdot \varepsilon^{(3)}_\mu = p^0, \quad p^\mu \cdot \varepsilon^{(1)}_\mu = p^\mu \cdot \varepsilon^{(2)}_\mu = 0.$$

Specifically, $\varepsilon^{(0)}_\mu$ is called the **scalar polarization**, $\varepsilon^{(3)}_\mu$ is the **longitudinal polarization**, and $\varepsilon^{(1,2)}_\mu$ are the **transverse polarizations**. So if we have a massless particle $k^\mu = \begin{bmatrix} k \\ 0 \\ 0 \\ k \end{bmatrix}$, we can parameterize $\varepsilon^{(i)}$s to be the standard basis vectors in the first, second, third, and fourth coordinates. We'll also normalize to avoid dependence on momentum, so that

$$\varepsilon^\mu_{(\lambda)}(p)\varepsilon_{(\lambda'),\mu}(p) = \eta_{\lambda\lambda'}.$$

(We should think about having one field per polarization, so we have a "separate photon" in each of those directions because they're independent particles.) We claim we have the usual commutation relation

$$\boxed{[A^\mu(\vec{x}), \Pi^\mu(\vec{y})] = ig^{\mu\nu}\delta^{(3)}(\vec{x}-\vec{y})}.$$

Indeed, plugging in our mode expansion,

$$[A^\mu(\vec{x}), \Pi^\mu(\vec{y})] = -\int \frac{d^3p}{(2\pi)^3}\frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_q}}\frac{1}{\sqrt{2\omega_p}}\sum_{\lambda,\lambda'}\Bigg[e^{ipx-iqy}\varepsilon^\mu_{(\lambda)}(p)\varepsilon^{*\nu}_{(\lambda')}(q)(-iq^0)[a^{(\lambda)}(p), a^{\dagger(\lambda')}(q)]$$

$$+e^{-ipx+iqy}\varepsilon^{*\mu}_{(\lambda)}(p)\varepsilon^\nu_{(\lambda')}(q)(iq^0)[a^{\dagger(\lambda)}(p), a^{(\lambda')}(q)]\Bigg],$$

and we see that this means we require

$$\boxed{[a^{(\lambda)}(p), a^{\dagger(\lambda')}(q)] = -(2\pi)^3\eta^{\lambda\lambda'}\delta^{(3)}(\vec{q}-\vec{p})},$$

so that the commutator above becomes

$$= -i \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} \sum_{\lambda,\lambda'} -\eta^{\lambda\lambda'} \varepsilon^{\mu}_{(\lambda)}(p) \varepsilon^{*\nu}_{(\lambda')}(p) p^0 \left[e^{i\vec{p}(\vec{x}-\vec{y})} + e^{-i\vec{p}(\vec{x}-\vec{y})} \right],$$

which is indeed $ig^{\mu\nu}\delta^{(3)}(\vec{x}-\vec{y})$ as we wish because $\sum_{\lambda,\lambda'} -\eta^{\lambda\lambda'}\varepsilon^{\mu}_{(\lambda)}(p)\varepsilon^{*\nu}_{(\lambda')}(p) = -g^{\mu\nu}$. (Here remember η and g are the same.)

We can now write down the Hamiltonian by finding the Noether current: we have

$$\begin{aligned} j^{\mu\nu} &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\sigma)} \partial^\nu A_\sigma - g^{\mu\nu}\mathcal{L} \\ &= -\partial^| mu A_\sigma \partial^| nu A_\sigma + q^{\mu\nu} \frac{1}{2} (\partial_\mu A^\sigma)(\partial^\mu A_\sigma), \end{aligned}$$

so that

$$H = \int d^3x j^{00} = \int d^3x \left[-\frac{1}{2}\Pi^\mu \Pi_\mu - \frac{1}{2}\partial_i A^\sigma \partial^i A_\sigma \right].$$

Plugging in mode expansions and doing some lengthy calculations, we find that

$$H = \frac{1}{2} \int \frac{d^3p}{(2\pi)^3} \omega_p \sum_{\lambda,\lambda'=0}^{3} -\eta_{\lambda,\lambda'} \left[a^{(\lambda)}(p) a^{\dagger(\lambda')}(p) + a^{\dagger(\lambda)}(p) a^{(\lambda')}(p) \right],$$

and as usual we can rewrite this in normal ordering:

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_p \left[\sum_{\lambda=1}^{3} a^\dagger_{(\lambda)} a_{(\lambda)} - a^\dagger_{(0)} a_{(0)} \right] + \text{vacuum energy}$$

where we ignore the last term. But this should be concerning, because now we are enumerating the modes of polarization zero and those can contribute negative energy. But this is where the Gupta-Blueler condition comes into play: if we want

$$0 = \langle \psi_{\text{phys}} | \partial_\mu A^\mu | \psi_{\text{phys}} \rangle,$$

then plugging in our mode expansion shows that we are requiring (here we are using k instead of p for momentum)

$$0 = \sum_{\lambda=0}^{3} k^\mu \varepsilon^{(\lambda)}_\mu(k) a_{(\lambda)}(k) \left| \psi_{\text{phys}} \right\rangle.$$

But we already imposed the condition on polarization earlier that $\varepsilon^{\mu(1,2)}(k)k_\mu = 0$, and $\varepsilon^{\mu(0)}k_\mu = k^0$ and $\varepsilon^{\mu(3)}(k)k_\mu = -k^0$. Thus we are really requiring that

$$k^0(a^{(0)}(k) - a^{(3)}(k)) \left| \psi_{\text{phys}} \right\rangle = 0 \implies \left\langle \psi_{\text{phys}} \middle| a^{\dagger(0)} a^{(0)} \middle| \psi_{\text{phys}} \right\rangle = \left\langle \psi_{\text{phys}} \middle| a^{\dagger(3)} a^{(3)} \middle| \psi_{\text{phys}} \right\rangle.$$

So plugging this in we see that

$$\langle \psi_{\text{phys}} | :H: | \psi_{\text{phys}} \rangle = \left\langle \psi_{\text{phys}} \middle| \int \frac{d^3k}{(2\pi)^3} k^0 \sum_{\lambda=1}^{2} a^{\dagger(\lambda)} a^{(\lambda)} \middle| \psi_{\text{phys}} \right\rangle,$$

and the physical state of observables now has well-defined time-evolution and positive Hamiltonian – the other part of the state space is just coming from the additional symmetry and the way that we set up the problem.

We can now write down our Fock space by writing down particle states (this means we have a particle with momentum k and polarization λ)

$$|(k,\lambda)\rangle = -\sqrt{2E_k} a^{(\lambda)\dagger}(k) |0\rangle$$

and also writing down two-particle states

$$|(k_1,\lambda),(k_2,\lambda)\rangle = \sqrt{2E_{k_1}}\sqrt{2E_{k_2}}a^{\dagger(\lambda)}(k_1)a^{\dagger(\lambda')}(k_2)\,|0\rangle\,.$$

For $k_1 \neq k_2$ we can flip their roles and interchange the particles, so our states are bosons. We can also create twice the same particle

$$\frac{1}{\sqrt{2}}(2E_k)(a^{\dagger(\lambda)}(k))^2\,|0\rangle = |2(k,\lambda)\rangle$$

and there is no Pauli exclusion principle (we can put in the same particle multiple times); generalizing, we get n-particle states satisfying the equations

$$\sqrt{2E_{k_i}}a^{\dagger(\lambda_i)}(k_i)\,|n_1(k_1,\lambda_1),\cdots,n_i(k_i,\lambda_i)\rangle = \sqrt{n_i+1}\,|n_1(k_1,\lambda_1),\cdots,(n_i+1)(k_i,\lambda_i),\cdots\rangle\,,$$

$$\frac{1}{\sqrt{2E_{k_i}}}a^{(\lambda_i)}(k_i)\,|n_1(k_1,\lambda_1),\cdots,n_i(k_i,\lambda_i)\rangle = \sqrt{n_i}\,|n_1(k_1,\lambda_1),\cdots,(n_i-1)(k_i,\lambda_i),\cdots\rangle\,.$$

We should think of annihilation as **absorption** (reducing the number of photons by one gives us proportionality to $\sqrt{n}$), so the absorption probability is zero if there are no photons. On the other hand, creation is emission and is proportional to $\sqrt{1+n}$ (**stimulated emission**), so we can emit photons even if $n=0$ (this is called **spontaneous emission**). So in a cavity, having more and more photons makes it easier to emit more photons and that's how lasers work.

So now to solve our theory, we can commute the non-equal-time commutator

$$[A_\mu(x),A_\nu(y)] = -g_{\mu\nu}\int\frac{d^3k}{(2\pi)^3}\frac{1}{2k^0}\left[e^{-ik(x-y)}-e^{ik(x-y)}\right] = -g^{\mu\nu}\left[D(x-y)-D(y-x)\right],$$

so that the Feynman propagator is

$$D_F^{\mu\nu}(x-y) = \langle 0|T\{A^\mu(x)A^\nu(y)\}|0\rangle = \int\frac{d^4k}{(2\pi)^4}e^{-ik(x-y)}\frac{-ig^{\mu\nu}}{k^2+i0},$$

which is the Green's function of Maxwell's theory:

$$\Box D_F^{\mu\nu}(x-y) = ig^{\mu\nu}\delta^{(4)}(x-y).$$

So we can calculate the time-ordered expectation value just like for scalar field theory, with the only complication being sums over polarizations. Then the Feynman propagator is basically the same as the massless scalar Feynman propagator times $-g^{\mu\nu}$. (And we're assuming that photons are massless, and gauge theory looks naively wildly violated if we have nonzero mass, so we won't get into that here.)

So summarizing everything, we see that we have the Dirac Lagrangian $\overline{\psi}(i\not{\partial}-m)\psi$ and the electromagnetism Lagrangian $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}-\frac{1}{2}(\partial_\mu A^\mu)^2$, and now we will just add an **interacting Lagrangian** $-e\overline{\psi}\gamma^\mu\psi A_\mu = -e\overline{\psi}\not{A}\psi$ (where $e=\sqrt{4\pi\alpha}$ and α is the fine structure constant). The overall **quantum electrodynamics** Lagrangian is then the sum of these three terms.

Writing down this particular interaction term is not super motivated at first, but we'll see that it has the properties that we want. We can then calculate the S-matrix elements:

$$S = U(T,-T) = \sum_{n=0}^{\infty}\frac{1}{n!}(-ie)^n T\int d^4x_1\cdots d^4x_n\overline{\psi}(x_1)\not{A}(x_1)\psi(x_1)\cdots\overline{\psi}(x_n)\not{A}(x_n)\psi(x_n).$$

We wish to calculate the overlap $S_{fi}=\langle f|S|i\rangle$, which we still claim (but haven't proven) is ${}_0\langle f|U(T,-T)|i\rangle_0$ where

we only sum over connected, amputated diagrams but compute in the free theory states. We write down

$$\psi^\dagger(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2k^0}}\sum_{s=\pm 1/2} v_s e^{ipx} b_s^\dagger(k)$$

for creating a positron at x, and similarly $\overline{\psi}^\dagger$ creates an electron, ψ^- absorbs an electron, and $\overline{\psi}^-$ absorbs a positron. We also have that A^+ and A^- create and absorb photons. We can represent these all in Feynman diagrams, with positrons and electrons being emitted or absorbed based on outgoing or incoming vertices, and photons represented with wavy lines. Wick's theorem tells us that we can contract $\psi(x)\overline{\psi}_y = S_F(x-y)$, and we similarly see here that $A^\mu(x)A^\nu(y) = D_F^{\mu\nu}(x-y)$. So in the Feynman diagram formulation, we claim that $-ie\overline{\psi}\not{A}\psi$ corresponds to a vertex with a photon and an incoming and outgoing arrow. Then the first-order S-matrix is

$$S^{i0} = -ie\int d^4x T\{\overline{\psi}(x)\not{A}(x)\overline{\psi}(x)\} = -ie\int d^4x :\overline{\psi}(x)\not{A}(x)\psi(x): + \cdots.$$

The normal-ordered product basically means that we can "direct time" in a variety of ways at vertices – we have three-point interactions such as a photon creating a positron and an electron. But

$$(p_{e^+} + p_{e^-})^2 = p_\gamma^2 = 0 = 2m_e^2 + 2p_{e^-}p_{e^+} = 2m_e^2 + 2E_{e^-}2E_{e^+} - 2|\vec{p}_{e^-}||\vec{p}_{e^+}|\cos\theta,$$

but that cannot happen and thus there is no three-particle scattering at this first order. So we have to expand the S-matrix some more to get the interactions to show up, and we'll see this more next time.

15 November 17, 2022

We introduced quantum electrodynamics last time, writing down the QED Lagrangian

$$\mathcal{L}_{\text{QED}} = \overline{\psi}(i\not{\partial} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial_\mu A^\mu)^2 - e\overline{\psi}\not{A}\psi,$$

where the first term corresponds to the free Dirac equation (fermions – electrons and positrons), the next two describe free electromagnetism and photons, and the last term describes interactions between them. We found that the Feynman propagators look like (these are contractions os $\psi\overline{\psi}$ and $A^\mu A^\nu$, respectively):

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}\frac{i(\not{p}+m)}{p^2-m^2+i0}, \quad D_F^{\mu\nu}(x-y) = \int \frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}\frac{-iq^{\mu\nu}}{p^2+i0}.$$

We want to do calculations in perturbation theory, since we can relate the electric charge e to the fine structure constant and find that e is quite small. So we can do perturbation theory to study scattering processes; using all the tools we found that the overlap of the S-matrix S_{fi} can be calculated in terms of the the time-ordered product of the interacting Lagrangians, integrating over n points at nth order. But then we found that there's no three-particle on-shell scattering ($\gamma = e^+ + e^-$) by kinematic momentum calculations at first order.

Example 30

We'll study **Compton scattering** today, which is the idea of shining a photon on an electron at rest and seeing what happens. We'll generalize this to a process of the form $\gamma + e^- \to \gamma + e^-$, where the initial and final momentum and polarization of the photon are $p_1, \lambda \to p_3, \lambda'$, and the initial and final momentum and spin of the electron are $p_3, s \to p_2, s'$.

Our initial state is thus

$$|i\rangle = \left|e^-(p_2,s)\gamma(p_1,\lambda)\right\rangle = \sqrt{2E_1 2E_2}a^\dagger_{(\lambda)}(p_1)a^\dagger_s(p_2)\left|0\right\rangle,$$

and similarly we have

$$|f\rangle = \sqrt{2E_3E_4}a^\dagger_{(\lambda')}(p_3)a^\dagger_{s'}(p_4)\left|0\right\rangle.$$

We thus want to find $S_{fi} = {}_0\langle f|S|i\rangle_0$ to help calculate scattering amplitude, and we care about the nontrivial part where something actually happened to the electron and photon. It'll thus look like $\delta_{if} + i(2\pi)\delta^{(4)}(p_i - p_f)M$ for some constant M which we want to find, and the first time we will actually see a nontrivial term like that is if we expand to second order (first order vanishes as we saw last time). So we want to calculate the term

$$\frac{1}{2}(-ie)^2\sqrt{16E_1E_2E_3E_4}\left\langle 0\middle|a_{(s')}(p_4)a_{(\lambda)}(p_3)\int dxdyT\left\{\overline{\psi}_x\not{A}_x\psi_x\overline{\psi}_y\not{A}_y\psi_y\right\}a^\dagger_{(s)}(p_2)a^\dagger_{(\lambda)}(p_1)\middle|0\right\rangle.$$

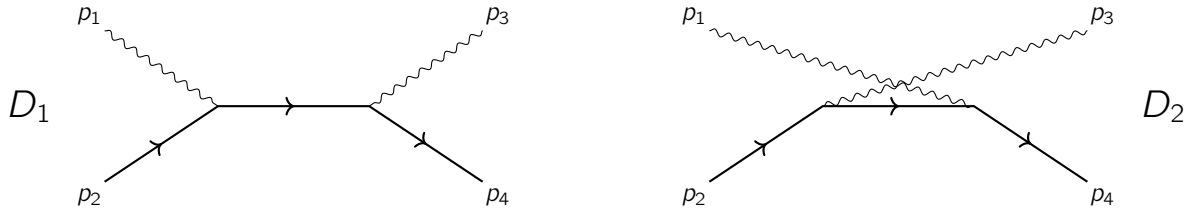
We're only getting something nonzero when we contract all fields together, and we can only contract operators of the same species of particle. For example, $a_{(s')}(p_4)$ could be contracted with $a^\dagger_s(p_2)$, and $a_{(\lambda)}(p_3)$ could be contracted with $a^\dagger_{(\lambda')}(p_1)$, and then we can contract $\not{A}_x\not{A}_y$, $\overline{\psi}_x\psi_x$, and $\overline{\psi}_y\psi_y$. But that's really just free propagation of the photon going to itself and the electron going to itself, so it shouldn't be included (the diagram is not connected, and we just add this to the overall vacuum energy).

Next, we can contract $a_{(s')}(p_3)$ with $\overline{\psi}_x$ and ψ_x with $a^\dagger_{(s)}(p_2)$, $\overline{\psi}_y$ with ψ_y, and then connect $a_{(\lambda)}(p_3)$ with $a^\dagger_{(\lambda')}(p_1)$ and contract the fields $\not{A}_x, \not{A}_y$. (So at x, we create a photon, which connects to a spacetime point y where a fermion line is connected to itself.) But this falls into the "amputated" consideration – the propagator is coming from outside, and again we don't want to count this case.

So now we can turn to something more nontrivial – we're learning that **whatever we have on the outside should be contracted with something we have in our S-matrix**. So we take our fermion $a_{s'}(p_4)$ and contract it with $\overline{\psi}_x$, and we take our photon $a_{(\lambda)}(p_3)$ and contract it with $\not{A}_x$. We can contract ψ_x and $\overline{\psi}_y$ in the middle, and finally contract $\not{A}_y$ with $a^\dagger_{(\lambda)}(p_1)$ and ψ_y with $a^\dagger_{(s)}(p_2)$. Similarly we can do the same but have $a_{(\lambda)}(p_3)$ connect to $\not{A}_y$ instead of $\not{A}_x$ and have $a^\dagger_{(\lambda)}(p_1)$ connect to $\not{A}_x$. (In fact there are two of each instance giving the same Feynman diagram, which compensates for the $\frac{1}{2}$ factor in front – this is down to being able to exchange x and y without changing anything else.) We could then do a lot of calculations, but it's a fairly tedious process, and it's easier to do these calculations with Feynman rules:

- Represent electrons going in with arrows in the upper right direction, corresponding to a $u_s(p)$ spinor term, and represent positrons going in in the lower left direction, corresponding to a $\overline{v}_s(p)$ term. Similarly an electron going out is a $\overline{u}_s(p)$ term and a positron going out is a $v_s(p)$ term.
- A photon coming in is a $\varepsilon_\mu(p)$ term (corresponding to a wavy line entering a node on the right), and a photon going out is a $\varepsilon^*_\mu(p)$ term.
- Propagators of fermions correspond to $\frac{i(\not{p}+m)}{p^2-m^2+i0}$ terms, and propagators of photons correspond to $-\frac{iq^{\mu\nu}}{p^2+i0}$.
- At every vertex we have a $-ie\gamma^\mu\delta_{ss'}$ term corresponding to a three-particle interaction.
- Momentum conservation must hold at every vertex.
- Closed loops correspond to a $\int d^4p$ integral.
- We get a (-1) factor for each closed fermion loop $\overline{\psi}_1\psi_1\overline{\psi}_2\psi_2 = (-1)\text{tr}(\psi_1\overline{\psi}_1\psi_2\overline{\psi}_2) = (-1)\text{tr}(S_F(x_1 - x_2)S_F(x_2 - x_1))$.
- Orientation of arrows matters, and they must run in a consistent direction around loops.

So in Compton scattering, we have γ and e^- coming in and γ and e^- coming out, and there are basically two potential diagrams to consider:



There are no negative signs or other issues here, so we can just use our graphical Feynman rules. We always start with a fermion line and look at an outgoing arrow to write down what we get from each: for the first diagram D_1 we get (here we require $p_3 + p_4 = p_{34}$ to be the momentum at the right vertex, and $p_1 + p_2 = p_{34}$ by the conservation at the left vertex)

$$= \overline{u}_{s'}(p_4) \cdot (-ie)\gamma^\mu \cdot \frac{i(\not{p}_{34} + m)}{p_{34}^2 - m^2 + i0}(-ie)\gamma^\nu u_s(p_2) \cdot \varepsilon^{*\mu}_{(\lambda')}(p_3)\varepsilon^\nu_{(\lambda)}(p_1).$$

(So we've "gone off-shell" in the middle of the diagram so that momentum is conserved, but $(p_3+p_4)^2 \neq 0$ in general.) Similarly, in the second diagram D_2 we get

$$\overline{u}_{s'}(p_4)(-ie)\gamma^\nu \frac{i\not{p}_{14} + m}{p_{14}^2 - m^2 + i0}(-ie)\gamma^\mu u_s(p_2)\varepsilon^{*\mu}_{\lambda'}(p_3)\varepsilon^\nu_\lambda(p_1).$$

So this process of writing down all diagrams and using the Feynman rules gets us the total scattering amplitude more easily. (This is really the Born approximation, since we're only looking at tree-level diagrams.) And this tells us that

$$M = D_1 + D_2 \implies \sigma = \frac{1}{2\sqrt{(2p_1p_2)^2 - 4m_1^2m_2^2}} \frac{1}{N_{s_1}} \frac{1}{N_{s_2}} \int \frac{d^4p_3}{(2\pi)^4} \frac{d^4p_4}{(2\pi)^4}$$

$$2\pi\delta_+(p_3^2 - m^2)2\pi\delta_+(p_4^2 - m^2)(2\pi)^4\delta^{(4)}(p_1 + p_2 - p_3 - p_4)|\overline{M}|^2,$$

where we'll discuss what the bar over M means soon. (We derived this for scalar scattering early on, and it's the same idea here.) Here note that $m_1 = 0, m_2 = m_e, m_3 = 0, m_4 = m_e$, so we really need to have $p_3^2 = 0$, and $N_{s_1}N_{s_2}$ correspond to the number of spin states for particles of the two types (1 for scalar, 2 for photons because of polarization, and 2 for fermions because of spin-up and spin-down – we'd have something bigger if we had higher spins.) And the point is that $\overline{M}$ is the matrix element in scattering over all final states, but we also want to **sum over all initial polarizations**, but then we have to average out (divide by the total) because we just have a single state when we start. So we should take the average of our scattering probabilities if it (for example) doesn't matter what the initial polarization actually is, which is why we have the factors of N_{s_1} and N_{s_2}. This means

$$|\overline{M}|^2 = \sum_{\lambda,\lambda'} \sum_{s,s'} |M|^2,$$

where $|M|^2 = D_1D_1^\dagger + D_1D_2^\dagger + D_1^\dagger D_2 + D_2D_2^\dagger$ and

$$D_1 = -\frac{ie^2}{s - m^2}\overline{u}_4\gamma^\mu(\not{p}_{12} + m)\gamma^\nu u_2\varepsilon_\mu\varepsilon_\nu^* \implies D_1^\dagger = \frac{ie^2}{s - m^2}u_2^\dagger\gamma^{\nu\dagger}(\not{p}_{12} + m)\gamma^{\dagger\mu}\gamma^0 u_4\varepsilon_\mu^*\varepsilon_\nu.$$

But now remember that $\gamma^0\gamma^0 = 1$, we can insert γ^0 into various places and use that $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$ to write

$$D_1^\dagger = \frac{ie^2}{s - m^2}\overline{u}_2\gamma^\nu(\not{p}_{12} + m)\gamma^\mu u_4\varepsilon_\mu^*\varepsilon_\nu.$$

So we can calculate the first term in the matrix element $|M|^2$ now:

$$\sum_{\lambda,\lambda'}\sum_{s,s'} D_1 D_1^\dagger = \frac{e^4}{(s-m^2)^2}\sum_{\lambda,\lambda'}\varepsilon_\mu(p_3)\varepsilon_\rho^*(p_3)\varepsilon_\nu^*(p_1)\varepsilon_\sigma(p_1)$$

$$\sum_{s,s'}\overline{u}_{\psi,s}\gamma^\mu(\not{p}_{12}+m)\gamma^\nu u_{r,s}\overline{u}_{r,s}\gamma^\sigma(\not{p}_{12}+m)\gamma_\rho u_{\psi,s'}).$$

We can make use of the identities now that

$$\sum_\lambda \varepsilon_\lambda^{*\mu}(p_1)\varepsilon_\lambda^\nu(p_1) = -g^{\mu\nu}, \quad \sum_s u_{r,s}\overline{u}_{r,s} = \not{p}_2 + m, \quad \sum_{s'} u_{\psi,s}\overline{u}_{\psi,s'} = \not{p}_4 + m,$$

so we can contract Lorentz indices between γ matrices and end up with

$$= \frac{e^4}{(s-m^2)^2}\mathrm{tr}\left[(\not{p}_4+m)\gamma^\mu(\not{p}_{12}+m)\gamma^\nu(\not{p}_2+m)\gamma_\nu(\not{p}_{12}+m)\gamma_\mu\right].$$

One of the things we can do for simplification here is to set all masses to zero to make our life easier (for example high-energy scattering), so that we want to calculate $\frac{e^4}{s^2}\mathrm{tr}\left[\not{p}_4\gamma^\mu\not{p}_{12}\gamma^\nu\not{p}_2\gamma_\nu\not{p}_{12}\gamma^\mu\right]$. Then $\gamma^\mu\gamma_\mu = 4I$ and $\gamma^\mu\not{p}\gamma_\mu = -2\not{p}$, and remembering that $S = (p_1+p_2)^2 = p_{12}^2$ is the center-of-mass energy, we have

$$D_1 D_1^\dagger = \frac{d^4}{s^2}4\mathrm{tr}[\not{p}_4\not{p}_1\not{p}_2\not{p}_1] = -\frac{e^4}{s^2}8su,$$

where $u = (p_1+p_4)^2 = p_{14}^2$. We then find similarly that

$$D_1 D_2^\dagger = D_1^\dagger d D_2 = 0, \quad D_2 D_{\textcircled{e}}^\dagger = -\frac{e^4}{u^2}8su,$$

so that $|\vec{M}|^2 = -\frac{e^4}{su}(s^2+u^2)$. Plugging this back in and integrating out all the delta functions, we find that the scattering cross-section is

$$\sigma = \frac{1}{64\pi^2 s}\int d\Omega_3 - \frac{e^4(s^2+u^2)}{su}$$

where Ω_3 is the angular element in three dimensions. Plugging in $p_2\cdot p_4 = 2E_2E_3(1-\cos\theta)$, we find that

$$-\frac{u^2+s^2}{su} = \frac{1+\cos^2\theta}{\cos\theta},$$

so the differential cross-section for no mass is

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s}\frac{1+\cos^2\theta}{\theta}.$$

On the other hand, if we had kept the mass term we would have found a similar result with additional corrections:

$$|\overline{M}|^2 = 8e^4\left[\frac{m^2-u}{s-m^2} + \frac{m^2-s}{u-m^2} + 4\left(\frac{m^2}{s-m^2}+\frac{m^2}{u-m^2}\right) + 4\left(\frac{m^2}{s-m^2}+\frac{m^2}{u-m^2}\right)^2\right].$$

16 November 29, 2022

We've now discussed a few different quantum field theories with different Lagrangians: looking first at the interaction-free case, we have the spin-0 $\mathcal{L}_0^{\text{scalar}} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2$, the spin-$\frac{1}{2}$ $\mathcal{L}_0^{\text{Dirac}} = \overline{\psi}(i\not{\partial} - m)\psi$, and the spin-1 $\mathcal{L}_0^{\text{Maxwell}} = \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial_\mu A^\mu)^2$. We then calculated the two-point correlation functions for each of those cases, which are represented with arrows in Feynman diagrams, to be $\frac{i}{p^2-m^2+i0}$, $\frac{i(\not{p}+m)}{p^2-m^2+i0)}$, and $-\frac{iq^{\mu\nu}}{p^2+i0}$. (Higher spins can arise, but in

the standard model we just have these three cases.) Introducing interactions means that we have to introduce new local operators: for quantum electrodynamics we add a $-ie\overline{\psi}\not{A}\psi$ term to the Dirac and Maxwell Lagrangians, and to get the Higgs boson and τ-leptons we add a $-\frac{m_\tau}{v}h\overline{\psi}\psi$ term instead.

We then computed the S-matrix using perturbation theory as a sum over amputated, connected Feynman diagrams, in which we associate particular pictures (vertices in quantum electrodynamics, for example) with a correspponding mathematical expression, which allows us to write down expressions for probabilities of particular scattering processes.

Everything so far has been classical – our next step is to figure out quantum corrections, which come up when we expand further in perturbation theory. (Remember that there have been factors of $\hbar$ in our expansions of time-ordered series, and taking $\hbar \to 0$ is recovering the classical limit.) So we'll be discussing **renormalization** throughout the rest of this course. We'll be looking at our ϕ^4 Lagrangian for simplicity

$$\mathcal{L}_{\phi^4} = \mathcal{L}_0^{\text{scalar}} - \frac{\lambda}{4!}\phi^4,$$

and the idea is that we want to look at Feynman diagrams with loops, which give us expressions that look like $\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i0}$ (if we attach a loop of momentum k to a segment with endpoints p_1 – this is called a tadpole graph) or $\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i0}\frac{i}{(k+p_1+p_2)^2-m^2+i0}$ (if we have incoming and outgoing segments of momentum p_1, p_2 on both sides, and the momentum is k and $k+p_1+p_2$ on the two sides of our loop). And the point is that as $k \gg m, p_1, p_2$, the tadpole graph contribution approximately becomes $\int \frac{d^4k}{k^2} \sim \int_0^\infty \frac{d|k||k|^3}{|k|^2}\int d\Omega^3 \sim |k|^2$, and similarly the other expression diverges as $\int \frac{d^4k}{k^4} \sim \log(k)$. So we get quadratic and logarithmic divergence respectively, and we'll now discuss how bad that really is for probability calculations.

Suppose we have a Feynman diagram with n vertices, E external lines, I internal lines, and L loops. (So in the tadpole graph above, we have $(n, E, I, L) = (1, 2, 1, 1)$.) By counting graph degree, for the ϕ^4 theory, we must have $4n = E + 2I$. If we then work in d spacetime dimensions, we now have $\int d^dk$ integrals, so the integration measure of any diagram goes as $d \cdot L$ (Since we get an $\int d^dk$ for each loop)

Definition 31

The **degree of divergence** of a diagram is $D = d \cdot L - 2I$ (where we're basically counting powers of k).

For example, we can check that $D = 2$ and $D = 0$ in the two cases above. It turns out we also have $L = I - (n-1)$ because of the requirements of the momentum conservation at each vertex. And this is good, because it means we can write everything in terms of n and E (which are actually the values that are relevant in a physical situation where we are looking for scattering to occur):

$$D = d - \left(\frac{d}{2} - 1\right)E + n(d-4)$$

and in particular this is $4 - E$ when $d = 4$, which is good because we do not make our diagrams worse as we add more and more vertices. (This doesn't mean there will be no divergence if E is large – this is a superficial degree just looking at what happens to k, and it's possible that we have subdiagrams that diverge and are only isolated to specific parts.)

So we can speak of **renormalizable** quantum field theories (in which D is independent of n, for example if $d = 4$ in our specific theory above), **super renormalizable** quantum field theories (in which D decreases with n, for example if $d = 2$), and **non-renormalizable** quantum field theories (where D is increases with n, for example if $d = 6$). On the other hand, gravity in $d = 4$ turns out to be non-renormalizable – it's one of the reasons why we think we don't have a quantum theory of gravity.

What we want to talk about is what happens in QED. There, we have n vertices, p_e external γ photon lines, p_i internal γs, E_e external electron lines, E_i internal electron lines, and L loops. The same kind of graph-theoretic

considerations (remembering that at each vertex we have one photon line and two electron lines) leads us to $n = p_e + 2p_i$ and $2n = E_e + 2E_i$. Then because the photon propagator goes essentially as $\frac{1}{p^2}$ and the electron one goes as $\frac{\not{p}}{p^2} \sim \frac{1}{p}$, we have the superficial degree of divergence

$$D = d \cdot L - E_i - 2p_i = d + n \cdot \frac{d-4}{2} - \frac{d-1}{2} E_e - \frac{d-2}{2} p_e,$$

which is $4 - \frac{3}{2}E_e - p_e$ for $d = 4$. (So in particular, this means that QED is renormalizable.) The divergent diagrams are those where we basically have "**electron self-energies**," meaning that we have internal photon lines branching off of a single electron propagator (so $D = 4 - 2 \cdot \frac{3}{2} = 1$), or where we have "**photon self-energies**," where $D = 4 - 2 = 2$. We also have the case where we have "**vertex corrections**" (for instance, taking a usual interaction vertex and connecting the two electron lines with a photon), where $D = 4 - \frac{3}{2} \cdot 2 - 1 = 0$ (logarithmic). And of course, these are all divergent if they are embedded into a larger Feynman diagram as well. There are other diagrams as well, but they will cancel out when we implement them into larger diagrams.

So it's just the three types above that we need to worry about, and our goal now is to do something to deal with them. We'll do this using something called **regularization** – one method is to perform cutoff regularization, where if we have some loop integration $\int_0^\infty \frac{d|p|}{|p|}$, we can define that to be $\lim_{\Lambda \to \infty} \int_0^\Lambda \frac{d|p|}{|p|}$ and rearrange all the diagrams so that the Λ-dependence actually goes away. But the common technique these days is to use **dimensional regularization** and say that we're working in $d = 4 - 2\varepsilon$ dimensions, taking $\varepsilon \to 0$.

Example 32

With this method, our tadpole graph contribution is now $-\frac{i\lambda}{2} \int \frac{d^d k}{(2\pi)^d} \frac{i}{k^2 - m^2 + i0}$.

In order to turn this into a Euclidean integral, we do a **Wick rotation** where $k^0 = ik_E^0$, so that our poles in the k^0 plane are now rotated by 90 degrees. Then $k^2 = (k^0)^2 - \vec{k}^2$ is now $-k_E^2$ (so that we don't have the complicated Minkowski metric with a minus sign). Our integral then becomes

$$\frac{i\lambda}{2} \int \frac{d^d k_E}{(2\pi)^d} \frac{1}{-k_E^2 - m^2 + i0};$$

introducing spherical coordinates (this is the equivalent of $d^3x = r^2 dr d\Omega$ in three dimensions) this becomes

$$= \frac{i\lambda}{2(2\pi)^d} \int_0^\infty d|k_E| |k_E|^{d-1} \int d\Omega_d \frac{1}{-k_E^2 - m^2 + i0}.$$

To compute this, we will use the useful identity

$$\int_0^\infty \frac{d|k_E|^2}{2} \frac{(k_E^2)^{d/2-1}}{k_E^2 + \Delta} = (\Delta)^{d/2-1} \Gamma\left(\frac{d}{2} - 1\right) \Gamma\left(2 - \frac{d}{2}\right),$$

where Γ is the usual gamma function (related to the factorial) with $\Gamma(n+1) = n!$ and $\Gamma(d+1) = d\Gamma(d)$ for all d. This function turns out to help us with calculating the area of a sphere in d dimensions – remembering that $\sqrt{\pi} = \int dx e^{-x^2}$, raising both sides to the dth power tells us that

$$\pi^{d/2} = \int d^d x e^{-\sum_{i=1}^d x_i^2} = \int d\Omega_d \int_0^\infty d|\vec{x}| |\vec{x}|^{d-1} e^{-|\vec{x}|^2}$$

where again we've split into spherical coordinates. But this last expression is

$$\int d\Omega_d \int_0^\infty \frac{d(|\vec{x}|)^2}{2} (|\vec{x}|^2)^{(d-1)/2} e^{-|\vec{x}|^2} = \int d\Omega_d \frac{1}{2} \Gamma\left(\frac{d}{2}\right),$$

so we find that $\boxed{\Omega_d = \frac{2(\sqrt{\pi})^d}{\Gamma\left(\frac{d}{2}\right)}}$. So plugging everything back in, we find that the contribution from our tadpole graph is

$$\text{contribution}_{\text{tadpole}} = \frac{-i\lambda\pi^{d/2}}{2(2\pi)^d\Gamma(\frac{d}{2})}\Delta^{\frac{d}{2}-1}\Gamma\left(1-\frac{d}{2}\right)\Gamma\left(2-\frac{d}{2}\right),$$

and now if we plug in $\Delta = m^2$ (as it is in our particular integral) and $d = 4-2\varepsilon$ we get

$$\Omega = -\frac{i\lambda}{2(4\pi)^2}\left(\frac{m^2}{4\pi}\right)^{-\varepsilon} m^2\frac{\Gamma(1-\varepsilon)}{\varepsilon}.$$

We can now use the series expansion of Γ as

$$\Gamma(1+\varepsilon) = e^{-\varepsilon\gamma_E}\left(1+\frac{\zeta_2}{2}\varepsilon^2-\frac{1}{3}\varepsilon^3+\cdots\right)$$

where γ_E is the Euler-Mascheroni constant (around 0.577) and ζ_n is in terms of the Riemann zeta function: $\zeta_n = \sum_{k=1}^{\infty} k^{-n}$. So at leading order we see that

$$\Omega = -\frac{im^2\lambda}{2(4\pi)^2}\frac{1}{\varepsilon},$$

and thus the divergence shows up as a **pole** in $\frac{1}{\varepsilon}$. So we can now calculate in this fractional dimension because we've managed to make the divergence manifest. (And also, in dimensional regularization, we'll set $\int d^dk\frac{1}{k^2} = 0$ because there's no "relevant scale.") So the idea of renormalization in general is the following: we start with a Lagrangian $\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4$, where we can say that λ can be the λ_R in experiments times a factor Z_Λ, and the fiedl ϕ can be ϕ_R from experiments times some factor $\sqrt{Z_\phi}$, and m^2 is then $m_R^2 Z_m$. The point is that these Zs will encode the divergences, and then redefining our Lagrangian in terms of the **renormalized** constants makes the singularities go away. So we now have

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

and the correlation functions in vacuum (in momentum space) now become $\frac{iZ_\phi}{p^2-m_R^2Z_m+i0}$ instead of $\frac{i}{p^2-m^2+i0}$. We can now do a perturbative expansion in terms of our coupling constant: say $Z_i = 1+\lambda_R\delta_i^{(1)}+\lambda_R^2\delta_i^{(2)}+\cdots$ for each i. Then if we want to calculate all possible interaction terms occurring from a single line segment, we get the tree-level two-point function, plus the one-loop correlation function, plus order λ^2 terms, which now looks like

$$\frac{iZ_\phi}{p^2-m_R^2Z_m+i0}+\frac{iZ_\phi}{p^2-m_R^2Z_m+i0}\left(\frac{\lambda}{2(4\pi)^2}\left(\frac{m^2}{4\pi}\right)^{-\varepsilon}\frac{\Gamma(1+\varepsilon}{\varepsilon}\right)\frac{iZ_\phi}{p^2-m_R^2Z_m+i0}+O(\lambda^2).$$

This can be broken up in terms of our perturbative factors into

$$=\frac{i}{p^2-m_R^2+i0}\left(1+\lambda_R\delta_\phi^{(1)}+\frac{i\delta_m^{(1)}m_R^2}{p^2-m_R^2+i0}-\frac{i\lambda_Rm_R^2}{2(4\pi)^2}\left(\frac{m_R^2}{4\pi}\right)^{-\varepsilon}\frac{\Gamma(1+\varepsilon)}{\varepsilon}\frac{i}{p^2-m_R^2+i0}+O(\lambda^2)\right),$$

and now if we **choose** $\delta_\phi^{(1)} = 0$, $\delta_m^{(1)} = \frac{i\lambda_R}{(2(4\pi)^2}\left(\frac{m_R^2}{4\pi}\right)^{-\varepsilon}\frac{\Gamma(1+\varepsilon)}{\varepsilon}$, then we find that the renormalized two-point correlation function is the renormalized tree-level diagram plus $O(\lambda^2)$, which is just $\frac{i}{p^2-m_R^2+i0}$. So we've removed the tadpole diagram contribution from our calculations entirely by choosing the appropriate δ coefficients. And in all calculations, we can use this kind of redefinition universally – we'll get the same δ because the same singularity appearing in the two-point function can appear in more complicated diagrams as well, and it turns out we will only ever have a finite number of such problems that we need to remove. We'll take a more systematic approach to all of this next time!

17 December 1, 2022

Last time, we started discussing divergences and renormalization – we saw that certain diagrams with loops give us k-dependence like k^2 or $\log k$. To characterize the degree of divergence (at least at a superficial level), we define for any diagram a number depending on the spacetime dimension, number of vertices, and the number of external legs. For example for the ϕ^4 theory, we found that $D_{\phi^4} = d - \left(\frac{d}{2} - 1\right) E + n(d-4) = 4 - E$, so at least the theory is normalizable (D is independent of n). Similarly for QED, we found that $D_{\text{QED}} = 4 - \frac{3}{2}E_e - p_e$. The way we deal with these divergences is with dimensional regularization, where we consider our spacetime dimension as $d = 4 - 2\varepsilon$ and take $\varepsilon \to 0$. Then the renormalization corresponds to redefining couplings, masses, and fields with power series, perturbatively expanding with coefficients $\delta_i^{(j)}$, which will be "counter terms" to absorb the singularities that we see in our problems. (Indeed, we saw last time that choosing a specific $\delta_\phi^{(1)}$ and $\delta_m^{(1)}$) allow us to avoid the tadpole diagram in our renormalized two-point correlation function.) And the constants like λ_R in which we are doing perturbation theory can be actually physically measured by looking at scattering cross-sections and comparing to real experiments. But in most quantum field theories we cannot get exact formulas for these expansions (calculations usually use around 5 or 6 loops for things like QED or the ϕ^4 theory).

Today, we'll think about QED in d spacetime dimensions. We have $S = \int d^d x \mathcal{L}$ and S should be unitless, so $\mathcal{L}$ should have mass dimension $[\mathcal{L}] = d$. Writing down the Lagrangian

$$\mathcal{L} = \overline{\psi}(i\not{\partial} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\overline{\psi}\not{A}\psi - \frac{1}{2}(\partial_\mu A^\mu)^2,$$

we know that $[m] = 1$ (that's how units are defined) so $[\psi] = \frac{d-1}{2}$, and $[\partial_\mu] = 1$ so $[A_\mu] = \frac{d}{2} - 1$; plugging into the $e\overline{\psi}\not{A}$ term shows that we must have $[e] = 2 - \frac{d}{2}$. Thus, we will replace e with $e(\mu_R)^{2-d/2}$ so that we have a dimensionless coupling constant, and μ_R will carry our mass dimension. And the mass dimension ends up telling us how badly operators diverge if we put them in a loop (higher means worse).

Remark 33. *For Dirac algebras in dimension d, we then see that $\eta^{\mu\nu}$ has entries $(-1, 1, \cdots, 1)$ on the diagonal with $\eta^\mu{}_\mu = d$. We'll still have the Clifford algebra identities $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$, $\gamma^\mu\gamma_\mu = d$, $\gamma^\mu\gamma^\nu\gamma_\mu = (2-d)\gamma^\nu$, but importantly the trace of the unit element $tr(\mathbf{1})$ is still 4. So we're not changing the dimensions of our γ matrices themselves, and that's bad because we don't have a good γ_5 in general. But we won't go into that much here.*

We'll thus redefine our parameters via

$$\psi = \sqrt{Z_2}\psi_R, \quad A_\mu = \sqrt{Z_3}A_R, \quad m = m_r + \delta m, e = \frac{Z_1}{Z_2\sqrt{Z_3}} e_R \mu_R^{d/2-2},$$

where $Z_i = 1 + \frac{\alpha_R}{\pi}\delta_i^{(1)} + \cdots$ are again defined as perturbative expansions. Our QED Lagrangian is then

$$\mathcal{L}_{\text{QED}} = \overline{\psi}_R i(\not{\partial} - M_R)\psi_R - \frac{1}{4}F_{\mu\nu R}F_R^{\mu\nu} - \frac{1}{2}(\partial_\mu A_R^\mu)^2 - e_R\overline{\psi}_R\not{A}_R\psi_R$$

$$-\frac{1}{4}\delta_3 F_{\mu\nu R}F_R^{\mu\nu} - \frac{\delta_3}{2}(\partial_\mu A_R^\mu)^2 + \overline{\psi}_R(i\not{\partial}\delta_2 - \delta m - m_R\delta_2)\psi_R - \mu^\varepsilon e_R\overline{\psi}_R\not{A}_\mu\psi_R$$

(everything in the second line is extra terms from renormalization) and we can now adjust the counter terms to get rid of the divergent diagrams. The two-point correlation function $\int d^4x \left\langle 0\middle|T\{\overline{\psi}(x)\psi(0)\}\middle|0\right\rangle$ contains a contribution from just a straight directed line segment, then at order α we have connecting two points on that segment with a photon line, at order α^2 two different contributions from two photon lines coming off of that segment, and so on. We'll define a concept here: we are **1PI** (one-particle irreducible) if whenever we cut one line, the diagram does not fall apart. (So one of the two α^2 contributions is 1PI but the other is not.) Then let the series consisting of all 1PI diagrams be

Σ (they basically consist of diagrams where all photon lines are concentrically nested) – our correlation function then consists of just the fermion propagator term with no additional lines, $\frac{i(\not p+m)}{p^2-m^2+i0} = \frac{i}{\not p - m} = i(\not p - m)^{-1}$, and then the diagrams from all 1PI contributions, then the diagrams where we have two disjoint blobs of 1PI contributions, and so on: thus the contributions to our integral are

$$\frac{i}{\not p - m} + \frac{i}{p^2-m}\Sigma\frac{i}{p^2-m} + \frac{i}{p^2-m}\Sigma\frac{i}{p^2-m}\Sigma\frac{i}{p^2-m} + \cdots = \frac{i}{\not p - m - i\Sigma + i0}$$

by doing a geometric series. We now need to set up our renormalization schemes (conditions): we want, at all orders,

$$\int d^4x e^{ipx}\left\langle 0\middle|T\{\overline{\psi}(x)\psi(0)\}\middle|0\right\rangle\Big|_{p^2=m_R^2} = \left.\frac{i}{\not p - m_R + i0}\right|_{\not p = m_R}.$$

We thus need (plugging in renormalized quantities on the left-hand side)

$$\left.\frac{i(1+\delta_2)}{\not p - m - \delta m + i0}\right|_{\not p = m} = \frac{i}{\not p - m - i\Sigma(p^2-m^2) - (\not p - m)\left.\frac{d\Sigma}{d\not p}\right|_{\not p = m} + (\not p - m)^2} = \left.\frac{i(1+\delta_2)}{(\not p - m_R)\left(1 - i\left.\frac{d\Sigma}{d\not p}\right|_{\not p = m}\right)}\right|_{\not p = m},$$

so the two conditions we must have for singularities to go away are

$$\delta m = i\Sigma(p^2 = m^2), \quad \delta_2 = -i\left.\frac{d\Sigma}{d\not p}\right|_{\not p = m}$$

at **all** orders. (This is called the **on-shell scheme**.) So we'll now go through the self-energy corrections and see the calculations more explicitly:

Example 34

Consider the diagram with a single photon line loop, where the momentum is $p - k$ on along the photon line.

Then the contributions within the loop are part of Σ, and collecting terms gives us

$$\begin{aligned}\Sigma^{(1)} &= \int \frac{d^dk}{(2\pi)^d}(-ie\gamma^\mu)\frac{i(\not k + m)}{k^2-m^2}(-ie\gamma^\nu)\frac{-i\eta_{\mu\nu}}{(p-k)^2} \\ &= -e^2\int\frac{d^dk}{(2\pi^d)}\frac{\gamma^\mu(\not k + m)\gamma_\mu}{(k^2-m^2)(p-k)^2} = -\frac{e^2}{(2\pi)^d}\int d^dk\frac{(2-d)\not k + dm}{(k^2-m^2)(p-k)^2},\end{aligned}$$

where we've used the γ matrix identities. We'll now introduce Feynman parameters, where the trick is basically that

$$\frac{1}{A} = \int_0^\infty dx e^{-xA} \implies \frac{1}{AB} = \int_0^\infty\int_0^\infty dx_1dx_2e^{-x_1A-x_2B} = \int_0^\infty\int_0^\infty\int_0^\infty dydx_1dx_2\delta(y-x_1-x_2)e^{-x_1A-x_2B},$$

and now rescaling x_1 and x_2 by y times yields

$$= \int_0^\infty\int_0^\infty\int_0^\infty dydx_1dx_2y\delta(1-x_1-x_2)e^{-Ax_1y-Bx_2y} = \int_0^\infty dx_1dx_2\frac{\delta(1-x_1-x_2)}{(Ax_1+Bx_2)^2}.$$

Plugging back in, we see that the integral we want to evaluate is

$$I_1 = \int\frac{d^dk}{(2\pi)^d}\frac{1}{(k^2-m^2)(k-p)^2} = \int\frac{d^dk}{(2\pi)^d}\int dx_1dx_2\delta(1-x_1-x_2)\left[(k^2-m^2)x_1 + (k^2-2kp+p^2)x_2\right]^{-2};$$

since $x_2 = 1-x_1$ the bracketed term is $k^2 - 2kp(1-x_1) - m^2x_1 + p^2(1-x_1)$. Doing a linear shift $k \mapsto k + p(1-x_1)$, we thus have

$$= \int\frac{d^dk}{(2\pi)^d}\int dx_1dx_2\delta(1-x_1-x_2)\left[k^2 + x_1((1-x_1)p^2 - m^2)\right]^{-2}.$$

Then performing a Wick rotation like before, this simplifies to

$$= \int_0^1 dx_1 i \int \frac{d(k_E)^2}{2(2\pi)^d} (k_E^2)^{d/2-1} d\Omega_D [-k_E^2 - \Delta]^{-2},$$

where $\Delta = -x_1((1-x_1)p^2 - m^2)$, and now we can just calculate this integral because we've converted to the Euclidean metric: we have

$$I_1 = \frac{i\Omega_d}{2(2\pi)^d} \frac{\Gamma(d/2)\Gamma(2-d/2)}{\Gamma(2)} \int_0^1 dx_1 (-(p^2(1-x_1) - m^2)x_1)^{d/2-2}.$$

The point was that we have something quadratic in k^2 – after some manipulation there's no mixed term kp and thus we can turn our problem into a spherical integral, using that $\int dx x^a (x+\Delta)^{-b} = \Delta^{1+a-b} \frac{\Gamma(1+a)}{\Gamma(b-a-1)} \Gamma(b)$. We only really care about the counter terms here, though – expanding the result around $p^2 = m^2$ (because that's where our pole is), we have

$$\int_0^1 dx_1 (x_1^{d-4} m^{d-4} + \frac{p^2 - m^2}{m^2}(1-x_1) m^{d-4} x_1^{d-5} + O((p^2-m^2)^2)),$$

and so we have

$$I_1 = \frac{i\Gamma(1+\varepsilon)}{(1-2\varepsilon)\varepsilon(4\pi)^2} \left(\frac{m^2}{4\pi}\right)^{-2\varepsilon} \left[1 - (p^2-m^2)\frac{1}{2m^2} + O((p^2-m^2)^2)\right).$$

But now we also have to deal with the fact that there is a $\not{k}$ in the denominator, and we do so by defining

$$I_2 = \int \frac{d^d k}{(2\pi)^d} \frac{\not{k}}{(k^2-m^2)(p-k)^2};$$

doing exactly the same thing yields $\int \frac{d^d k}{(2\pi)^d} \int_0^1 dx_1 \left[\not{k}(k^2 - 2kp(1-x_1) - m^2 x_1 + p^2(1-x_1)\right]^{-2}$, which (after the shift $k \mapsto k + p(1-x_1)$) becomes

$$\int \frac{d^d k}{(2\pi)^d} \int_0^1 dx_1 (\not{k} + \not{p}(1-x_1)) \left[k^2 + x_1(p^2(1-x_1) - m^2)\right]^{-2},$$

and in fact the contribution from the only remaining $\not{k}$ in this expression will now cancel out because the parts with k and $-k$ do. Thus

$$I_2 = \frac{1}{2(1-\varepsilon)} \not{p} \frac{i\Gamma(1+\varepsilon)}{(1-2\varepsilon)\varepsilon(4\pi)^2} \left(\frac{m^2}{4\pi}\right)^{-\varepsilon} \left[1 + \frac{p^2-m^2}{m^2} + \cdots\right].$$

We then also find that

$$\Sigma|_{p^2=m^2} = \frac{i\alpha_R}{4\pi\varepsilon} \left(\frac{m^2}{4\pi\mu^2}\right)^{-\varepsilon} \Gamma(1+\varepsilon) \frac{3-2\varepsilon}{1-2\varepsilon} = i\delta m$$

by definition, and thus in order to satisfy the conditions we must have

$$\delta_2 = -i \frac{d\Sigma}{dp}\bigg|_{\not{p}=m_R} = \frac{\delta m}{m}.$$

So the point is that in the process of renormalization, we introduce a finite number of parameters that redefine couplings, masses, and fields, and in regularization we set the number of dimensions to $4-2\varepsilon$. We then want to get rid of the counter terms. The on-shell scheme then shows that the fermion propagator (vacuum energy of the electron) is $\frac{i}{\not{p}-m_R+i0}$ to all orders in peturbation theory, the photon propagator is $-\frac{iq^{\mu\nu}}{p^2+i0}$ in all orders, and a vertex (as $q^2 \to 0$) is given by $-ieF(q^2 \to 0, m^2)\overline{u}\gamma^0 u\phi + \delta(q^2)$, where ϕ is the electric potential. And in particular, $F(0,m^2) = 1$, so we're not changing the electric charge to all orders in perturbation theory – the point is that $E \neq E_R$ but F_R is just constantly 1. That makes $\mathcal{L} = \mathcal{L}_R + \mathcal{L}_{\text{counter}}$, and by defining our appropriate constants makes the Lagrangian look the same as we're physically used to – we cancel out the singularities in all diagrams that we can construct.

18 December 6, 2022

We'll start today by summarizing the ideas of renormalization once again. Basically, when we have divergent diagrams (such as loops in a photon propagator), we characterize how badly they diverge using a superficial degree of divergence D (which was $4-\frac{3}{2}E_e-P_e$ for QED, where E_e, p_e are the number of external e^-/e^+ lines and γ lines, respectively). We then regulate this divergence by setting dimension to $d=4-2\varepsilon$ and take $\varepsilon\to 0$ in a controlled way – we can then redefine (renormalize) parameters and fields in our theory to cancel out the counter-terms. Specifically, in QED we set $\psi=\sqrt{Z_2}\psi_R$, $A^\mu=\sqrt{Z_3}A^\mu_R$, $m=m_R+\delta m$, and $e=\frac{Z_1}{Z_2\sqrt{Z_3}}\mu_R^\varepsilon e_R$, where each $Z_i=1+\delta_i$ can be expanded in perturbation theory. The point is that these corrections from δs only come into play at higher order, and we found the renormalization conditions for QED last time (demanding that all potential loop contributions from the electron propagator at $p^2=m_R^2$ should yield $\frac{i}{p^2-m_R^2+i0}$, and that all potential contributoins from the photon propagator at $p^2=0$ together are $-\frac{i\eta^{\mu\nu}}{p^2+i0}$, and a third condition coming from the vertices) – we then ended up with the equations

$$\frac{\alpha}{\pi}\delta m^{(1)}=i\Sigma|_{p^2=0},\quad \frac{\alpha}{\pi}\delta_2^{(1)}=-i\frac{d\Sigma}{d\not{p}}|_{p^2=m^2}.$$

Then because the self-energy of electrons, photons, and vertices are the only situations in this theory with a superficial degree of divergence, every potential problem diagram will have one of them as a subdiagram. So this renormalizes the whole theory properly. (We should remember that the observable actual values of our parameters are what correspond to physical values – the original ones are just badly divergent and incorrect.)

The problem is that there are other issues with quantum field theory that we've ignored and only hinted at so far in this course. We've dealt with **ultraviolet** singularities (loop momenta going to ∞) using renormalization, but we can encounter **infrared** (also called **collinear**) singularities at low energies as well. For example, introducing a connecting photon line of momentum k near a vertex by connecting the two electron lines give us $\int d^4k\frac{1}{k^2(k+p_1)^2(k-p_2)^2}$, and when $p_1^2=p_2^2=0$ and $k^2\sim 0$ (remember these are still four-vectors) our integral is basically

$$\int d^4k\frac{1}{(k^2(2kp_1)(-2kp_2)}=\int\frac{d|k|^2}{2}|k|^2\frac{d\Omega_4}{k^2(2kp_1)(-2kp_2)}\sim\int\frac{dE^2}{E^2}\frac{d\Omega_4}{(1-\cos\theta_{k1})(1-\cos\theta_{k2})}$$

because $kp\approx -E_kE_p(1-\cos\theta_{kp})$. So if we take $E_k\to 0$ (so very low energy and very high wavelength, corresponding to photon interactions across huge distances), or if that photon line is almost collinear with p_1 or p_2, connecting the diagram also diverges and this is a different issue from that in renormalization (in which we imagine that the photon line is getting very close to the vertex).

The solution is basically to not just add loop diagrams but also define our observables accordingly. A scattering process $e^-+e^-\to\gamma\to\mu+\mu^-$ with an additional γ line connecting μ and μ^- must also include corrections where an actual photon is emitted – these might look completely different, but in the limit where the photon energy goes to zero the processes will look experimentally identical, and there is a singularity that corresponds to the infrared singularity in the original diagram. The same is true of collinearity (since detector can not resolve the difference between the muon and photon, and adding the quantum numbers gives us back the original muon). So the point is to think about the cross-section σ as a sum of the real diagrams and the virtual diagrams (we need to make sure the "infinitely-long distances" don't end up contributing to our physical result), and if we want to learn more about this we should search up the **KLN theorem**.

What's nice, though, is that dimensional regularization can regularize both ultrviolet and infrared singularities. In the previous lecture, we looked at the on-shell scheme, and when we subtract off only the $\frac{1}{\varepsilon_{UV}}$ singularities from ultraviolet that's instead called the **minimal subtraction scheme** (only removing poles). We may also use the **modified minimal subtraction scheme** ($\overline{\text{MS}}$) in which we add some constants such as modifying the coupling constant $e=$

$\frac{Z_1}{Z_2\sqrt{Z_3}}\left(\frac{\mu_R^2}{4\pi}\right)^{\varepsilon/2} e^{\varepsilon\gamma_E} e_R(\mu_R^2)$ to make our expressions nicer. (And this is in fact most widely used in calculations.) In this $\overline{\text{MS}}$ scheme, we have

$$\delta_M^{\overline{\text{MS}}} = i\Sigma|_{P^2=m^2,UV} = -\frac{3\alpha_R}{4\pi\varepsilon}, \quad \delta_1^{\overline{\text{MS}}} = -F(0,m^2)|_{(UV)} = -\frac{\alpha_R}{4\pi\varepsilon},$$

$$\delta_2^{\overline{\text{MS}}} = -i\frac{d\Sigma}{d\not{p}}|_{p^2=m^2,UV} = -\frac{\alpha_R}{4\pi\varepsilon}, \quad \delta_3^{\overline{\text{MS}}} = \Pi_2(0) = -\frac{\alpha_R}{3\pi\varepsilon}.$$

(This last expression is the total contributions from the photon self-propagator at $p^2 = 0$ in the ultraviolet correction.) The point is that these expressions look much neater than before, since there are no finite terms and we just have $\frac{1}{\varepsilon}$ type terms. And this is valid at order α_R – it turns out we need to calculate infrared finite observables to figure out how to deal with these cross-terms to remove the remaining singularities and figure out the counter terms.

If we translate between different renormalization schemes, we can relate the parameters between them. For example, we have that

$$m_R^{\overline{\text{MS}}} = m_R^{\text{OS}} + \delta m^{\text{OS}} - \delta m^{\overline{\text{MS}}} = m_R^{\text{OS}}\left[1 - \frac{\alpha_R^{\text{OS}}}{\pi}\left(1 + \frac{3}{4}\log\frac{m_R^2}{m_R^{\text{OS}}}\right) + O((\alpha_R^{\text{OS}})^2)\right]$$

(here OS means "on-shell"). And groups like the Particle Data Group always define real physical quantities relative to specific renormalization schemes – for different purposes, different schemes may be more useful, and because our series in renormalization aren't completely divergent some schemes may do worse than others in that sense.

We can now talk about the **running coupling** – the fine-structure constant $\alpha = \frac{e^2}{4\pi}$ becomes the renormalized α in the $\overline{\text{MS}}$ scheme:

$$\alpha = \frac{\alpha_R^{\overline{\text{MS}}}}{Z_3}\left(\frac{\mu_R^2}{4\pi}\right)^{\varepsilon} e^{\varepsilon\gamma_E}$$

(here notice that α_R must depend on μ_R^2) because in this case it turns out that $Z_1 = Z_2$ and $\frac{1}{Z_3} = 1 + \frac{\alpha_R^{\overline{\text{MS}}}}{3\pi\varepsilon} + O(\alpha_R^2)$. But now the left-hand side is independent of μ_R^2, so if we apply $\mu_R^2\frac{\partial}{\partial\mu_R^2}$ to both sides we get $\boxed{0 = \beta(\alpha_R(\mu_R^2))}$, where β is the **beta function** of our theory (telling us how th ecoupling constant changes in terms ofour scale μ_R).

$$\beta(\mu_R^2) = -\alpha_R\left[\beta_0\frac{\alpha_R}{\pi} + \beta_1\left(\frac{\alpha_R}{\pi}\right)^2 + \left(\frac{\alpha_R}{\pi}\right)^3\beta_2 + \cdots\right]$$

with (matching coefficients with the form of $\frac{1}{Z_3}$) $\beta_0 = -\frac{1}{3}$. So that means that

$$\mu_R^2\frac{\partial}{\partial\mu_R^2}\alpha_R^{\overline{\text{MS}}}(\mu_R^2) = \frac{\alpha_R^2(\mu_R^2)}{3\pi} + O(\alpha_R^2).$$

We can then try to solve this **renormalization group equation**

$$\frac{\partial\alpha_R}{\partial\log(\mu_R^2)} = -\frac{\beta_0}{\pi}\alpha_R^2 \implies \frac{d\alpha_R}{\alpha_R^2} = -\frac{\beta_0}{\pi}d\log(\mu_R^2),$$

and integrating both sides yields

$$-\frac{\beta_0}{\pi}\log\frac{\mu_R^2}{\mu_0^2} = -\frac{1}{\alpha_R(\mu_R^2)} + \frac{1}{\alpha_R(\mu_0^2)} \implies \alpha_R(\mu_R^2) = \frac{\alpha_R(\mu_0^2)}{1 + \frac{\alpha_R(\mu_0^2)}{\pi}\beta_0\log\frac{\mu_R^2}{\mu_0^2}}.$$

So comparing with the on-shell scheme, we find that

$$\alpha_R^{\overline{\text{MS}}}(\mu_R^2) = \alpha_S^{\text{OS}}\frac{1-\delta_3^{\overline{\text{MS}}}}{1-\delta_3^{\text{OS}}} = \alpha_R^{\text{OS}}\left(1 + \frac{\alpha_R}{3\pi}\log\frac{\mu_R^2}{m_e^2}\right),$$

so that $\alpha_R^{\overline{\text{MS}}}(m_e)$ is actually just the ordinary fine-structure constant $\approx \frac{1}{137}$ in the on-shell scheme. So α_R starts off

at $\frac{1}{137}$ at m_{e^-} and goes to ∞ as μ_R^2 increases (in fact there is a **Landau singularity** at $\lambda_{QE} = m_e \exp\left(-\frac{2\pi}{\beta_0 \alpha_R(m_e)}\right) \approx 1.2 \times 10^{278}$GeV, which is ridiculously high), or we can also think about μ_R^2 as decaying proportional to R^2 and approaching $\frac{1}{137}$ as $R \to \infty$. So QED becomes stronger and stronger coupled depending on our renormalization, and the modifications come through this kind of calculation. But the pole means that at some point we cannot think about particle degrees of freedom anymore – each order in perturbation then starts to contribute a similar amount and the probability of any number of particles interacting is essentially the same.

Fact 35

In QED, because $\beta < 0$, the Landau pole occurs in the UV regime. Meanwhile, some theories have $\beta = 0$ – they are called **conformal theories**, where the size of interaction doesn't change and there is no intrinsic scale coming from quantum corrections. Finally, there are some systems with $\beta > 0$ where the Landau pole occurs in the IR regime, such as quantum chromodynamics. (This means that as the energy gets lower and lower, we need more and more diagrams to build into our amplitudes. Indeed, in QCD we work with quarks and gluons, and there we only have protons and neutrons and cannot talk about individual degrees of freedom for quarks and gluons.) There are some two-dimensional quantum field theories similar to the Ising model where we can see the transition between different coupling regimes, but this is still limited to toy examples.

19 December 8, 2022

Last lecture, we discussed the renormalization of QED in more detail, highlighting the differences between UV (short distance) singularities, which come from just a finite set of diagrams, and infrared (long distance) singularities. Specifically, we thought about different renormalization schemes beyond the on-shell scheme such as $\overline{\text{MS}}$ (the modified minimal subtraction scheme), in which our renormalization only absorbs the $\frac{1}{\varepsilon}$ poles coming from dimensional regularization $d = 4 - 2\varepsilon$, as well as some constants that make our calculations cleaner. And it's important to remember that physical parameters like masses are dependent on the scheme that we use. In particular, there is no "actual mass" – in actual experiments we do measure the invariant masses by looking at a decay process and looking at the sum of the resulting momenta, plotting cross-section σ as a function of that. Then there will be a peak at the top quark mass, and we could call that the preferred "mass," but it's still fundamentally a parameter in the Lagrangian.

Today, we'll look at **LSZ (Lehmann-Symanzik-Zimmermann) reduction**. In the course of our renormalization procedures, we calculated the two-point correlation function (or specifically its Fourier transform): in scalar field theory this looked like $\int d^4x e^{ipx} \langle \Omega | T\{\phi(x)\phi(0)\} | \Omega \rangle |_{p^2=m^2}$, and we found that this looks like $\frac{iZ}{p^2 - m_R^2 + i0}$. This pole at $p^2 = m^2$ was then interpreted as a particle having a well-defined mass m, which is an isolated particle at a particular location. If we then consider n points and calculate an expression like

$$J = \int d^4x e^{ipx} \langle \Omega | T\{\phi(x)\phi(z_1)\cdots\phi(z_n)\} | \Omega \rangle ,$$

we now wish to identify the poles in p^0 associated with on-shell particles. To do this, we can split our time integration range into three pieces, writing $\int_\infty^\infty dx^0 = \int_\infty^{-T} dx^0 + \int_{-T}^{T} dx^0 + \int_T^\infty dx^0$, calling those three regions region I, II, and III. (Specifically, choose T so that its magnitude is much larger than that of any z_i^0, so region I and region III correspond to times "much earlier" or "much later" than the significant contributions to the correlation functions.) In region III,

where $x^0 \gg z_i^0$, we can pull x out of the time-ordered operator (because it's "happening much later")

$$\begin{aligned} J_{III} &= \int_T^\infty dx^0 \int d^3x e^{ipx} \langle\Omega|\phi(x)T\{\phi(z_1)\cdots\phi(z_n)\}|\Omega\rangle \\ &= \int_T^\infty dx^0 \int d^3x e^{ipx} \int \frac{d^3q}{(2\pi)^3}\frac{1}{2E(q)} \langle\Omega|\phi(x)|q\rangle \langle q|T\{\phi(z_1)\cdots\phi(z_n)\}|\Omega\rangle \end{aligned}$$

where we've inserted a complete set of states. We then have

$$\langle\Omega|\phi(x)|q\rangle = \left\langle\Omega\middle|e^{ipx}\phi(0)e^{-ipx}\middle|q\right\rangle = \langle\Omega|\phi(0)|q\rangle\, e^{-iqx}|_{q^0=E(\vec{q})}$$

where $E = \sqrt{\vec{q}^2 + m^2}$, so we can rewrite

$$J_{III} = \int_T^\infty dx^0 \int \frac{d^3q}{(2\pi)^3}\frac{1}{2E(q)} e^{i(p^0-q^0)x^0} \int d^3x e^{i(\vec{p}-\vec{q})\cdot\vec{x}} \langle\Omega|\phi(0)|q\rangle \langle q|T\{\phi(z_1)\cdots\phi(z_n)\}|\Omega\rangle$$

and the point is that the only x-dependence is a delta-function $\int d^3x e^{i(\vec{p}-\vec{q})\cdot\vec{x}} = (2\pi)^3\delta^{(3)}(\vec{p}-\vec{q})$, which we can then use to also calculate the d^3q momentum integral. This thus all simplifies to

$$= \int_T^\infty dx^0 \frac{e^i(p^0 - E(\vec{p}))x^0}{2E(\vec{p})} \langle\Omega|\phi(0)|\vec{p}, E(\vec{p})\rangle \langle\vec{p}, E(\vec{p})|T\{\phi(z_1)\cdots\phi(z_n)\}|\Omega\rangle\,,$$

where we're writing the momentum eigenstate in a funny way because we haven't specified p^0 yet. We can then carry out the last x^0 integration and end up with

$$J_{III} = \frac{ie^{i(p^0-E(\vec{p})T}}{2E(\vec{p})(p^0 - E(\vec{p}) + i0)} \langle\Omega|\phi(0)|\vec{p}, E(\vec{p})\rangle \langle\vec{p}, E(\vec{p})|T\{\phi(z_1)\cdots\phi(z_n)\}|\Omega\rangle\,.$$

So if there is a pole at $p^0 = E(\vec{p})$, then the contribution to the correlation function as $T \to \infty$ for $p^0 \to E(\vec{p})$ is of the form

$$J \sim \frac{i\sqrt{Z}}{p^2 - m_R^2 + i0} \langle p|T\{\phi(z_1)\cdots\phi(z_n)\}|\Omega\rangle\,,$$

where we've used that $\langle\Omega|\phi(0)|p\rangle = \sqrt{Z}$, $\frac{1}{2E(\vec{p})(p^0-E(\vec{p})+i0} = \frac{1}{p^2-m_R^2+i0}\Big|_{p^2=m_R^2}$. (So the point is that poles propagate a single particle, and we're saying that we generate a one-particle state at isolated future times if we take the Fourier transform and evaluate it near an on-shell pole.)

Next, we can look at region I, in which

$$J_I = \int_{-\infty}^{-T} dx^0 \int d^3x e^{ipx} \langle\Omega|T\{\phi(z_1)\cdots\phi(z_n)\}\phi(x)|\Omega\rangle$$

because now we're in a region where x^0 is much earlier than the z_i^0s. Again inserting a complete set of states and doing the same calculations as before, we see that

$$\lim_{p^0\to -E(\vec{p})} J = \lim_{p^0\to -E(\vec{p})} \frac{i\sqrt{Z}}{p^2 - m_R^2 + i0} \langle\Omega|T\{\phi(z_1)\cdots\phi(z_n)\}|-p\rangle\,,$$

so if we are forcing p^0 to be the negative energy (in this definition of the Fourier transform of the momentum p), we're picking up a particle pole sitting in the initial state with momenutm $-p$. And now we can repeat this for all n of our particles, and we find that

$$\lim_{p_f^0\to E_f(p_f), k_i^0\to E_i(\vec{k})} \int \prod_{i=1}^n d^4x e^{ix_ik_i} \prod_{p=1}^m e^{-iy_pk_f} dy_f^4 \langle\Omega|T\{\phi(x_1)\cdots\phi(x_i)\cdots\phi(x_n)\phi(y_1)\cdots\phi(y_f)\cdots\phi(y_m)\}|\Omega\rangle$$

that is, if we send the final-state momenta to what we want them to be, we get single-particle-state energies. And

specifically sending the final and initial p_f^0 and p_i^0 to the corresponding energies $E_f(p_f)$ and $E_i(p_i)$, we get

$$\prod_{i=1}^{n} \frac{i\sqrt{Z}}{p_i^2 - m_R^2 + i0} \prod_{i=1}^{m} \frac{i\sqrt{Z}}{p_f^2 - m_R^2 + i0} \langle k_1, \cdots, k_f | T\{1\} | p_1, \cdots, p_n \rangle$$

and thus we can take an n-point correlation function (which is what we learned how to calculate in the first half of the course, at least for scalar fields), doing a Fourier transform, demanding that the momenta with respect to our on-shell one-particle states can be related to **some propagator factors times an S-matrix element**. And in fact we can turn this around and use it to define an S-matrix element:

$$\langle k_1, \cdots, k_f | T\{1\} | p_1, \cdots, p_n \rangle = \lim_{k_f^0 \to E(k_f), p_i^0 \to E(p_i)} \prod_{i=1}^{n} \left(\frac{i\sqrt{Z}}{p^2 - m_R^2 + i0} \right)^{-1} \prod_{i=1}^{m} \left(\frac{i\sqrt{Z}}{p^2 - m_R^2 + i0} \right)^{-1}$$
$$\int \prod_{i=1}^{n} d^4x e^{ip_i x_i} \prod_{f=1}^{m} dy_f e^{-ik_f x_f} \langle \Omega | T\{\phi(p_1) \cdots \phi(p_n) \phi(k_1) \cdots \phi(k_m)\} | \Omega \rangle .$$

Graphically, we can think of the $\frac{i\sqrt{Z}}{p^2 - m_R^2 + i0}$ terms as the inverses of diagrams coming from a single propagator (accounting for all potential loops and other interactions inside), one for each of the n incoming particles and m outgoing particles, and the final term coming from sending in n particles and getting m particles out. So basically what we want to count is only the **amputated** $(n+m)$-momentum-space diagrams for the S-matrix element. So the key is to take the n-point correlation function that we know how to derive, Fourier transform it with momenta corresponding to on-shell particle momenta, and then if we're in a situation with an isolated n-particle state we must pick up the right residue for the pole. (But if we think about this situation for QED, the photon is massless and everything becomes messy.)

There's now one more thing we need to clear up, the **Ward-Takahashi identity** (this is a totally new topic). Suppose we have a scattering of a bunch of particles including a photon, so that we can write the scattering matrix element $M = \varepsilon^\mu M_\mu$. Then based on gauge invariance, we find that $k^\mu M_\mu$ if k is corresponding to the photon momentum $\varepsilon(k)$, and we want to understand where this comes from in generality for QFT. To do this, start with a fermion line and suppose we have n photon lines going into it of momentum q_1, scattering in and "adding momentum." So if the initial fermion momentum is p, we get the momentum $p_1 = p + q_1, p_2 = p_1 + q_2, \cdots, p' = p_{n-1} + q_n$ at various points on the line. Now suppose we add another photon of momentum k in between q_i and q_{i+1} on the line, so that p_i becomes $p_i + k$ and our final momentum is $p' + k$. We then want to replace the polarization vector k^μ with its momentum ε^μ. The insertion of a vertex can then be written as $-ie((\not{p}_i + \not{k} + m) - (\not{p}_i - m))$, so if we just look at the new vertex inserted we get

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

For the general fermion line in which the diagram continues, we can then write the contribution as

$$\cdots \frac{i}{\not{p}_i + \not{k} - m} \gamma^{\mu_{i+1}} \left(\frac{ie}{\not{p}_i - m} - \frac{ie}{\not{p}_i + \not{k} - m} \right) \gamma^{\mu_i} \frac{i}{\not{p}_i - m} \cdots .$$

Meanwhile, if we insert into the neighboring location between q_{i-1} and q_i instead, we end up with

$$\cdots \frac{i}{\not{p}_{i+1} + \not{k} - m} \gamma^{\mu_{i+1}} \frac{i}{\not{p}_i + \not{k} - m} \gamma^{\mu_i} \left(\frac{ie}{\not{p}_{i-1} - m} - \frac{ie}{\not{p}_{i-1} + \not{k} - m} \right) \gamma^{\mu_{i-1}} \cdots .$$

We can then compare the two diagrams and notice that there will be lots of cancellations if we add them together by telescoping. Adding across all possible insertion points, we get e times the original fermion line without any insertion, minus e times the fermion line with shifted momenta $p + k$ everywhere in place of p. And indeed if we think about our photon as part of scattering, it will interact with our scattering amplitude in some way, and that's exactly summing

over all insertion points. We thus find that if we have a scattering process with n photon points, we have a loop amplitude (where momentum starts off as p_1 and everything else is conserved by momentum conservation at vertices)

$$k^\mu M_\mu = -e(-ie)^n \int \frac{d^d p_i}{(2\pi)^d} \left(\text{tr}\left[\frac{i}{\not{p}_n - m} \gamma^{\mu_n} \cdots \frac{i}{\not{p}_1 - m} \gamma^{\mu_1} \right] - \text{tr}\left[\frac{i}{\not{p}_n + \not{k} - m} \gamma^{\mu_n} \cdots \frac{i}{\not{p}_1 + \not{k} - m} \gamma^{\mu_1} \right] \right),$$

where one trace corresponds to momentums shifted and one not. But then by shift-invariance of the loop momentum, we can replace $p_1 \mapsto p_1 - k$ in one of the two traces, so we indeed have $k^\mu M_\mu = 0$ for loops. However, if we have an external fermion coming into our amplitude, we shouldn't write the amplitude with propagators because we should be amputating our diagrams – we should write

$$M^\mu = \sum_{\text{insertion points}} \overline{u}(p' + k)(-i(\not{p}' + k - m)) \cdot (\text{insertion diagram with photon } k \text{ inserted}) \cdot (-i(\not{p} - m))u(p)$$

where u is some spinor. Then the graphical identity we just derived simplifies this to

$$= e\overline{u}(p'+k)(-i(\not{p}'+\not{k}-m))\cdot[(\text{original insertion diagram}) - (\text{original insertion diagram with shifted momenta}]\,(-i(\not{p}-m))u(p).$$

(where "original insertion diagram" means we don't include the photon k). Then we can use the Dirac equation

$$(\not{p} - m)u(p) = \overline{u}(p^2 + k)(\not{p} + \not{k} - m = 0$$

and indeed find that $k^\mu M_\mu = 0$ in this case as well even if we don't have a loop. So when we did calculations with Compton scattering, for example, we calculated square matrix elements and had to substitute for polarization sum and brute-force replaced $\sum_{\lambda=1}^{2} \varepsilon^{\mu_{(\lambda)}}(k)\varepsilon^{\nu}_{(\lambda)}(k) = -\eta^{\mu\nu}$. But we should really be summing over physical polarizations (only transverse in the direction of motion), so that wasn't really correct. (If we have $\varepsilon^{\mu}_{(1)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $\varepsilon^{\nu}_{(2)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, we instead get the matrix $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.) So what's happening here is that if we have some $k^\mu = \begin{bmatrix} k \\ 0 \\ 0 \\ k \end{bmatrix}$ with $\varepsilon_{(\lambda)} k = 0$, we can introduce an auxiliary gauge vector n^μ that also satisfies $n \cdot \varepsilon = 0$ and is independent of k. We can then create a polarization sum that is actually useful to us: consider

$$\frac{k^\mu n^\nu + k^\nu n^\mu}{n \cdot k} - \frac{k^\mu k^\nu}{(n \cdot k)^2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

and then we instead find that

$$\sum_{\lambda=1}^{2} \varepsilon^{\mu_{(\lambda)}}(k)\varepsilon^{\nu}_{(\lambda)}(k) = -\eta^{\mu\nu} - \frac{k^\mu n^\nu + k^\nu n^\mu}{n \cdot k} + \frac{k^\mu k^\nu}{(n \cdot k)^2}.$$

But when we calculate squared matrix elements like for Compton scattering, we're basically using expressions like $M^\mu P_{\mu\nu} M^{\dagger\nu}$, and by the Ward identity we just derived, we really only need the metric term because $k^\mu M_\mu = 0$. So we indeed don't need to deal with complicated additional terms because of this graphical proof!