QUANTUM FIELD THEORY

IAN LIM
LAST UPDATED JANUARY 21, 2019

These notes were taken for the *Quantum Field Theory* course taught by Ben Allanach at the University of Cambridge as part of the Mathematical Tripos Part III in Michaelmas Term 2018. I live-TEXed them using Overleaf, and as such there may be typos; please send questions, comments, complaints, and corrections to itel2@cam.ac.uk.
Many thanks to Arun Debray for the LATEX template for these lecture notes: as of the time of writing, you can find him at https://web.ma.utexas.edu/users/a.debray/.

Contents

Lecture 1.

Thursday, October 4, 2018

$2 = \pi = i = -1$ *in these lectures.* –a former lecturer of Prof. Allanach's.

To begin with, some logistic points. The notes and much of the course material will be based on David Tong's QFT notes plus some of Prof. Allanach's on cross-sections and decay rates. See http://www.damtp.cam.ac.uk/user/examples/indexP3.html and in particular http://www.damtp.cam.ac.uk/user/examples/3P11.pdf for the notes on cross-sections. In revising these notes, I'll be cross-referencing the Tong

QFT notes as well as my copy of Anthony Zee's *Quantum Field Theory in a Nutshell*, which takes a different pedagogical order in starting from the path integral formalism and introducing second quantization (the approach described here) later. Any good education in QFT requires an understanding of both formalisms, and we'll see the path integral next term in *Advanced Quantum Field Theory*.[1]

After Tuesday's lecture, we'll be assigned one of four course tutors:

- Francesco Careschi, fc435@cam.ac.uk
- Muntazir Abidi, sma74
- Khim Leong, lkw30
- Stefano Vergari, sv408

Also, the Saturday, November 24th lecture has been moved to 1 PM Monday 26 November, still in MR2. That's it for logistics for now.

Definition 1.1. A *quantum field theory* (QFT) is a field theory with an infinite number of degrees of freedom (d.o.f.). Recall that a field is a function defined at all points in space and time (e.g. air temperature is a scalar field in a room wherever there's air). The states in QFT are in general multi-particle states.

Special relativity tells us that energy can be converted into mass, and so particles are produced and destroyed in interactions (particle number is in general not conserved). This reveals a conflict between SR and quantum mechanics, where particle number is fixed. Interaction forces in our theory then come from additional structure in the theory, depending on things like

- symmetry
- locality
- "renormalization group flow."

Definition 1.2. A *free QFT* is a QFT that has particles but no interactions. The classic free theory is a relativistic theory with which treats particles as excitations of infinitely many quantized harmonic oscillators.

Free theories are generally not realistic but they are important, as interacting theories can be built from these with perturbation theory. The fact we can do this means the particle interactions are somehow weak (we say these theories have *weak coupling*), but other theories of interest (e.g. the strong force) have strong coupling and cannot be described with perturbation theory.

Units in QFT In QFT, we usually set $c = \hbar = 1$. Since $[c] = [L][T]^{-1}$ and $[\hbar] = [L]^2[M][T]^{-1}$, we find that in natural units,

$$[L] = [T] = [M]^{-1} = [E]^{-1}$$

(where the last equality follows from $E = mc^2$ with $c = 1$, for example). We often just pick one unit, e.g. an energy scale like eV, and describe everything else in terms of powers of that unit. To convert back to metres[2] or seconds, just reinsert the relevant powers of c and $\hbar$.

Example 1.3. The de Broglie wavelength of a particle is given by $\lambda = \hbar/(mc)$. An electron has mass $m_e \simeq 10^6$ eV, so $\lambda_e = 2 \times 10^{-12}$ m.

If a quantity x has dimension $(mass)^d$, we write $[x] = d$, e.g.

$$G = \frac{\hbar c}{M_p^2} \implies [G] = -2.$$

$M_p \approx 10^{19}$ GeV corresponds to the Planck scale, $\lambda_p \sim 10^{-33}$ cm, the length/energy scales where we expect quantum gravitational effects to become relevant. We note that the problems associated with relativising the Schrödinger equation are fixed in QFT by particle creation and annihilation.

[1] Note that the path integral formulation from *Statistical Field Theory*, also taught this term, is precisely equivalent to the path integral that appears in QFT under the identification of one of the Euclidean dimensions of a statistical field theory with the imaginary time dimension of a QFT. This will be more obvious in hindsight.

[2] As a USAmerican, I am likely to be bewilderingly inconsistent with regards to using American versus British spellings. Please bear with me.

Classical field theory Before we do QFT, let's review classical field theory. In classical particle mechanics, we have a finite number of generalized coordinates $q_a(t)$ (where a is a label telling you which coordinate you're talking about), and in general they are a function of time t. But in field theory, we instead have continuous fields $\phi_a(\mathbf{x}, t)$, where a labels the field in question and $\mathbf{x}$ is no longer a coordinate but a label like a.[3]

In our classical field theory, there are now an infinite number of degrees of freedom, at least one for each position in space $\mathbf{x}$, so position has been demoted from a dynamical variable to a mere label.

Example 1.4. The classical electromagnetic field is a vector field with components $E_i(x,t), B_i(x,t)$ such that $i, j, k \in \{1, 2, 3\}$ label spatial directions. In fact, these six fields are derived from four fields (or rather four field components), the four-potential $A_\mu(x,t) = (\phi, \mathbf{A})$ where $\mu \in \{0, 1, 2, 3\}$.

Then the classical fields are simply related to the four-potential by

$$E_i = \frac{\partial A_i}{\partial t} - \frac{\partial A_0}{\partial x_i} \text{ and } B_i = \frac{1}{2}\epsilon_{ijk}\frac{\partial A_k}{\partial x_j} \tag{1.5}$$

with ϵ_{ijk} the usual Levi-Civita symbol, and where we have used the Einstein summation convention (repeated indices are summed over).

The dynamics of a field are given by a *Lagrangian* L, which is simply a function of $\phi_a(x,t), \dot{\phi}_a(x,t)$, and $\boldsymbol{\nabla}\phi_a(x,t)$. This is in precise analogy to the Lagrangian of a discrete system, which is a function of the coordinates $q_a(t)$ and their derivatives $\dot{q}_a(t)$.

Definition 1.6. We write

$$L = \int d^3x \mathcal{L}(\phi_a, \partial_\mu \phi_a), \tag{1.7}$$

where we call $\mathcal{L}$ the *Lagrangian density*, or by a common abuse of terminology simply the Lagrangian.

Definition 1.8. We may then also define the *action*

$$S \equiv \int_{t_0}^{t_1} L dt = \int d^4x \mathcal{L}(\phi_a, \partial_\mu \phi_a) \tag{1.9}$$

Let us also note that in these units we take the action S to be dimensionless, $[S] = 0$ (since it appears alone in an exponent, for instance, e^{iS}), and so since $[d^4x] = -4$ we have $[\mathcal{L}] = 4$.

The dynamical principle of classical field theory is that fields evolve such that S is stationary with respect to variations of the field that don't affect the initial or final values (boundary conditions). That is, $\delta S = 0$. A general variation of the fields produces a variation in the action

$$\delta S = \sum_a \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a}\delta\phi_a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\delta(\partial_\mu \phi_a) \right\}.$$

Integrating the second term by parts, we find that the variation in the action becomes

$$\delta S = \sum_a \int d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a}\delta\phi_a + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\delta\phi_a \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \delta\phi_a \right\}.$$

The integral of the total derivative term vanishes for any term that decays at spatial ∞ (i.e. $\mathcal{L}$ is reasonably well-behaved) and has $\delta\phi_a(x, t_1) = \delta\phi_a(x, t_0) = 0$, as guaranteed by our boundary conditions. Therefore the boundary term goes away and we find that stationary action, $\delta S = 0$, implies the *Euler-Lagrange equations*,

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0. \tag{1.10}$$

Example 1.11. Consider the Klein-Gordon field ϕ, defined as the real-valued field ϕ which has a Lagrangian

$$\mathcal{L} = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2. \tag{1.12}$$

Here $\eta^{\mu\nu}$ is the standard Minkowski metric[4].

[3] See for instance Anthony Zee's *QFT in a Nutshell* to see a more detailed description of how we go from discrete to continuous systems.

[4] We use the mostly minus convention here, but honestly the sign conventions are all arbitrary and relativity often uses the other one where time gets the minus sign.

To compute the Euler-Lagrange equation for this field theory, we see that

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2\phi \text{ and } \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi.$$

The Euler-Lagrange equations then tell us that ϕ obeys the equation of motion

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0,$$

which we call the *Klein-Gordon equation*. It has wave-like solutions $\phi = e^{-ipx}$ with $(-p^2 + m^2)\phi = 0$ (so that $p^2 = m^2$, which is what we expect in units where $c = 1$).

Non-lectured aside: on functional derivatives If you're like me, you get a little anxious about taking complicated functional derivatives. The easiest way to manage these is to rewrite the Lagrangian so that all terms precisely match the form of the quantity you are taking the derivative with respect to, and remember that matching indices produce delta functions.

Here's a quick example. To compute $\frac{\partial}{\partial(\partial_\alpha \phi)}\left[\partial_\mu \phi \partial^\mu \phi\right]$, rewrite the term in the brackets as $\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$ (since we are deriving with respect to a function of the form $\partial_\alpha \phi$) and make sure to take the derivative with respect to a new index not already in the expression, e.g. $\partial_\alpha \phi$. Then

$$\begin{aligned}
\frac{\partial}{\partial(\partial_\alpha \phi)}\left[\partial_\mu \phi \partial^\mu \phi\right] &= \frac{\partial}{\partial(\partial_\alpha \phi)}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \\
&= \eta^{\mu\nu}(\delta^\alpha_\mu)\partial_\nu\phi + \eta^{\mu\nu}\partial_\mu\phi(\delta^\alpha_\nu) \\
&= 2\partial^\alpha\phi,
\end{aligned}$$

where we have raised the index with $\eta^{\mu\nu}$ and written the final expression in terms of α using the delta function. The functional derivative effectively finds all appearances of the denominator exactly as written, including indices up or down, and replaces them with delta functions so the actual indices match. This is especially important in computing the Euler-Lagrange equations for something like Maxwell theory, where one may have to derive by $\partial_\mu A_\nu$ and both those indices must match exactly to their corresponding appearances in the Lagrangian.

No one ever taught me exactly how to approach such variational problems, so I wanted to record my strategy here for posterity. It may take a little longer than just recognizing that $\frac{\partial}{\partial(\partial_\mu\phi)}\frac{1}{2}\partial_\nu\phi\partial^\nu\phi = \partial^\mu\phi$, but this approach always works and it has the benefit of helping avoid careless mistakes like forgetting the factor of 2 in the example above.

Lecture 2.

Saturday, October 6, 2018

Last time, we derived the Euler-Lagrange equations for Lagrangian densities:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0. \tag{2.1}$$

Today, we'll look at some more simple Lagrangians. We'll introduce Noether's theorem as it applies to fields and also derive the energy-momentum tensor in a field theory context.

Example 2.2. Consider the Maxwell Lagrangian,

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A^\mu)^2. \tag{2.3}$$

Plugging into the E-L equations, we find that $\frac{\partial \mathcal{L}}{\partial A_\nu} = 0$ and

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = \partial^\mu A^\nu + \eta^{\mu\nu}\partial_\rho A^\rho. \tag{2.4}$$

Thus E-L tells us that

$$0 = -\partial^2 A^\nu + \partial^\nu(\partial_\rho A^\rho) = -\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu). \tag{2.5}$$

Defining the field strength tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, we can write the E-L equation for Maxwell as the simple

$$0 = \partial_\mu F^{\mu\nu},$$

which written explicitly is equivalent to Maxwell's equations in vacuum (we'll revisit this when we do QED).

The Lagrangians we'll consider here and afterwards are all *local*– in other words, there are no couplings $\phi(\mathbf{x}, t)\phi(\mathbf{y}, t)$ with $\mathbf{x} \neq \mathbf{y}$. There's no reason a priori that our Lagrangians have to take this form, but all physical Lagrangians seem to do so.

Lorentz invariance Consider the Lorentz transformation on a scalar field $\phi(x) \equiv \phi(x^\mu)$. The coordinates x transform as $x' = \Lambda^{-1}x$ with $\Lambda^\mu{}_\sigma \eta^{\sigma\tau} \Lambda^\nu{}_\tau = \eta^{\mu\nu}$. Under Λ, our field transforms as $\phi \to \phi'$ where $\phi'(x) = \phi(x')$. Recall that Lorentz transformations generically include boosts as well as rotations in $\mathbb{R}^3$. As we've discussed in Symmetries, Fields and Particles, Lorentz transformations form a Lie group ($O(3,1)$, or specifically the proper orthochronous Lorentz group) under matrix multipication. They have a representation given on the fields (i.e. a mapping to a set of transformations on the fields which respects the group multiplication law).

For a scalar field, this is $\phi(x) \to \phi(\Lambda^{-1}x)$ (an active transformation). We could have also used a passive transformation where we re-label spacetime points: $\phi(x) \to \phi(\Lambda x)$. It doesn't matter too much– since Lorentz transformations form a group, if Λ is a Lorentz transformation, so is Λ^{-1}. In addition, most of our theories will be well-behaved and Lorentz invariant.

Definition 2.6. *Lorentz invariant* theories are ones where the action S is unchanged by Lorentz transformations.

Example 2.7. Consider the action given by

$$S = \int d^4x \left[\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - U(\phi)\right],$$

where $U(\phi)$ is some potential density. $U \to U'(x) \equiv U(\phi'(x)) = U(x')$ means that U is a scalar field (check this!) and we see that

$$\partial_\mu\phi' = \frac{\partial}{\partial x^\mu}\phi(x') = \frac{\partial x'^\sigma}{\partial x^\mu}\partial'_\sigma\phi(x') = (\Lambda^{-1})^\sigma{}_\mu\partial'_\sigma\phi(x')$$

where $\partial'_\sigma \equiv \frac{\partial}{\partial x'^\sigma}$. Thus the kinetic term transforms as

$$\mathcal{L}_{kin} \to \mathcal{L}'_{kin} = \eta^{\mu\nu}\partial_\mu\phi'\partial_\nu\phi' = \eta^{\mu\nu}(\Lambda^{-1})^\sigma{}_\mu(\Lambda^{-1})^\tau{}_\nu\partial'_\sigma\phi(x')\partial'_\tau\phi(x') = \eta^{\sigma\tau}\partial'_\sigma\phi(x')\partial'_\tau\phi(x') = L_{kin}(x).$$

Thus we see that the action overall transforms as

$$S \to S' = \int d^4x\mathcal{L}(x') = \int d^4x\mathcal{L}(\Lambda^{-1}x).$$

Under a change of variables $u \equiv \Lambda^{-1}x$, we see that $\det(\Lambda^{-1}) = 1$ (from group theory) so the volume element is the same, $d^4y = d^4x$ and therefore

$$S' = \int d^4y\mathcal{L}(y) = S.$$

We conclude that S is invariant under Lorentz transformations.

We also remark that under a LT, a vector field A_μ transforms like $\partial_\mu\phi$, so

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

This is enough to attempt Q1 from example sheet 1.[5]

[5] Copied here for quick reference: Show directly that if $\phi(x)$ satisfies the Klein-Gordon equation, then $\phi(\Lambda^{-1}x)$ also satisfies this equation for any Lorentz transformation Λ.

Theorem 2.8. *Every continuous symmetry of $\mathcal{L}$ gives rise to a current J^μ which is conserved, $\partial_\mu j^\mu = 0$. Each j^μ has a conserved charge $Q = \int_{\mathbb{R}^3} j^0 d^3x$.*

This conserved charge appears because $\frac{dQ}{dt} = \int_{\mathbb{R}^3} d^3x \partial_0 j^0 = -\int_{\mathbb{R}^3} d^3x \boldsymbol{\nabla}\cdot\mathbf{j} = 0$ by the divergence theorem, assuming $|\mathbf{j}| \to 0$ as $|\mathbf{x}| \to \infty$.

Let us define an infinitesimal variation of a field ϕ, $\phi(x) \to \phi'(x) = \phi(x) + \alpha\Delta\phi(x)$ with α an infinitesimal change. If S is invariant, we call this a *symmetry* of the theory.

Since S is invariant up to adding a total 4-divergence (a total derivative ∂_μ) to the Lagrangian, our symmetry doesn't affect the Euler-Lagrange equations. L transforms as

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha\partial_\mu X^\mu(x), \tag{2.9}$$

and expanding to leading order in α we have

$$\mathcal{L} \to \mathcal{L}(x) + \alpha\frac{\partial\mathcal{L}}{\partial\phi}\Delta\phi + \alpha\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\Delta\phi) + O(\alpha^2). \tag{2.10}$$

We can rewrite this in terms of a total derivative $\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Delta\phi\right)$ so that

$$\mathcal{L}' = \mathcal{L}(x) + \alpha\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Delta\phi\right) + \alpha\left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\Delta\phi. \tag{2.11}$$

By Euler-Lagrange, the second term in parentheses vanishes, so we identify the first term in parentheses as none other than $\alpha\partial_\mu X^\mu(x)$ from Eqn. 2.9 (in other words, $\partial_\mu\left(\frac{\partial L}{\partial(\partial_\mu\phi)}\Delta\phi\right) = \partial_\mu X^\mu$) and recognize

$$j^\mu \equiv \frac{\partial L}{\partial(\partial_\mu\phi)}\Delta\phi - X^\mu \tag{2.12}$$

as our conserved current (such that $\partial_\mu j^\mu = 0$).

Example 2.13. Take a complex scalar field

$$\psi(x) = \frac{1}{\sqrt{2}}(\phi_1(x) + i\phi_2(x)).$$

We can then treat ψ, ψ^* as independent variables and write a Lagrangian

$$L = \partial_\mu\psi^*\partial^\mu\psi - V(|\psi|^2).$$

Then we observe that under $\psi \to e^{i\beta}\psi, \psi^* \to e^{-i\beta}\psi^*$, the Lagrangian is invariant. The differential changes are $\Delta\psi = i\psi$ (think of expanding $\psi \to e^{i\beta}\psi$ to leading order) and similarly $\Delta\psi^* = -i\psi^*$ (here we find that $X^\mu = 0$).

We add the currents from ψ, ψ^* to find

$$j^\mu = i\{\psi\partial_\mu\psi^* - \psi^*\partial_\mu\psi\}.$$

This is enough to do questions 2 and 3 on the example sheet.

Example 2.14. Under infinitesimal translation $x^\mu \to x^\mu - \alpha\epsilon^\mu$, we have $\phi(x) \to \phi(x) + \alpha\epsilon^\mu\partial_\mu\phi(x)$ by Taylor expansion (similar for $\partial_\mu\phi$). If the Lagrangian doesn't depend explicitly on x, then $\mathcal{L}(x) \to \mathcal{L}(x) + \alpha\epsilon^\mu\partial_\mu\mathcal{L}(x)$.

Rewriting to match the form $\mathcal{L} + \alpha\partial_\mu X^\mu$, we see that our new Lagrangian takes the form $L(x) + \alpha\epsilon^\nu\partial_\mu(\delta^\mu_\nu L)$. We get one conserved current for each component of ϵ^ν, so that

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

with $\partial_\mu(j^\mu)_\nu = 0$. We write this as $j^\mu{}_\nu \equiv T^\mu{}_\nu$, the energy-momentum tensor.

Definition 2.15. The *energy-momentum tensor* (sometimes *stress-energy tensor*) is the conserved current corresponding to translations in time and space. It takes the form

$$T^{\mu\nu} \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L},$$

where we have raised an index with the Minkowski metric as is conventional. The conserved charges from integrating $\int d^3x T^{0\nu}$ end up being the total energy $E = \int d^3x T^{00}$ and the three components of momentum $P^i = \int d^3x T^{0i}$.[6]

Lecture 3.

Tuesday, October 9, 2018

Last time, we used Noether's theorem to find the stress-energy tensor

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial^\nu \phi - \eta^{\mu\nu} \mathcal{L}. \tag{3.1}$$

To better understand this object, we might ask: what is $T^{\mu\nu}$ for free scalar field theory? Recall the Lagrangian for this theory is

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2. \tag{3.2}$$

Then by explicit computation, the stress-energy tensor is

$$T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \eta^{\mu\nu}\mathcal{L}.$$

The energy is given by

$$E = \int d^3x \left[\frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2\right]$$

(from integrating the T^{00} component) and the conserved momentum components are (from T^{0i})

$$p^i = \int d^3x \dot{\phi}(\partial^i \phi).$$

Note that the original Lagrangian terms don't show up here, since $\eta^{\mu\nu}$ is diagonal.

We'll note that $T^{\mu\nu}$ for this theory is symmetric in μ, ν, but a priori it doesn't have to be. If $T^{\mu\nu}$ is not symmetric initially, we can massage it into a symmetric form by adding $\partial_\rho \Gamma^{\rho\mu\nu}$ where $\Gamma^{\mu\rho\nu} = -\Gamma^{\rho\mu\nu}$ (antisymmetric in the first two indices). Then $\partial_\mu \left(\partial_\rho \Gamma^{\rho\mu\nu}\right) = 0$, which means that adding this term will not affect the conservation of $T^{\mu\nu}$. This is sufficient to attempt questions 1-6 of the first examples sheet.

Canonical quantization Here, we'll follow Dirac's lead and attempt to quantize our field theories. Recall that the Hamiltonian formalism also accommodates field theories (as well as our garden-variety QM).

Definition 3.3. We define the *conjugate momentum*

$$\pi(x) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}}$$

where a $\cdot$ denotes a time derivative d/dt, and the *Hamiltonian density* corresponding to a Lagrangian $\mathcal{L}$ is then

$$\mathcal{H} \equiv \pi(x)\dot{\phi}(x) - \mathcal{L}(x).$$

As in classical mechanics, we eliminate the time derivative $\dot{\phi}$ in favor of the conjugate momentum π everywhere in $\mathcal{H}$.

Example 3.4. For $\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi)$ (and writing in terms of $\pi(x) = \dot{\phi}(x)$) we get

$$\begin{aligned}\mathcal{H} &= (\pi)(\dot{\phi}) - \left(\frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi)\right) \\ &= \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi).\end{aligned}$$

---

[6] The definition of the energy-momentum tensor here is slightly different from the one used in general relativity. Here, we have used time and space translations to derive $T^{\mu\nu}$, but in general relativity, we use variations of the metric $g^{\mu\nu}$ instead. The benefit of the GR definition is that the resulting tensor is always symmetric, whereas the $T^{\mu\nu}$ from spacetime translations is not guaranteed to be symmetric. We'll see an example of this in the example sheets, but the $T^{\mu\nu}$ defined by spacetime translations can always be *made* symmetric by defining the "Belinfante-Rosenfeld tensor." The construction isn't anything too special, but relativists insist that variations of the action with respect to the metric is the correct way to define the energy-momentum tensor.

The Hamiltonian is just the integral of the Hamiltonian density: $H = \int d^3x\mathcal{H}$. Hamilton's equations then yield the equations of motion:

$$\dot{\phi} = \frac{\partial H}{\partial \pi}, \dot{\pi} = -\frac{\partial H}{\partial \phi}.$$

Working these out explicitly for the free theory will give us back the Klein-Gordon equation. Note that H agrees with the total field energy E that we computed above.

There's a slight complication in working in the Hamiltonian formalism– because t is special in our equations, the theory is not manifestly Lorentz invariant (compare to the ∂_μs and variations with respect to $\delta\partial_\mu\phi$ in the Lagrangian formalism). Our original theory was Lorentz invariant, so our rewritten theory is still Lorentz invariant– it's just not immediately obvious from how we've written it.

Now let's recall that in quantum mechanics, canonical quantization takes the (classical) coordinates q_a and momenta p_a and promotes them to (quantum) operators. We also replace the Poisson bracket $\{,\}$ with commutators $[,]$. In QM, we had

$$[q_a, p^b] = i\delta_a^b,$$

working in units where $\hbar = 1$. We'll do the same for our fields ϕ_a and the conjugate momenta π_b.

Definition 3.5. A *quantum field* is an operator-valued function of space $\phi_a(\mathbf{x})$ obeying the commutation relations

$$\begin{aligned} [\phi_a(\mathbf{x}), \phi_b(\mathbf{y})] &= 0 && (3.6)\\ [\pi_a(\mathbf{x}), \pi_b(\mathbf{y})] &= 0 && (3.7)\\ [\phi_a(\mathbf{x}), \pi^b(\mathbf{y})] &= i\delta^3(\mathbf{x}-\mathbf{y})\delta_a^b. && (3.8) \end{aligned}$$

The subscript a labels which field we are talking about, and the point $\mathbf{x}$ denotes where in space we are looking.

It's no coincidence that these precisely replicate the commutation relations of the operators $\hat{x}$ and $\hat{p}$ in ordinary quantum mechanics, except that now we have an additional label $\mathbf{x}$ on the fields. If you like, quantum mechanics is just a $0+1$-dimensional QFT– there are no spatial labels to keep track of, only coordinates and momenta q_a, p^a. Note that $\phi_a(x), \pi^b(x)$ don't depend on t, since we are in the Schrödinger picture. All the t dependence sits in the states which evolve by the usual time-dependent Schrödinger equation

$$i\frac{d}{dt}|\psi\rangle = H|\psi\rangle.$$

We have an infinite number of degrees of freedom, at least one for each point x in space. For some theories (free theories), different solutions ϕ can be added together and will evolve independently– free field theories have L quadratic in ϕ_a (plus derivatives thereof), which implies linear equations of motion.

We saw that the simplest free theory leads to the classical Klein-Gordon equation for a real scalar field $\phi(\mathbf{x}, t)$, i.e. $\partial_\mu\partial^\mu\phi + m^2\phi = 0$. To see explicitly why this is a free theory, take the Fourier transform of $\phi(\mathbf{x}, t)$ to write the equations of motion in momentum space:

$$\phi(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}}\phi(\mathbf{p}, t).$$

Then we get the equation of motion

$$\left[\frac{\partial^2}{\partial t^2} + (|\mathbf{p}|^2 + m^2)\right]\phi(\mathbf{p}, t) = 0.$$

We see that the solution is a harmonic oscillator with frequency $\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$, so the general solution is a superposition of simple harmonic oscillators each vibrating at different frequencies $\omega_{\mathbf{p}}$. To quantize our field $\phi(\mathbf{x}, t)$, we have to quantize these harmonic oscillators.

Review of 1D harmonic oscillators Recall that the Hamiltonian for the simple harmonic oscillator is

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2,$$

subject to the canonical commutation condition

$$[q, p] = i,$$

where p and q are the momentum and position operators as usual. It's certainly possible to solve this system by the series method, but the algebraic method is much more elegant by far and will generalize better. Our approach is as follows– we'd like to factor the Hamiltonian (since if p and q were classical quantities we could just write it as $\frac{1}{2}(p + i\omega q)(p - i\omega q)$, for instance) but we know that this doesn't quite work because p and q do not commute. Therefore, we define the following operators:

- The creation or raising operator, $a^\dagger \equiv -\frac{i}{\sqrt{2\omega}}p + \sqrt{\frac{\omega}{2}}q$
- The annihilation or lowering operator, $a \equiv +\frac{i}{\sqrt{2\omega}}p + \sqrt{\frac{\omega}{2}}q$.

Note that we can equivalently solve for p and q in terms of a and $a^\dagger$: $q = \frac{1}{\sqrt{2\omega}}(a + a^\dagger)$ and $p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger)$. Substituting p and q into the quantization condition yields the commutator of $a, a^\dagger$,

$$[a, a^\dagger] = 1.$$

We'll then factorize the Hamiltonian into a and $a^\dagger$, picking up an extra term from the commutation relation of p and q– a little more algebra allows us to rewrite the Hamiltonian as

$$H = \frac{1}{2}\omega(aa^\dagger + a^\dagger a) = \omega\left(a^\dagger a + \frac{1}{2}\right).$$

Computing the commutators $[H, a]$ and $[H, a^\dagger]$ reveals that

$$[H, a^\dagger] = \omega a^\dagger, [H, a] = -\omega a,$$

which tells us that the operators $a, a^\dagger$ take us between energy eigenstates.[7] More specifically, they take us up and down a ladder of equally spaced energy eigenstates so that if we have one eigenstate with energy E, then we can reach a whole set of eigenstates with energy $\ldots E + 2\omega, E + \omega, E, E - \omega, E - 2\omega, \ldots$.

If we further postulate that the energy is bounded from below, this implies the existence of a ground state $|0\rangle$ such that the lowering operator acting on $|0\rangle$ kills the state: $a|0\rangle = 0$.[8] In our original Hamiltonian, this ground state has energy given by

$$H|0\rangle = \omega\left(a^\dagger a + \frac{1}{2}\right)|0\rangle = \frac{\omega}{2}|0\rangle,$$

so the ground state energy (or *zero point energy*) of the system is $\omega/2$. For our quantum theory it's really differences in energy which matter more than their absolute values,[9] so we could have just as easily written an equivalent Hamiltonian $H = \omega a^\dagger a$ and set the ground state energy to 0.

We only need one state to construct our full ladder of energy eigenstates, and we can do so by passing our equation back to q-space (real coordinates) and further writing $p = i\frac{\partial}{\partial q}$. If we plug these back into the Hamiltonian, having set $H|0\rangle = 0$, we can then solve for the ground state and find that it is a Gaussian in q with some appropriate variance and normalization. Then we simply need to apply $a^\dagger$ repeatedly to get all the other states, labeling them as $|n\rangle \equiv (a^\dagger)^n|0\rangle$ with $H|n\rangle = n\omega|n\rangle$. (Here we've disregarded normalization, but it's easy enough to add some scaling factor in the definition of $|n\rangle$ so that $\langle n|\, m\rangle = \delta_{nm}$.)

[7] Explicitly, consider an eigenstate $|E\rangle$ with energy E. Then $Ha^\dagger|E\rangle = (a^\dagger H + \omega a^\dagger)|E\rangle = (E + \omega)a^\dagger|E\rangle$, so $a^\dagger|E\rangle$ is an eigenstate with energy $E + \omega$. The computation for a is similar.

[8] As a fun aside, the theory of angular momentum is similar, except that there for angular momentum, there is a *maximum* eigenvalue as well. In fact, angular momentum is a special example of a representation of the $SU(2)$ Lie algebra– this same structure of a ladder or lattice of states is everywhere in representation theory. See the notes for *Symmetries* for more details.

[9] Remark: gravity is different! Gravity couples directly to energy, not to differences in energy. But in a simple theory like the 1D harmonic oscillator, all we care about is the spacing of the energy levels.

That's about all there is to the quantum harmonic oscillator! We have recovered the quantized energy levels and defined operators a and $a^\dagger$ to move between them. Next time, we'll repeat the same procedure with quantum fields.

Lecture 4.

Thursday, October 11, 2018

Today, we'll introduce the second quantization procedure, which generalizes the quantum harmonic oscillator to our free scalar field. We'll find that the Hamiltonian for a free scalar field takes the form of an integral over momentum of infinitely many uncoupled harmonic oscillator Hamiltonians with some characteristic frequencies $\omega_{\mathbf{p}}$. From this Hamiltonian, we'll recover the particle interpretation of the excitations of these harmonic oscillators.

Recall that we can write the Fourier transform of a free scalar field,

$$\phi(\mathbf{x},t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}}\phi(\mathbf{p},t),$$

where the momentum-space field obeys

$$\left[\frac{\partial^2}{\partial t^2} + (\mathbf{p}^2 + m^2)\right]\phi(\mathbf{p},t) = 0.$$

We also defined $\omega_{\mathbf{p}}^2 \equiv \mathbf{p}^2 + m^2$, and observed that our theory has plane wave solutions, i.e. the field in momentum space is simply $\phi(\mathbf{p},t) = e^{i\omega_{\mathbf{p}}t}$.

Last time, we rewrote the coordinate q and momentum p in terms of creation and annihilation operators $a^\dagger, a$ as

$$\begin{aligned} q &= \frac{1}{\sqrt{2\omega}}(a + a^\dagger), \\ p &= -i\sqrt{\frac{\omega}{2}}(a - a^\dagger). \end{aligned}$$

Let's repeat this process to free fields now, defining our field ϕ and its associated conjugate momentum π to be

$$\phi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}(a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} + a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}) \tag{4.1}$$

$$\pi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\mathbf{p}}}{2}}(a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} - a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}). \tag{4.2}$$

in terms of some new creation and annihilation operators $a_{\mathbf{p}}^\dagger, a_{\mathbf{p}}$. These operators now depend explicitly on momentum, as do the characteristic "frequencies" $\omega_{\mathbf{p}}$. Note that if our quantum field theory was $0+1$-dimensional, the Fourier integral over momentum would be trivial and we would simply recover q and p from the 1D harmonic oscillator.[10]

We've therefore defined new creation and annihilation operators in order to rewrite the field and its conjugate momentum as Fourier integrals over momentum space. This process is called *second quantization*. With our new $a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger$ in hand, we now want to impose the canonical commutation relations,

$$[a_{\mathbf{p}}, a_{\mathbf{q}}] = [a_{\mathbf{p}}^\dagger, a_{\mathbf{q}}^\dagger] = 0$$

and

$$[a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] = (2\pi)^3\delta^3(\mathbf{p} - \mathbf{q}).$$

[10] The minus sign in the exponential for $a^\dagger$ is just convention, I believe. Since the $d^3\mathbf{p}$ integral is over all three-momenta and $\mathbf{p}$ is therefore just a dummy integration variable, we can certainly rewrite the integral to have the same factor $e^{i\mathbf{p}\cdot\mathbf{x}}$ in the second term. However, this choice of sign will make the canonical commutation relations manifest when we compute commutators of fields, etc.

However, in terms of the fields, we can show that the commutation relations for the operators $a_\mathbf{p}, a^\dagger_\mathbf{q}$ are actually equivalent to the field commutation relations

$$[\phi(\mathbf{x}), \phi(\mathbf{y})] = [\pi(\mathbf{x}), \pi(\mathbf{y})] = 0$$

and

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^3(\mathbf{x} - \mathbf{y}).$$

It's a good exercise to check this explicitly. For instance, we can check one way: assume the $a, a^\dagger$ commutation relations. By definition,

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{d^3\mathbf{q}}{(2\pi)^3} \frac{(-1)}{2} \sqrt{\frac{\omega_\mathbf{q}}{\omega_\mathbf{p}}} \{-[a_\mathbf{p}, a^\dagger_\mathbf{q}]e^{i\mathbf{p}\cdot\mathbf{x}-i\mathbf{q}\cdot\mathbf{y}} + [a^\dagger_\mathbf{p}, a_\mathbf{q}]e^{-i\mathbf{p}\cdot\mathbf{x}+i\mathbf{q}\cdot\mathbf{y}}\}.$$

Using the $a, a^\dagger$ commutation relations, we can rewrite their commutators as delta functions, $(2\pi)^3\delta^3(\mathbf{p} - \mathbf{q})$. We then do the integral over $\mathbf{q}$ to get

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3} \left(\frac{-i}{2}\right) \{-e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} - e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\} = i\delta^3(\mathbf{x} - \mathbf{y})$$

since $\delta^3(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}}$ and $\mathbf{p}$ is a dummy integration variable, so we can freely switch the sign in the exponent.

Now we compute H in terms of the operators $a_\mathbf{p}, a^\dagger_\mathbf{p}$ to find (after some work with δ functions which you should check) that

$$\begin{aligned}
H &= \frac{1}{2} \int d^3x \left(\pi^2 + (\nabla\phi)^2 + m^2\phi^2\right) \\
&= \frac{1}{2} \int d^3x \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{d^3\mathbf{q}}{(2\pi)^3} \left[\frac{-\sqrt{\omega_\mathbf{p}\omega_\mathbf{q}}}{2} (a_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} - a^\dagger_\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}})(a_q e^{iq\cdot x} - a^\dagger_q e^{-iq\cdot x}) \right. \\
&\quad \left. + \frac{1}{2\sqrt{\omega_\mathbf{p}\omega_\mathbf{q}}} (ipa_\mathbf{p} e^{ip\cdot x} - ipa^\dagger_\mathbf{p} e^{-ip\cdot x}) \right]
\end{aligned}$$

There's a lot of algebraic manipulation here (details in David Tong's notes) but the net result is that

$$H = \frac{1}{2} \int \frac{d^3p}{(2\pi)^3} \omega_\mathbf{p} (a_\mathbf{p} a^\dagger_\mathbf{p} + a^\dagger_\mathbf{p} a_\mathbf{p}).$$

This is simply the Hamiltonian for an infinite number of uncoupled simple harmonic oscillators with frequency $\omega_\mathbf{p}$, just as expected.

Now we can define a vacuum state $|0\rangle$ as the state which is annihilated by all annihilation operators $a_\mathbf{p}$:

$$a_\mathbf{p} |0\rangle = 0 \forall \mathbf{p}.$$

Then computing the vacuum state energy $H|0\rangle$ yields

$$\begin{aligned}
H|0\rangle &= \int \frac{d^3p}{(2\pi)^3} \omega_\mathbf{p} (a^\dagger_\mathbf{p} a_\mathbf{p} + \frac{1}{2}[a_\mathbf{p}, a^\dagger_\mathbf{p}]) |0\rangle \\
&= \frac{1}{2} \int \frac{d^3p}{(2\pi)^3} \omega_\mathbf{p} [a_\mathbf{p}, a^\dagger_\mathbf{p}] |0\rangle \\
&= \frac{1}{2} \int d^3p\, \omega_\mathbf{p} \delta^3(\mathbf{0}) |0\rangle,
\end{aligned}$$

which is infinite. Oh no!

What's happened is that $\int d^3p \left(\frac{1}{2}\omega_\mathbf{p}\right)$ is the sum of ground state energies for each harmonic oscillator, but $\omega_\mathbf{p} = \sqrt{|\mathbf{p}|^2 + m^2} \to \infty$ as $|\mathbf{p}| \to \infty$, so we call this a high-frequency or *ultraviolet divergence*. That is, at very high frequencies/short distances, our theory breaks down and we should really cut off the validity of our theory at high momentum.[11] Of course, there's another way to handle this divergence in our theory– just redefine the Hamiltonian to set the ground state energy to zero.[12]

Thus, we redefine the Hamiltonian for our free scalar field theory to be

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} a^\dagger_{\mathbf{p}} a_{\mathbf{p}},$$

such that $H|0\rangle = 0$. Nice. Subtractin' infinities. Because we're physicists.

More formally, the difference between the old and new Hamiltonians can be seen as due to an ordering ambiguity in moving from the classical theory to the quantum one, since our quantum operators (critically) do not commute. We could have written the classical Hamiltonian as

$$H = \frac{1}{2}(\omega q - ip)(\omega q + ip)$$

which is classically the same as the original simple harmonic oscillator but just becomes

$$\omega a^\dagger a$$

when we quantize.

Definition 4.3. We define a *normal ordered* string of operators $\phi_1(\mathbf{x}_1)\phi_2(\mathbf{x}_2)\ldots\phi_n(\mathbf{x}_n)$ as follows. We write colons around the operators to be normal ordered,

$$:\phi_1(\mathbf{x}_1)\phi_2(\mathbf{x}_2)\ldots\phi_n(\mathbf{x}_n):,$$

and simply move all annihilation operators to the righthand side of the expression (so all the creation operators are on the left). Note that we totally ignore commutation relations in normal ordering! Just move the operators around.[13]

Normal-ordered strings of operators are nice to work with because they make it easy to see what initial particle states will be annihilated and what final particle states will be created. We will see a theorem shortly which relates normal-ordered strings to the more physically relevant time-ordered strings of operators.

Example 4.4. For our free scalar field Hamiltonian, the normal-ordered version looks like

$$\begin{aligned} :H: &= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega_{\mathbf{p}} :(a_{\mathbf{p}}a^\dagger_{\mathbf{p}} + a^\dagger_{\mathbf{p}}a_{\mathbf{p}}): \\ &= \int \frac{d^3p}{(2\pi)^3}\omega_{\mathbf{p}} a^\dagger_{\mathbf{p}} a_{\mathbf{p}}. \end{aligned}$$

We'd like to recover particles from this theory. Recall that $\forall \mathbf{p}, a_{\mathbf{p}}|0\rangle = 0$, so $H|0\rangle = 0$ (where now H means the normal-ordered version of the Hamiltonian). It's easy to verify (exercise) that

$$[H, a^\dagger_{\mathbf{p}}] = \omega_{\mathbf{p}} a^\dagger_{\mathbf{p}}$$

and similarly

$$[H, a_{\mathbf{p}}] = -\omega_{\mathbf{p}} a_{\mathbf{p}}.$$

Let us define the state $|\mathbf{p}'\rangle = a^\dagger_{\mathbf{p}'}|0\rangle$. Then

$$H|\mathbf{p}'\rangle = \int \frac{d^3p}{(2\pi)^3}\omega_{\mathbf{p}} a^\dagger_{\mathbf{p}}[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}]|0\rangle = \omega_{\mathbf{p}'}|\mathbf{p}'\rangle.$$

Therefore $|\mathbf{p}'\rangle$ is an eigenstate of H with an energy given by $\omega_{\mathbf{p}'} = \sqrt{\mathbf{p}'^2 + m^2}$, the relativistic dispersion relation for a particle of mass m and momentum $\mathbf{p}'$. The creation operator $a^\dagger_{\mathbf{p}}$ can therefore be thought of as creating a single particle of mass m and momentum $\mathbf{p}$ when it acts on the vacuum state $|0\rangle$. Recognizing $\omega_{\mathbf{p}}$ as the energy, we'll write $E_{\mathbf{p}}$ instead of $\omega_{\mathbf{p}}$.

[11] This sort of cutoff behavior becomes especially important in the *renormalization group*, a method of studying the relationships of different field theories under special scaling transformations. We'll see this in *Statistical Field Theory*.

[12] "We're not interested in gravity, only energy differences, so we can just subtract ∞." –Ben Allanach

[13] Well, there are sign flip subtleties when we come to working with fermions because of their antisymmetrization properties, but we won't worry about them for now.

We can also define the (single-particle) momentum operator P such that

$$\mathbf{P}\,|\mathbf{p}\rangle = \mathbf{p}\,|\mathbf{p}\rangle\,.$$

P is simply the quantized version of the momentum operator from the stress-energy tensor:

$$\mathbf{P} = -\int \pi(x)\boldsymbol{\nabla}\phi(\mathbf{x})d^3x = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\mathbf{p}a^\dagger_{\mathbf{p}}a_{\mathbf{p}}.$$

Lecture 5.

Saturday, October 13, 2018

We previously found that we could write the field momentum operator (not the conjugate momentum!) as

$$\mathbf{P} = -\int \pi(x)\boldsymbol{\nabla}\phi(\mathbf{x})d^3x = \int \frac{d^3p}{(2\pi)^3}\mathbf{p}a^\dagger_{\mathbf{p}}a_{\mathbf{p}}.$$

We could also act on our momentum eigenstates with the equivalent of the angular momentum operator J^i, and what we find is that

$$J^i\,|\mathbf{p}\rangle = 0,$$

so the scalar field theory represents a spin 0 (scalar) particle.

In general we could imagine cooking up the multi-particle state

$$|\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_n\rangle = a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}\ldots a^\dagger_{\mathbf{p}_n}\,|0\rangle\,.$$

But it follows that

$$|\mathbf{p},\mathbf{q}\rangle = |\mathbf{q},\mathbf{p}\rangle\,,$$

since the creation operators for different momenta commute, $[a^\dagger_{\mathbf{p}_1},a^\dagger_{\mathbf{p}_2}] = 0$. So our states are symmetric under interchange, which means these particles are bosons. The full Hilbert space is spanned by

$$|0\rangle\,, a^\dagger_{\mathbf{p}}\,|0\rangle\,, a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}\,|0\rangle\,,\ldots$$

and this space of states is called *Fock space*.

If we use the number operator

$$N \equiv \int \frac{d^3p}{(2\pi)^3}a^\dagger_{\mathbf{p}}a_{\mathbf{p}}$$

which counts the number of particles in a state, we find[14]

$$N\,|\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle = n\,|\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle\,.$$

[14] The calculation is brief. Consider $N\,|\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle$. We rewrite the multi-particle state as $a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}\ldots a^\dagger_{\mathbf{p}_n}\,|0\rangle$, and then

$$\begin{aligned}
N\,|\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle &= \int \frac{d^3p}{(2\pi)^3}a^\dagger_{\mathbf{p}}a_{\mathbf{p}}a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}\ldots a^\dagger_{\mathbf{p}_n}\,|0\rangle \\
&= \int \frac{d^3p}{(2\pi)^3}a^\dagger_{\mathbf{p}}(2\pi^3\delta(\mathbf{p}-\mathbf{p}_1) + a_{\mathbf{p}_1}a_{\mathbf{p}})a^\dagger_{\mathbf{p}_2}\ldots a^\dagger_{\mathbf{p}_n}\,|0\rangle \\
&= |\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle + \int \frac{d^3p}{(2\pi)^3}a^\dagger_{\mathbf{p}}a^\dagger_{\mathbf{p}_1}a_{\mathbf{p}}a^\dagger_{\mathbf{p}_2}\ldots a^\dagger_{\mathbf{p}_n}\,|0\rangle \\
&= |\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle + a^\dagger_{\mathbf{p}_1}N\,|\mathbf{p}_2,\ldots,\mathbf{p}_n\rangle\,,
\end{aligned}$$

so proceeding by induction we see that for each $a^\dagger_{\mathbf{p}_i}$ we commute through, we pick up a copy of $|\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle$. When the $a_{\mathbf{p}}$ has commuted all the way to the vacuum state $|0\rangle$, it simply annihilates it, leaving behind n copies of the initial state $|\mathbf{p}_1,\ldots,\mathbf{p}_n\rangle$.

But using the commutation relations, it's easy to check that[15]

$$[N,H]=0,$$

which means that the number of particles in a given state is conserved in the free theory. Crucially, this is not true once we add interactions.

Let's also note that our momentum eigenstates are *not* localized in space. We can describe a spatially localized state by a Fourier transform,

$$|\mathbf{x}\rangle = \int \frac{d^3p}{(2\pi)^3} e^{-i\mathbf{p}\cdot\mathbf{x}} |\mathbf{p}\rangle .$$

More generally we describe a wavepacket partially localized in position and momentum space by a Fourier integral of the form

$$|\psi\rangle = \int \frac{d^3p}{(2\pi)^3} e^{-i\mathbf{p}\cdot\mathbf{x}} \psi(\mathbf{p}) |\mathbf{p}\rangle .$$

Note that neither $|\mathbf{x}\rangle$ nor $|\psi\rangle$ are eigenstates of the Hamiltonian like in QM.

We consider now relativistic normalization. We define the vacuum such that $\langle 0|\, 0\rangle = 1$, which certainly must be Lorentz invariant (1 is just a number). So in general our momentum eigenstates have the inner product

$$\langle \mathbf{p}|\, \mathbf{q}\rangle = \langle 0| [a_\mathbf{p}, a^\dagger_\mathbf{q}] |0\rangle = (2\pi)^3\delta^3(\mathbf{p}-\mathbf{q}).$$

Is this quantity Lorentz invariant? Under the Lorentz transformation, four-momenta transform as

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

We want the momentum eigenstates $|\mathbf{p}\rangle, |\mathbf{p}'\rangle$ to be related by a unitary transformation so that the inner product $\langle \mathbf{p}|\, \mathbf{q}\rangle$ is Lorentz invariant (i.e. $\langle \mathbf{p}|\, \mathbf{q}\rangle \to \langle \mathbf{p}'|\, \mathbf{q}'\rangle = \langle \mathbf{p}| U(\Lambda)^\dagger U(\Lambda) |\mathbf{q}\rangle) = \langle \mathbf{p}|\, \mathbf{q}\rangle$ by unitarity). It turns out the normalization we've chosen is not quite right.

Let us begin by claiming that

$$\int \frac{d^3p}{2E_\mathbf{p}}$$

is Lorentz invariant.

Proof. First note that the integration measure $\int d^4p$ is Lorentz invariant, since $\Lambda \in SO(1,3)$ (i.e. $\det\Lambda = 1$). Therefore the factor of $\det\Lambda$ we would normally pick up from doing the coordinate transformation is just 1, so the four-volume element is Lorentz invariant, $\int d^4p = \int d^4p'$. It's also true that the quantity $p_0^2 = \mathbf{p}^2 + m^2$ is Lorentz invariant (in particular, it expresses the length of a four-vector $p_\mu p^\mu = m^2$). The solutions for p_0 have two branches, positive and negative:

$$p_0 = \pm\sqrt{\mathbf{p}^2+m^2}.$$

But our choice of branch is also Lorentz invariant (we can't go from the positive to negative solutions via Lorentz transformation). This means that $p_0^2 - \mathbf{p}^2 - m^2, p_0 > 0$ will be a Lorentz invariant quantity, and will remain so even if we put it inside, say, a delta function. Combining the last few facts, we find that

$$\int d^4p\,\delta(p_0^2 - \mathbf{p}^2 - m^2)|_{p_0>0} = \int \frac{d^3p}{2p_0}\bigg|_{p_0=E_p} = \int \frac{d^3p}{2E_p}$$

[15] Leaving out the integral and normalization factors, the Hamiltonian is $a_\mathbf{p} a^\dagger_\mathbf{p}$, and our number operator is similarly $a^\dagger_\mathbf{q} a_\mathbf{q}$. It follows that

$$\begin{aligned}
[N,H] &\sim a^\dagger_\mathbf{q} a_\mathbf{q} a_\mathbf{p} a^\dagger_\mathbf{p} - a_\mathbf{p} a^\dagger_\mathbf{p} a^\dagger_\mathbf{q} a_\mathbf{q} \\
&= [a^\dagger_\mathbf{q}, a_\mathbf{p}] a_\mathbf{q} a^\dagger_\mathbf{p} + a_\mathbf{p} a^\dagger_\mathbf{q} a_\mathbf{q} a^\dagger_\mathbf{p} - a_\mathbf{p} a^\dagger_\mathbf{p} a^\dagger_\mathbf{q} a_\mathbf{q} \\
&= (-(2\pi)^3\delta^3(\mathbf{p}-\mathbf{q})) a_\mathbf{q} a^\dagger_\mathbf{p} + a_\mathbf{p} a^\dagger_\mathbf{q} [a_\mathbf{q}, a^\dagger_\mathbf{p}] + a_\mathbf{p} a^\dagger_\mathbf{p} a^\dagger_\mathbf{q} a_\mathbf{q} - a_\mathbf{p} a^\dagger_\mathbf{p} a^\dagger_\mathbf{q} a_\mathbf{q} \\
&= -(2\pi)^3\delta^3(\mathbf{p}-\mathbf{q}) a_\mathbf{q} a^\dagger_\mathbf{p} + (2\pi)^3\delta^3(\mathbf{q}-\mathbf{p}) a_\mathbf{p} a^\dagger_\mathbf{q}.
\end{aligned}$$

If we put the integrals back in and integrate over d^3q, the delta functions set $\mathbf{q}=\mathbf{p}$ and therefore $[N,H] \sim -a_\mathbf{p} a^\dagger_\mathbf{p} + a_\mathbf{p} a^\dagger_\mathbf{p} = 0$.

is Lorentz invariant, where we have used the fact that

$$\delta(g(x)) = \sum_{x_i \text{ roots of g}} \frac{\delta(x - x_i)}{|g'(x_i)|}$$

to perform the dp_0 integral. ⊠

We make the next claim: $2E_p\delta^3(\mathbf{p} - \mathbf{q})$ is the Lorentz invariant version of a δ-function.

Proof. As we just showed, $\int d^3p/2E_p$ is Lorentz invariant. It's also trivial to compute that

$$\int \frac{d^3p}{2E_p} 2E_p\delta^3(\mathbf{p} - \mathbf{q}) = 1.$$

But since $\int d^3p/2E_p$ is Lorentz invariant and the RHS of the equation is certainly Lorentz invariant, it follows that $2E_p\delta^3(\mathbf{p} - \mathbf{q})$ must also be Lorentz invariant. ⊠

We therefore learn that the correctly normalized states are

$$|p\rangle \equiv \sqrt{2E_p}\,|\mathbf{p}\rangle = \sqrt{2E_p}a^\dagger_{\mathbf{p}}\,|0\rangle\,,$$

(where p is now the four-vector p, not the three-vector $\mathbf{p}$) so that these momentum states have the Lorentz invariant inner product

$$\langle p|\, q\rangle = (2\pi)^3 2\sqrt{E_pE_q}\delta^3(\mathbf{p} - \mathbf{q}).$$

Note that in the basis of the old three-momentum eigenstates, we could have written the one-particle identity operator as an integral,

$$1 = \int \frac{d^3p}{(2\pi)^3}\,|\mathbf{p}\rangle\,\langle\mathbf{p}|\,.$$

We can now rewrite the 1-particle identity operator[16] as an integral over the normalized states,

$$1 = \int \frac{d^3p}{2E_p(2\pi)^3}\,|p\rangle\,\langle p|\,.$$

Free $\mathbb{C}$ scalar field We could also look at a free complex scalar field ψ, with Lagrangian

$$\mathcal{L} = \partial_\mu\psi^*\partial^\mu\psi - \mu^2\psi^*\psi.$$

We can compute the Euler-Lagrange equations varying ψ, ψ^* separately to find

$$\partial_\mu\partial^\mu\psi + \mu^2\psi = 0 \text{ and } \partial_\mu\partial^\mu\psi^* + \mu^2\psi^* = 0,$$

where the second equation is simply the complex conjugate of the first. Now we ought to write our field as a sum of two *different* creation and annihilation operators:

$$\psi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_p}}(b_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} + c^\dagger_{\mathbf{p}}e^{-i\mathbf{p}\cdot x})$$

and similarly

$$\psi^\dagger(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_p}}(b^\dagger_{\mathbf{p}}e^{-i\mathbf{p}\cdot\mathbf{x}} + c_{\mathbf{p}}e^{+i\mathbf{p}\cdot x})$$

[16] To see this really is the identity, let's act on the normalized $|q\rangle$. It's basically a one-liner:

$$\int \frac{d^3p}{2E_p(2\pi)^3}\,|p\rangle\,\langle p|\,|q\rangle = \int \frac{d^3p}{2E_p(2\pi)^3}\,|p\rangle\,\left[(2\pi)^3 2\sqrt{E_pE_q}\delta^3(\mathbf{p} - \mathbf{q})\right] = |q\rangle\,,$$

since the delta function makes the integral trivial by setting $\mathbf{p} = \mathbf{q}$.

so that the conjugate momentum to the field ψ is[17]

$$\pi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} i\sqrt{\frac{E_p}{2}}(b^\dagger_{\mathbf{p}} e^{-i\mathbf{p}\cdot\mathbf{x}} - c_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}}).$$

The conjugate momentum to $\psi^\dagger$ is equivalently $\pi^\dagger$. The commutation relations are then[18]

$$[\psi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^3(\mathbf{x}-\mathbf{y}), \quad [\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = 0$$
$$\implies [b_{\mathbf{p}}, b^\dagger_{\mathbf{q}}] = (2\pi)^3\delta^3(\mathbf{p}-\mathbf{q}) = [c_{\mathbf{p}}, c^\dagger_{\mathbf{q}}].$$

The interpretation of these equations is that different types of particle are created by the $b^\dagger_{\mathbf{p}}$ and $c^\dagger_{\mathbf{p}}$ operators. They are both spin 0 and of mass μ, so we should interpret them as a particle-antiparticle pair. This doesn't work for electrons, which have spin 1/2 and therefore require a more sophisticated spinor treatment, but it would describe something like a charged pion.

Indeed, if we compute the conserved charges in this theory by applying Noether's theorem, we get a conserved charge of the form $Q = i\int d^3x\,(\dot\psi^*\psi - \psi^*\dot\psi)$ or equivalently in terms of the conjugate momentum (since $\pi = \frac{\partial\mathcal{L}}{\partial\dot\psi} = \dot\psi^*$)

$$Q = i\int d^3x[\pi\psi - \psi^\dagger\pi^\dagger].$$

After normal ordering (exercise) one can write

$$Q = \int \frac{d^3p}{(2\pi)^3}(c^\dagger_{\mathbf{p}}c_{\mathbf{p}} - b^\dagger_{\mathbf{p}}b_{\mathbf{p}}) = N_c - N_b,$$

which shows that our conserved quantity has the interpretation of particle number (counting antiparticles as -1).

Since there are two real scalar fields in this theory, the Hamiltonian for this theory takes the form

$$H = \int \frac{d^3p}{(2\pi)^3} E_p(b^\dagger_{\mathbf{p}}b_{\mathbf{p}} + c^\dagger_{\mathbf{p}}c_{\mathbf{p}}).$$

As an exercise one can check that $[Q, H] = 0$ using the commutation relations,[19] and therefore Q is conserved. This is also true in the interacting theory. N_c, N_b are individually conserved in the free theory, but in the interacting theory they aren't– instead, they can be created and destroyed in particle-antiparticle pairs so that $N_c - N_b$ is constant.

Non-lectured aside: commutation relations and normal ordering First, let's derive the commutation relations for our new creation and annihilation operators. From the field commutation relations, we know that

$$[\psi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^3(\mathbf{x}-\mathbf{y}),$$

[17] To actually derive this expression, note that the classical conjugate momentum to ψ is $\frac{\partial\mathcal{L}}{\partial\dot\psi} = \dot\psi^*$. These fields as we've defined them only depend on space through $\mathbf{x}$, but when we add back in time dependence, a time derivative of a field will bring down factors of $\pm iE_{\mathbf{p}}$. This is more obvious when we write our fields as integrals over $e^{\pm ip\cdot x}$, where p and x are four-vectors with $p_0 = E_{\mathbf{p}}$ and $x_0 = t$. In the next lecture, we'll show that (for example) the field ψ can be rewritten as

$$\psi(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_p}}\left(b_{\mathbf{p}}e^{-ip\cdot x} + c^\dagger_{\mathbf{p}}e^{+ip\cdot x}\right),$$

so time derivatives will as promised produce a factor of $-iE_p$ for the $b_{\mathbf{p}}$ term and a factor of $+iE_p$ for the $c^\dagger_{\mathbf{p}}$ term. The signs in the exponents are correct here, since we're working in the mostly minus convention. For now, we assert that this is the correct conjugate momentum by fiat.

[18] Stated as an exercise in class. This computation is longer, see non-lectured aside.

[19] Leaving off the integrals, we have

$$[Q, H] \sim [c^\dagger_{\mathbf{p}}c_{\mathbf{p}}, b^\dagger_{\mathbf{q}}b_{\mathbf{q}}] + [c^\dagger_{\mathbf{p}}c_{\mathbf{p}}, c^\dagger_{\mathbf{q}}c_{\mathbf{q}}] - [b^\dagger_{\mathbf{p}}b_{\mathbf{p}}, b^\dagger_{\mathbf{q}}b_{\mathbf{q}}] - [b^\dagger_{\mathbf{p}}b_{\mathbf{p}}, c^\dagger_{\mathbf{q}}c_{\mathbf{q}}].$$

The commutators of bs and cs are zero since bs and cs always commute. The other two terms with only bs or only cs must cancel since b and c have the same commutation relations, so any commutators of cs and $c^\dagger$s will be equal to those same commutators with cs replaced by bs and $c^\dagger$s replaced by $b^\dagger$s everywhere. We conclude that $[Q, H] = 0$.

and if we write out this commutator explicitly in terms of the creation and annihilation operators, we find that it is

$$[\psi(\mathbf{x}), \pi(\mathbf{y})] = \frac{i}{2}\int \frac{d^3p\, d^3q}{(2\pi)^6}\Big([b_\mathbf{p}, b^\dagger_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} - [b_\mathbf{p}, c_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, b^\dagger_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})} - [c^\dagger_\mathbf{p}, c_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})}\Big). \quad (5.1)$$

Proving that the field relations hold given the creation and annihilation commutation relations is easy– if we know that $[b_\mathbf{p}, b^\dagger_\mathbf{q}] = [c_\mathbf{p}, c^\dagger_\mathbf{q}] = (2\pi)^3\delta^3(\mathbf{p}-\mathbf{q})$ and all other commutators are zero, then Eqn. 5.1 reduces to

$$\begin{aligned}
[\psi(\mathbf{x}), \pi(\mathbf{y})] &= \frac{i}{2}\int \frac{d^3p\, d^3q}{(2\pi)^3}\Big(\delta^3(\mathbf{p}-\mathbf{q})e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + \delta^3(\mathbf{q}-\mathbf{p})e^{-i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})}\Big) \\
&= i\int \frac{d^3p}{(2\pi)^3}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \\
&= i\delta^3(\mathbf{x}-\mathbf{y}).
\end{aligned}$$

Proving the other direction (b, c commutation relations given the field relations) takes a little more work. We also know that the commutator of two different fields vanishes,

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = 0.$$

Computing this commutator, we find that

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \int \frac{d^3p\, d^3q}{(2\pi)^6}\frac{1}{2\sqrt{E_\mathbf{p}E_\mathbf{q}}}\Big([b_\mathbf{p}, b^\dagger_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + [b_\mathbf{p}, c_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, b^\dagger_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, c_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})}\Big). \quad (5.2)$$

Since this integral is identically zero, the integrand must vanish. The factor $\frac{1}{2\sqrt{E_\mathbf{p}E_\mathbf{q}}}$ is nonzero for any finite values of $E_\mathbf{p}, E_\mathbf{q}$, so we learn that

$$0 = [b_\mathbf{p}, b^\dagger_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + [b_\mathbf{p}, c_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, b^\dagger_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, c_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})}. \quad (5.3)$$

This is a useful combination, since we can for instance add it to the commutator in Eqn. 5.1 to find that

$$[\psi(\mathbf{x}), \pi(\mathbf{y})] = i\int \frac{d^3p\, d^3q}{(2\pi)^6}\Big([b_\mathbf{p}, b^\dagger_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, b^\dagger_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})}\Big), \quad (5.4)$$

and since we know from the field relations that the left side is a delta function $i\delta^3(\mathbf{x}-\mathbf{y})$, we can pass back to the integral form of the delta function to find

$$\int \frac{d^3p}{(2\pi)^3}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} = \int \frac{d^3p\, d^3q}{(2\pi)^6}\Big([b_\mathbf{p}, b^\dagger_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, b^\dagger_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})}\Big),$$

or with a little forethought,

$$\int \frac{d^3p\, d^3q}{(2\pi)^3}\delta^3(\mathbf{p}-\mathbf{q})e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} = \int \frac{d^3p\, d^3q}{(2\pi)^6}\Big([b_\mathbf{p}, b^\dagger_\mathbf{q}]e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + [c^\dagger_\mathbf{p}, b^\dagger_\mathbf{q}]e^{-i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{y})}\Big).$$

Matching terms on left and right (which we can do since Fourier modes are orthogonal), we find that

$$[b_\mathbf{p}, b^\dagger_\mathbf{q}] = (2\pi)^3\delta^3(\mathbf{p}-\mathbf{q}), \quad [c^\dagger_\mathbf{p}, b^\dagger_\mathbf{q}] = 0.$$

A basically identical calculation (subtracting Eqn. 5.3 rather than adding it) yields the equivalent result for $[c_\mathbf{p}, c^\dagger_\mathbf{q}]$ and $[b_\mathbf{p}, c_\mathbf{q}]$. We find that

$$[b_\mathbf{p}, b^\dagger_\mathbf{q}] = [c_\mathbf{p}, c^\dagger_\mathbf{q}] = (2\pi)^3\delta^3(\mathbf{p}-\mathbf{q})$$

and all other commutators vanish. ☒

Next we'll show some properties of the particle number operator Q. First, normal ordering. Explicitly, we can write Q in terms of creation and annihilation operators as

$$\begin{aligned}
Q &= i\int d^3x[\pi(\mathbf{x})\psi(\mathbf{x}) - \psi^\dagger(\mathbf{x})\pi^\dagger(\mathbf{x})] \\
&= i\int d^3x \frac{d^3p\, d^3q}{(2\pi)^6}\frac{i}{2}\Big[(b_\mathbf{p}^\dagger e^{-i\mathbf{p}\cdot x} - c_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}})(b_\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{x}} + c_\mathbf{q}^\dagger e^{-i\mathbf{q}\cdot\mathbf{x}}) \\
&\qquad - (b_\mathbf{p}^\dagger e^{-i\mathbf{p}\cdot x} + c_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}})(-b_\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{x}} + c_\mathbf{q}^\dagger e^{-i\mathbf{q}\cdot\mathbf{x}})\Big] \\
&= -\int d^3x \frac{d^3p\, d^3q}{(2\pi)^6}[b_\mathbf{p}^\dagger b_\mathbf{q} e^{-i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x}} - c_\mathbf{p} c_\mathbf{q}^\dagger e^{i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x})}] \\
&= -\int \frac{d^3p\, d^3q}{(2\pi)^3}\delta^3(\mathbf{p}-\mathbf{q})[b_\mathbf{p}^\dagger b_\mathbf{q} - c_\mathbf{p} c_\mathbf{q}^\dagger] \\
&= \int \frac{d^3p}{(2\pi)^3}[c_\mathbf{p} c_\mathbf{p}^\dagger - b_\mathbf{p}^\dagger b_\mathbf{p}].
\end{aligned}$$

Applying normal ordering simply switches the c and $c^\dagger$ so that

$$:Q:= \int \frac{d^3p}{(2\pi)^3}(c_\mathbf{p}^\dagger c_\mathbf{p} - b_\mathbf{p}^\dagger b_\mathbf{p}) = N_c - N_b,$$

as desired. ⊠

Lecture 6.

Tuesday, October 16, 2018

We've been working in the Schrödinger picture where the states evolve in time, but now it will be useful to pass to the Heisenberg picture, where the states are fixed and the *operators* evolve in time.

In the Schrödinger picture, it's not obvious how our theory is Lorentz invariant. We seem to have picked out time as a special dimension when we write things down (even though we started with a Lorentz invariant theory, so our final theory should still be Lorentz invariant). The operators $\phi(\mathbf{x})$ don't depend on t, but the states evolve as

$$i\frac{d\left|p\right\rangle}{dt} = H\left|p\right\rangle = E_p\left|p\right\rangle \implies \left|p(t)\right\rangle = e^{-iE_pt}\left|p(0)\right\rangle.$$

In the Heisenberg picture, things look a bit better for covariance, since time dependence is moved into the operators. Denoting Heisenberg picture operators as O_H and Schrödinger picture operators as O_S, we have[20]

$$O_H(t) \equiv e^{iHt}O_S e^{-iHt}.$$

Taking the time derivative of each side, one finds that[21]

$$\frac{dO_H}{dt} = i[H, O_H].$$

This is the general time evolution of operators in the Heisenberg picture. It's clear that $O_H(t=0) = O_S$, so our operators agree at $t=0$ (but in general nowhere else). The field commutators then become *equal time commutation relations*:

$$[\phi(\mathbf{x},t),\phi(\mathbf{y},t)] = [\pi(\mathbf{x},t),\pi(\mathbf{y},t)] = 0$$

and

$$[\phi(\mathbf{x},t),\pi(\mathbf{y},t)] = i\delta^3(\mathbf{x}-\mathbf{y}).$$

Exercise 6.1. One should check (exercise) that $\frac{d\phi}{dt} = i[H,\phi]$ now means that the Heisenberg picture operator ϕ_H satisfies the Klein-Gordon equation, $\partial_\mu\partial^\mu\phi + m^2\phi = 0$.

[20] Here, the exponential of an operator is simply defined in terms of the series expansion of e, e.g. $e^{iHt} = \sum_{n=0}^\infty \frac{(iHt)^n}{n!}$.

[21] Explicitly, $\frac{dO_H(t)}{dt} = iHe^{iHt}O_Se^{-iHt} + e^{iHt}O_S(-iH)e^{-iHt} = ie^{iHt}[H,O_S]e^{-iHt} = i[H,O_H]$ since $e^{iHt}He^{-iHt} = H$. We also see from this computation that it doesn't matter to the Hamiltonian itself what picture we're in, since $H_S = H_H$.

We now write the Fourier transform of $\phi(x)$ (where x is now a four-vector) by deriving

$$e^{iHt}a_{\mathbf{p}}e^{-iHt} = e^{-iE_pt}a_{\mathbf{p}}$$

and

$$e^{iHt}a^{\dagger}_{\mathbf{p}}e^{-iHt} = e^{+iE_pt}a^{\dagger}_{\mathbf{p}}.$$

You should also check this (exercise) using the commutation relation $[H, a_{\mathbf{p}}] = -E_p a_{\mathbf{p}}$.

Therefore we can now write

$$\phi(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_p}}\{a_{\mathbf{p}}e^{-ip\cdot x} + a^{\dagger}_{\mathbf{p}}e^{+ip\cdot x}\}$$

where x and p are now four-vectors and $p_0 = E_p$.

Causality We might be concerned about the causal structure of this theory, since ϕ and π satisfy equal-time commutation relations. In general a Lorentz transform might mix up events which in one frame take place at "equal times." So what about arbitrary space-time separations? It turns out that causality requires that the commutators of spacelike separated operators is zero, i.e. two events which are spacelike separated cannot impact one another.

$$[O_1(x), O_2(y)] = 0 \quad \forall (x-y)^2 < 0.$$

Does this condition hold for our field operators? Let's define

$$\Delta(x-y) \equiv [\phi(x), \phi(y)]$$

and expand in the Fourier basis.

$$\begin{aligned}\Delta(x-y) &= \int \frac{d^3p}{(2\pi)^6}\frac{d^3p'}{\sqrt{4E_pE_{p'}}}\left([a_{\mathbf{p}}, a^{\dagger}_{\mathbf{p}'}]e^{-i(p\cdot x - p'\cdot y)} + [a^{\dagger}_{\mathbf{p}}, a_{\mathbf{p}'}]e^{i(p\cdot x - p'\cdot y)}\right)\\ &= \int \frac{d^3p}{2E_p(2\pi)^3}\left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right)\end{aligned}$$

Remarkably, this is just a c-number– it's not an operator at all but a (classical) number.[22] It is Lorentz invariant since the integration measure $d^3p/(2E_p)$ is Lorentz invariant and the integrand is too (it depends only on $p\cdot(x-y)$, which is the product of two four-vectors, and is therefore an invariant scalar). Moreover, each term is separately Lorentz invariant. In addition, if $x-y$ is spacelike then $x-y$ can be Lorentz transformed to $y-x$ in the first term, giving 0. It does not vanish for timelike separations, e.g.

$$[\phi(\mathbf{x}, 0), \phi(\mathbf{x}, t)] = \int \frac{d^3p}{(2\pi)^3 2E_p}(e^{-imt} - e^{+imt}) \neq 0.$$

And at equal times

$$[\phi(\mathbf{x}, t), \phi(\mathbf{y}, t)] = \int \frac{d^3p}{(2\pi)^3 2E_p}(e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} - e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}) = 0$$

(since we can send the integration variable $\mathbf{p} \to -\mathbf{p}$). One can also see in this way that the commutator for spacelike separated operators vanishes, since a general spacelike separation can always be transformed into a frame where the two events take place at equal times.

[22] Wikipedia says this terminology is due to Dirac, who coined it to contrast with q-numbers (quantum numbers), which are just operators.

Definition 6.2. We can then introduce the idea of a *propagator*– if we initially prepare a particle at point y, what is the amplitude to find it at x? We can write this as

$$\begin{aligned}
\langle 0| \phi(x)\phi(y) |0\rangle &= \int \frac{d^3p d^3p'}{(2\pi)^6\sqrt{4E_pE_{p'}}} \langle 0| a_{\mathbf{p}} a^\dagger_{\mathbf{p}'} |0\rangle e^{-ip\cdot x + ip'\cdot y} \\
&= \int \frac{d^3p d^3p'}{(2\pi)^6\sqrt{4E_pE_{p'}}} \langle 0| [a_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}] |0\rangle e^{-ip\cdot x + ip'\cdot y} \\
&= \int \frac{d^3p d^3p'}{(2\pi)^3\sqrt{4E_pE_{p'}}} \delta^3(\mathbf{p}-\mathbf{p}') e^{-ip\cdot x + ip'\cdot y} \\
&= \int \frac{d^3p}{(2\pi)^3 2E_p} e^{-ip\cdot(x-y)} \equiv D(x-y),
\end{aligned}$$

where we have used the fact that $a_{\mathbf{p}}$ kills the ground state (so we can freely replace $a_{\mathbf{p}} a^\dagger_{\mathbf{p}'}$ with the commutator $[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}]$) and used the resulting delta function to integrate over d^3p'.

In fact, one can show[23] that for spacelike separations $(x-y)^2 < 0$, the propagator decays as $D(x-y) \sim e^{-m|\mathbf{x}-\mathbf{y}|}$. The quantum field seems to "leak" out of the light cone. But we also computed that

$$\Delta(x-y) = [\phi(x), \phi(y)] = D(x-y) - D(y-x) = 0$$

if $(x-y)^2 < 0$. We can interpret this to mean that there's no Lorentz invariant way to order the two events at x and y. A particle can travel as easily from $y \to x$ as $x \to y$, so in a quantum measurement these two amplitudes cancel. With a complex scalar field, the story is more interesting. We find instead that the amplitude for a particle to go from $x \to y$ is cancelled by the amplitude for an anti-particle to go from $y \to x$.[24] This is also the case for the real scalar field, except the particle is its own antiparticle.

Definition 6.3. We now introduce the *Feynman propagator* Δ_F, which is like a regular propagator but with time ordering baked in. That is,

$$\Delta_F = \begin{cases} \langle 0| \phi(x)\phi(y) |0\rangle & \text{for } x^0 > y^0 \\ \langle 0| \phi(y)\phi(x) |0\rangle & \text{for } y^0 > x^0. \end{cases}$$

We claim the Feynman propagator can also be written as

$$\Delta_F = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip\cdot(x-y)}.$$

Note that this is Lorentz invariant– the volume element is certainly Lorentz invariant, and everything else is scalars. But there's an issue– this integral has a pole whenver $p^2 = m^2$, or equivalently for each value of $\mathbf{p}$, $p^2 - m^2 = (p^0)^2 - \mathbf{p}^2 - m^2 = 0$ when $p^0 = \pm E_{\mathbf{p}} = \pm\sqrt{\mathbf{p}^2 + m^2}$. We would like to integrate over p^0 to recover the earlier form of the propagator, so we can either deform the contour or push the poles of the real p^0 axis with an *$i\epsilon$ prescription*.

We'll finish the proof next time, but by analytically continuing p^0 to the complex plane, making this $i\epsilon$ prescription, and closing the contour appropriately we can do the p^0 integral and find that what we get is exactly the Feynman propagator as defined earlier in terms of time ordering.

Proof of Exercise 6.1 Let's find the equation of motion for ϕ. Recall that $[\phi(\mathbf{x}), \phi(\mathbf{y})] = 0$. We can also show that $\boldsymbol{\nabla}\phi(y)$ and $\phi(x)$ commute:

$$\boldsymbol{\nabla}\phi(\mathbf{y})\phi(\mathbf{x}) = \boldsymbol{\nabla}_y(\phi(\mathbf{y})\phi(\mathbf{x})) = \boldsymbol{\nabla}_y(\phi(\mathbf{x})\phi(\mathbf{y})) = \phi(\mathbf{x})\boldsymbol{\nabla}\phi(\mathbf{y})$$

[23] The easiest way to do this is to set $y = 0$ and take x and y at equal times, $x^0 = y^0 = 0$. This gets rid of p^0, and from here you can switch to spherical coordinates, rewriting $\mathbf{p}\cdot(x)$ as $|p||x|\cos\theta$.

[24] See also Wheeler's "one-electron universe"– `https://en.wikipedia.org/wiki/One-electron_universe`.

so the only term in the Hamiltonian we need to worry about is the π^2 term.

$$\begin{aligned}
\dot{\phi} = i[H,\phi] &= \frac{i}{2}\int d^3y\left[\pi^2(y) + (\nabla\phi(y))^2 + m^2\phi(y)^2, \phi(x)\right] \\
&= \frac{i}{2}\int d^3y(\pi^2(y)\phi(x) - \phi(x)\pi^2(y)) \\
&= \frac{i}{2}\int d^3y(\pi(y)(-[\phi(x),\pi(y)] + \phi(x)\pi(y)) - \phi(x)\pi^2(y)) \\
&= \frac{i}{2}\int d^3y(-i\delta^3(x-y)\pi(y) + \pi(y)\phi(x)\pi(y) - \phi(x)\pi^2(y) \\
&= \frac{i}{2}\int d^3y(-2i\delta^3(x-y)\pi(y)) \\
&= \pi(x).
\end{aligned}$$

We can also compute the time evolution for π. Here, we do have to worry about the $\boldsymbol{\nabla}\phi$ terms as well as the ϕ terms.

$$\begin{aligned}
\dot{\pi} = i[H,\pi] &= \frac{i}{2}\int d^3y\left[\pi^2(y) + (\nabla\phi(y))^2 + m^2\phi(y)^2, \pi(x)\right] \\
&= \frac{i}{2}\int d^3y\nabla\phi(y)\nabla_y(\phi(y)\pi(x)) - \nabla_y(\pi(x)\phi(y))\nabla\phi(y) + 2im^2\delta^3(x-y)\phi(y) \\
&= \frac{i}{2}\int d^3y\nabla\phi(y)\nabla_y([\phi(y),\pi(x)]) - \nabla_y(-[\phi(y),\pi(x)])\nabla\phi(y) + 2im^2\delta^3(x-y)\phi(y) \\
&= \frac{i}{2}\int d^3y\nabla\phi(y)\nabla_y(i\delta^3(y-x)) + \nabla_y(i\delta^3(y-x))\nabla\phi(y) + 2im^2\delta^3(x-y)\phi(y) \\
&= \frac{i}{2}\int d^3y\left(-2i\delta^3(x-y)\nabla^2\phi(y) + 2im^2\delta^3(x-y)\phi(y)\right) \\
&= \nabla^2\phi - m^2\phi.
\end{aligned}$$

(where we have integrated by parts to move the ∇ from the delta function to ϕ). Thus ϕ obeys the equation

$$\ddot{\phi} = \dot{\pi} = \nabla^2\phi - m^2\phi$$

or equivalently

$$\ddot{\phi} - \nabla^2\phi + m^2 = \partial_\mu\partial^\mu\phi + m^2 = 0.$$

Therefore ϕ satisfies the Klein-Gordon equation. (This is also in David Tong's notes.) ☒

We'll also make note of a potentially useful identity which can be proved by induction: if $[a,b] = \alpha$, then $[a^n, b] = n\alpha a^{n-1}$.

Proof of Heisenberg picture $a_p, a_p^\dagger$ Here, we'll show that

$$e^{iHt}a_p e^{-iHt} = e^{-iE_pt}a_p$$

using the commutation relation $[H, a_p] = -E_p a_p$. First, I'll claim that

$$H^n a_p = a(-E_p + H)^n.$$

Let's prove it by induction: for the base case, $n = 1$ and

$$Ha_p = [H, a_p] + a_pH = -E_pa_p + a_pH = a_p(-E_p + H).$$

Now the inductive step: suppose the hypothesis holds for n. Then

$$H^{n+1}a_p = H(H^na_p) = Ha_p(-E_p + H)^n = a_p(-E_p + H)^{n+1}.$$

Therefore we can use this in the expansion of e^{iHt}.

$$\begin{aligned} e^{iHt} a_p e^{-iHt} &= \sum_{n=0}^{\infty} \frac{(iHt)^n}{n!} a_p e^{-iHt} \\ &= a_p \sum_{n=0}^{\infty} \frac{(it(-E_p + H))^n}{n!} e^{-iHt} \\ &= a_p e^{-iE_p t} e^{iHt} e^{-iHt} \\ &= a_p e^{-iE_p t}. \end{aligned}$$

Rather than repeating this whole calculation, we can simply take the hermitian conjugate of each side (since H is hermitian) to get

$$e^{iHt} a_p^\dagger e^{-iHt} = e^{+iE_p t} a_p^\dagger.$$

Note that the sign flip in the exponent of $e^{\pm iHt}$ and the reversing of order from taking the hermitian conjugate cancel out. So the operators $a, a^\dagger$ do evolve in a nice way that allows us to write ϕ in terms of a four-vector product in the exponent, $p \cdot x$, and in turn this helps us to see that our theory has a sensible causal structure under Lorentz transformations. ⊠

Lecture 7.

Thursday, October 18, 2018

Today, we'll complete our initial discussion of propagators and then introduce interacting fields.

Last time, we claimed the Feynman propagator could be written as an integral over d^4p, and reduces to the regular propagator $D(x-y)$ or $D(y-x)$ depending on the sign of $x^0 - y^0$. The propagator $D(x-y)$ was an integral over d^3p only, so we need to integrate over the p^0 component. To evaluate the p^0 integral, one can make an $i\epsilon$ prescription and modify the pole to

$$\Delta_F = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)}$$

with $\epsilon > 0$ and small. This helps us to keep track of which pole is inside our contour, but we can also equivalently shift the contour (see picture). This shifts the pole at E_p to $E_p - i\epsilon$ and from $-E_p$ to $-E_p + i\epsilon$. This is a little quick, so I'll work it out more carefully in a footnote later.

Which way we close the contour depends on the sign of $x^0 - y^0$ since $(x^0 - y^0) > 0$ means that $e^{ip^0(x^0-y^0)} \to 0$ when $p^0 \to +i\infty$, and for $(x^0 - y^0) < 0$ it goes to 0 when $p^0 \to -i\infty$.

In any case, we can evaluate this with the Cauchy integral formula to find

$$\Delta_F(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} e^{-ip\cdot(x-y)} = D(x-y)$$

for $x^0 > y^0$ and

$$\Delta_F(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} e^{-ip\cdot(y-x)} = D(y-x)$$

for $y^0 > x^0$, where the sign flip has come from which way we close the contour and therefore which pole we pick up in the integration.

We can now observe that Δ_F is the *Green's function* of the Klein-Gordon equation. A Green's function (perhaps familiar from a class on PDEs or electrodynamics) is simply the inverse of a differential operator; it is a function which yields a delta function when you hit it with a given differential operator. You might have seen the Green's function for Poisson's equation, for instance.[25] In this case,

$$\begin{aligned}(\partial_t^2 - \nabla^2 + m^2)\Delta_F(x-y) &= \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2+i\epsilon}(-p^2+m^2)e^{-ip\cdot(x-y)} \\ &= -i\int \frac{d^4p}{(2\pi)^4}e^{-ip\cdot(x-y)} \\ &= -i\delta^4(x-y).\end{aligned}$$

It can be useful to choose other integration contours, e.g. for the retarded propagator which takes

$$\Delta_R(x-y) = \begin{cases}[\phi(x),\phi(y)] & : x^0 > y^0 \\ 0 & : y^0 > x^0\end{cases}$$

The advanced propagator is similarly defined but for $x^0 < y^0$. In any case, the Feynman propagator is the most applicable for our purposes.

Interacting fields Our free theories have made for nice, exactly solvable models. They have Lagrangians which are quadratic in the fields, which means that

- the equations of motions are linear
- we have exact quantization
- we can produce multi-particle states, but there is no scattering.

It's this third point which is not realistic– we know in general that particles should interact and scatter. Therefore, we guess that interactions must come from higher-order terms in the Lagrangian $\mathcal{L}$. For example, in a real scalar field ϕ we could more generally write

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \sum_{n=3}\frac{\lambda_n}{n!}\phi^n,$$

where the λ_ns are called *coupling constants*. Ideally, we'd like these corrections to be small so we can take a perturbative expansion about the free theory solutions, which already look like particles.

Naïvely, we might say that small perturbations means that $\lambda_n \ll 1$, but that only makes sense when λ_n is dimensionless. So let's do some dimensional analysis to figure out what the dimensions of λ_n are. Recall that the action S is dimensionless, $[S] = 0$. Since $S = \int d^4x\mathcal{L}$ and $[d^4x] = -4$, we find that $[\mathcal{L}] = 4$. From looking at the kinetic term $\partial_\mu\phi\partial^\mu\phi$ and using the fact that $[\partial_\mu] = +1$, we conclude that $[\phi] = 1$, $[m] = 1$, and

$$[\lambda_n] = 4 - n$$

(where this 4 comes from the fact we are working in $3+1$ spacetime dimensions).

What we discover is that there are three important cases here:

(a) $[\lambda_3] = 1$. The dimensionless parameter is λ_3/E, where E is the energy scale of the process of interest (e.g. the scattering energy, on the order of TeV at the LHC). If $\lambda_3/E \ll 1$, then $\lambda_3\phi^3/3!$ is a small perturbation at high energies. We call this a *relevant perturbation* because it is important at low energies. In a relativistic setting, $E > m$ so we can make the perturbation small by taking $\lambda_3 \ll m$. We call this class of theories with positive mass dimension coupling constants *renormalizable*, meaning that we can reasonably deal with the infinities which crop up from weak coupling.

(b) $[\lambda_4] = 0$. Here, $\lambda_4\phi^4/4!$ is small if $\lambda_4 \ll 1$. We call these *marginal* couplings, and these are also renormalizable.

(c) $[\lambda_n] = 4 - n$ for $n \geq 5$. These are called *irrelevant* couplings. The dimensionless parameter is $\lambda_n E^{n-4}$, and they are small at low energies but large at higher energies. These lead to non-renormalizable theories, where the infinities are bigger and scarier and we cannot sweep them under the rug by just subtracting off infinity.

[25] Green's functions are useful because they allow us to easily fit the boundary conditions. Consider the operator equation $\hat{O}\psi(x) = f(x)$ for some differential operator $\hat{O}$ and some given function $f(x)$. If we could just write down $\hat{O}^{-1}$, it would be easy enough to solve any equation of this form: $\psi(x) = \hat{O}^{-1}f(x)$. This is sort of what Green's functions let us do. If we know that $\hat{O}\Delta(x-y) = \delta(x-y)$, it follows that $\hat{O}\left[\int dy\Delta(x-y)f(y)\right] = \int dy\delta(x-y)f(y) = f(x)$ (where any derivatives in $\hat{O}$ are taken with respect to x), so $\int dy\Delta(x-y)f(y) = \psi(x)$ solves the differential equation.

On the one hand, the nature of irrelevant couplings means that we can describe (relatively) low-energy physics well by only looking at the first few terms in the perturbative expansion, but it also makes it very difficult to probe very high-energy physics (for instance, on the scale of quantum gravity).

Example 7.1. Let's consider ϕ^4 theory, with the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda\phi^4}{4!}; \lambda \ll 1.$$

We can already guess at the effects of this final term– in particular, $[H, N] \neq 0 \implies$ particle number is no longer conserved.[26] Expanding the last term, we expect some big integrals which will have terms like

$$\int \ldots ((a^\dagger_{\mathbf{p}})^4 \ldots) + \int \ldots a^{\dagger\,3}_{\mathbf{p}} a_{\mathbf{p}} + \ldots$$

which will destroy particles.

Example 7.2. We could also consider scalar Yukawa theory with two fields, $\psi \in \mathbb{C}, \phi \in \mathbb{R}$, and the Lagrangian

$$\mathcal{L} = \partial_\mu\psi^*\partial^\mu\psi + \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \mu^2\psi^*\psi - \frac{1}{2}m^2\phi^2 - g\psi^*\psi\phi.$$

In this theory, $[g] = 1$ and we take $g \ll m, g \ll \mu$. We get a Noether current by noticing that the Lagrangian is invariant under $\psi \to e^{i\theta}\psi$, and this current has the interpretation of charge conservation– the number of ψ particles minus the number of ψ anti-particles is conserved, but there is no such conservation law for the number of real scalar ϕs.

The interaction picture Previously, we saw the familiar Schrödinger picture where operators are time-independent and states evolve in time by the Schrödinger equation,

$$i\frac{d}{dt}\left|\psi\right\rangle_S = H\left|\psi\right\rangle_S.$$

We then introduced the Heisenberg picture, where we moved the explicit time dependence into the operators,

$$\left|\psi\right\rangle_H = e^{iHt}\left|\psi\right\rangle_S, O_H(t) = e^{iHt}O_S e^{-iHt}.$$

The interaction picture is a hybrid of the Heisenberg and Schrödinger pictures. It splits the Hamiltonian into a free theory part and an interaction part:

$$H = H_0 + H_{\text{int}}.$$

In the interaction picture, states evolve with the interacting Hamiltonian H_{int} and operators evolve by the free Hamiltonian H_0.

Example 7.3. In ϕ^4 theory, we have $\mathcal{L}_{int} = -\lambda\phi^4/4!$ with

$$H_{int} = -\int d^3x\mathcal{L}_{int} = +\lambda\int \phi^4/4!$$

and H_0 the standard free theory Hamiltonian $H_0 = \int d^3x\frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2$.

Non-lectured supplement: contour integration and the p^0 integral If you haven't seen contour integration before, it's basically an integration technique for certain real integrals which makes use of a theorem called the Cauchy residue theorem. I'll use some different notation here (ks instead of ps and ω_k instead of E_p), but all the physics is the same. I'm also setting $y = 0$ here since Δ_F only depends on the combination $|x - y|$.

[26] Those of you with some familiarity with Feynman diagrams can probably cook up a simple diagram which goes from one to three particles using the ϕ^4 interaction. The interaction has four lines so just put one on the left and three on the right (no need to worry about antiparticles since this is a scalar field).

Cauchy came up with a nice formula which says that if a function $f(z)$ is analytic[27] on and inside a simple[28] closed curve C, then the value of the following contour integral[29] along C is given by

$$\oint \frac{f(z)}{z-a} dz = 2\pi i f(a)$$

for $z = a$ a point inside C.

Mathematicians usually write this as a formula for the value of $f(a)$ in terms of the contour integral, but for our purposes it is more useful as a formula for the integral. The proof is not complicated and fits on a page or two (see for instance Boas Mathematical Methods 585-586 or `http://mathworld.wolfram.com/CauchyIntegralFormula.html`) but I will not repeat it here.

What's the practical use of this formula? Essentially, we can use it to compute real integrals which might have poles (singular points) along the integration path. Consider our expression for the propagator, and suppose $\varepsilon = 0$. Then the denominator becomes

$$k^2 - m^2 = (k^0)^2 - \mathbf{k}^2 - m^2 = (k^0)^2 - \omega_k^2$$

and written this way, it is clear that the integrand is going to become singular at $k^0 = \pm\omega_k$. Therefore, we make an "$i\varepsilon$ prescription," meaning that we add $i\varepsilon$ ($\varepsilon > 0$ and small) to push the poles off the real line into the complex plane so we can do the integral, and hope nothing bad happens as we let ε go to zero.

We'll need one more trick to compute this integral. You might have noticed that our integral isn't a closed curve yet (as required by the Cauchy formula)– it is an integral $\int_{-\infty}^{\infty} dk^0$. Therefore, we must close the contour by adding a curve whose final contribution to the overall integral will be zero. To warm up, suppose we want to compute

$$\int_{-\infty}^{\infty} dz \frac{e^{iz}}{z - iz_0}$$

for z_0 possibly complex. We can close the contour by adding a curve in the upper half-plane, "out at $+i\infty$." See Figure 1 for an illustration.

How do we decide whether to close the contour in the upper or lower half-plane? Notice that in the upper half-plane, $z = x + iy$ for $y > 0$, so $e^{iz} = e^{i(x+iy)} = e^{-y}e^{ix}$ with $y > 0$. Therefore, e^{iz} is exponentially damped in the upper half-plane and contributes basically zero to the overall integral. So we can close the curve "for free" and write

$$\int_{-\infty}^{\infty} dz \frac{e^{iz}}{z - z_0} = \oint dz \frac{e^{iz}}{z - z_0} = 2\pi i e^{iz_0}$$

if z_0 has imaginary part > 0 (is inside the contour) and 0 otherwise. Thanks, Cauchy integral formula.

In fact, the formula lends itself to an even better generalization, the *Cauchy residue theorem*. It states that

$$\oint_C f(z) dz = 2\pi i \cdot \text{sum of the residues of } f(z) \text{ inside } C,$$

where the integral around C is in the counterclockwise direction, and a *residue* is basically the value at the function at the pole if it didn't have that pole. Quick example: for the function $f(z) = \frac{z}{(1+z)(3-z)}$, $f(z)$ has a pole at $z = 3$. The residue $R(3)$ of $f(z)$ at $z = 3$ is simply $R(5) = \frac{z}{1+z}|_{z=3} = \frac{3}{4}$.

So to summarize, close the contour based on what the integrand is doing at $\pm i\infty$. Check which poles are inside your contour, and plug up the singularities one at a time to compute the residues. Sum up the residues, multiply by $2\pi i$, and you've got the value of your integral.

Returning to the problem at hand, we wish to compute the integral

$$\int dk^0 \frac{e^{ik^0 x^0}}{(k^0)^2 - (\omega_k^2 - i\varepsilon)}.$$

[27] not singular

[28] does not cross itself

[29] A fancy name for a closed line integral in the complex plane.

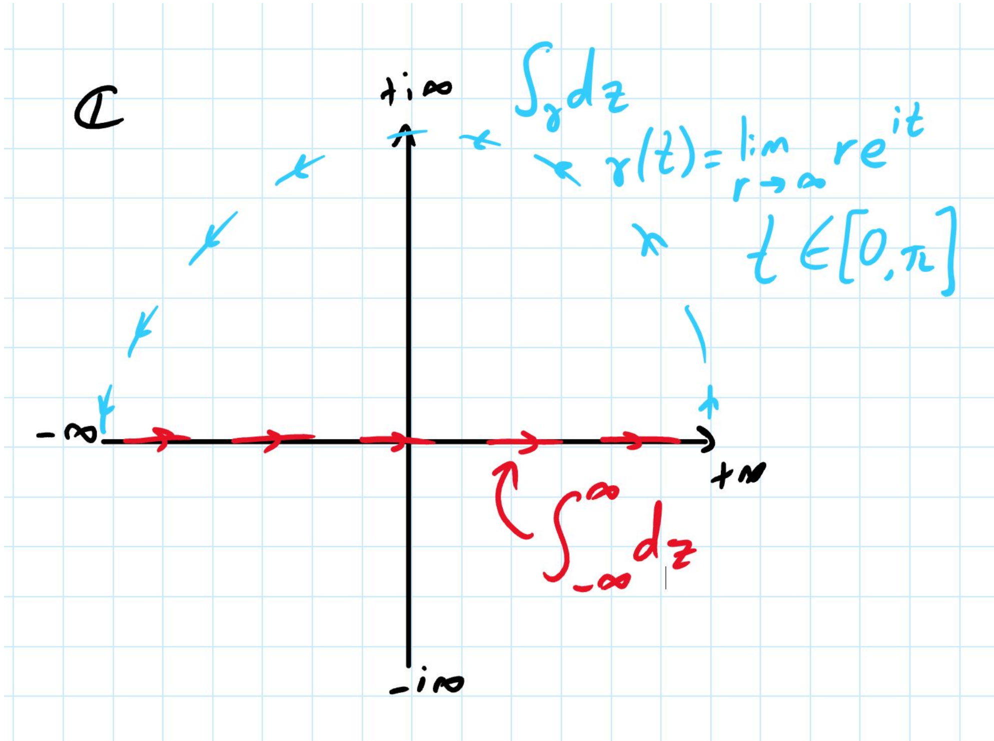

FIGURE 1. An illustration of closing the contour for a real integral (in red) so we can use the Cauchy integral formula to perform the integral. The integral along the light blue curve $\int_\gamma dz$ contributes nothing to the overall integration, so we can use it to close the contour out at $+i\infty$. Note the curve runs counterclockwise. If it ran clockwise, we would pick up a minus sign.

We have two poles at $\pm\sqrt{\omega_k^2 - i\varepsilon}$, which in the $\varepsilon \to 0$ limit become $+\omega_k - i\varepsilon$ and $-\omega_k + i\varepsilon$. Therefore, we rewrite as

$$\int_{-\infty}^{\infty} dk^0 \frac{e^{ik^0x^0}}{(k_0 - (\omega_k - i\varepsilon))(k_0 - (-\omega_k + i\varepsilon))}.$$

Since $e^{ik^0x^0}$ is exponentially damped in the upper half-plane, we close the contour at $+i\infty$, enclosing the pole at $-\omega_k + i\varepsilon$ (recall ω_k is real and ε is positive). Therefore, calculating the residue, this integral comes out to

$$2\pi i \frac{e^{-i\omega_k x^0}}{-2\omega_k}$$

(letting $\varepsilon \to 0$) and we conclude that for $x^0 > 0$,

$$\Delta_F(x) = -i \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-i(\omega_k t - \mathbf{k}\cdot\mathbf{x})}.$$

A similar calculation holds for $x^0 < 0$, so we recover our friend the Feynman propagator, which correctly accounts for the sign of x^0.

Lecture 8.

Saturday, October 20, 2018

Today, we'll take our first look at interacting theories in detail! Let's first complete our description of the interaction picture. Operators in the interaction picture evolve in time by the free Hamiltonian:

$$O_I(t) \equiv e^{iH_0t} O_S e^{-iH_0t},$$

with O_S the Schrödinger picture operator, while states evolve by

$$|\psi(t)\rangle_I = e^{iH_0 t} |\psi(t)\rangle_S = e^{-iH_{int}t} |\psi(0)\rangle_S .$$

Note that the interacting Hamiltonian also has an interaction picture counterpart,

$$H_I \equiv (H_{int})_I = e^{iH_0 t} H_{int} e^{-iH_0 t}. \tag{8.1}$$

In the context of our quantum fields,

$$\phi_I(x) = e^{iH_0 t}\phi(\mathbf{x})e^{-iH_0 t}$$

so that the interaction picture field ϕ_I obeys the Klein-Gordon equation

$$(\partial^2 + m^2)\phi_I = 0,$$

with solution

$$\phi_I(x) = \int \frac{d^3 p}{(2\pi)^3 \sqrt{2E_p}} (a_{\mathbf{p}} e^{-ip\cdot x} + a^\dagger_{\mathbf{p}} e^{ip\cdot x}).$$

Here, note that we're taking the four-vector inner product $p \cdot x$ as in the Heisenberg picture, with $p^0 = E_p$ and $x^0 = t$. We also see that $\phi_I(t = 0, \mathbf{x}) = \phi_S(\mathbf{x})$, so the fields at $t = 0$ agree with the Schrödinger picture fields.

As before,

$$[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}] = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'),$$

with other brackets vanishing. Note that the state $|0\rangle$ (satisfying $a_{\mathbf{p}} |0\rangle = 0$) is the vacuum of the free theory, not the interacting theory. This means we will have to be a little careful when we compute the transition amplitudes between states, since they are measured relative to vacuum fluctuations, which will start bubbling as soon as we turn on interactions.

As operators, interaction picture fields are related to the Heisenberg picture ones by

$$\phi_H(t, \mathbf{x}) = e^{iHt} e^{-iH_0 t} \phi_I(x) e^{iH_0 t} e^{-iHt},$$

where $e^{-iH_0 t}\phi_I(x)e^{iH_0 t} = \phi_S(\mathbf{x})$. We can also regroup the operators here to write

$$\phi_H(t, \mathbf{x}) = U(t, 0)^\dagger \phi_I(t, \mathbf{x}) U(t, 0),$$

where

$$U(t, t_0) \equiv e^{iH_0 t} e^{-iH(t-t_0)} e^{-iH_0 t_0} \tag{8.2}$$

is a unitary time evolution operator. U is defined such that[30]

$$U(t_1, t_2)U(t_2, t_3) = U(t_1, t_3)$$

and $U(t, t) = 1$. Equivalently, we see that[31]

$$|\psi(t)\rangle_I = U(t, t') |\psi(t')\rangle_I .$$

That is, U evolves interaction picture states in time.

Now, actually computing U in terms of operator exponentials would be a pain, especially since we might have multiple fields doing their thing (creating and destroying particles) at different points in time. Based

[30]We can verify this first property by a quick computation:

$$\begin{aligned} U(t_1, t_2)U(t_2, t_3) &= \left(e^{iH_0 t_1} e^{-iH(t_1-t_2)} e^{-iH_0 t_2}\right)\left(e^{iH_0 t_2} e^{-iH(t_2-t_3)} e^{-iH_0 t_3}\right) \\ &= e^{iH_0 t_1} e^{-iH(t_1-t_2)} e^{-iH(t_2-t_3)} e^{-iH_0 t_3} \\ &= e^{iH_0 t_1} e^{-iH(t_1-t_3)} e^{-iH_0 t_3} = U(t_1, t_3). \end{aligned}$$

Note that H and H_0 are by no means guaranteed to commute, so we cannot naïvely group them together in an exponent, i.e. $e^{iH_0 t}e^{-iHt} \neq e^{i(H_0 - H)t}$. Operator exponentials are different. In doing this calculation, we were only allowed to add the exponents when the operator in the exponent was the same in both terms.

[31]If this isn't obvious, note that since $|\psi(t)\rangle_S = e^{-iHt} |\psi(0)\rangle_S$, it follows that interaction picture states evolve as

$$|\psi(t)\rangle_I = e^{iH_0 t} |\psi(t)\rangle_S = e^{iH_0 t} e^{-iHt} |\psi(0)\rangle_S = U(t, 0) |\psi(0)\rangle_S = U(t, 0) |\psi(0)\rangle_I ,$$

since the states agree independent of picture at $t = 0$. Applying the property that $U(t_1, t_2)U(t_2, t_3)$, we see that $|\psi(t)\rangle_I = U(t, t') |\psi(t')\rangle_I$ for general t, t'. So U really does have the function of being a time evolution operator on interaction picture states.

on the properties of U we have defined, can we find a more tractable expression for the time evolution operator?

By differentiating our expression for U with respect to time, we see that

$$\begin{aligned} i\frac{dU(t,0)}{dt} &= i\left[iH_0 e^{iH_0 t}e^{-iHt} + e^{iH_0 t}(-iH)e^{-iHt}\right] \\ &= e^{iH_0 t}(H-H_0)e^{-iHt} \\ &= e^{iH_0 t}(H_{int})_S e^{-iH_0 t}e^{iH_0 t}e^{-iHt} \\ &= H_I(t)U(t,0), \end{aligned}$$

which tells us that the operator U obeys the equivalent of the Schrödinger equation.

If $(H_{int})_I = H_I$ were just a function, we could solve this by $U = \exp[-i\int_{t_0}^t H_I(t')dt']$. However, because H_i is an operator, life is not so simple, as we have ordering ambiguities. The issue becomes clear when we write out the first few terms in the exponential:

$$\exp[-i\int_{t_0}^t H_I(t')dt'] = 1 - i\int_{t_0}^t H_I(t')dt' + \frac{(-i)^2}{2!}\left(\int_{t_0}^t H_I(t')dt'\right)^2. \tag{8.3}$$

If we take the time derivative, Leibniz tells us that this quadratic term becomes

$$-\frac{1}{2}\int_{t_0}^t H_I(t')dt' H_I(t) - \frac{1}{2}H_I(t)\int_{t_0}^t H_I(t')dt'. \tag{8.4}$$

But this first term is a problem since the $H_I(t)$ is on the wrong side of the integral and we can't commute it through because $[H_I(t'), H_I(t'')] \neq 0$ for $t' \neq t''$.

However, our differential equation for U tells us that a solution for $U(t,t_0)$ is given by[32]

$$U(t,t_0) = 1 + (-i)\int_{t_0}^t dt' H_I(t')U(t',t_0).$$

Therefore we can substitute this expression for $U(t,t_0)$ back into itself to get the infinite series

$$U(t,t_0) = 1 + (-i)\int_{t_0}^t dt' H_I(t') + (-i)^2\int_{t_0}^t dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'') + \dots.$$

From the ranges of integration, it's clear that the H_Is are automatically time-ordered– for instance, $H_I(t'')$ always takes place at $t'' \leq t'$. Note that we could have rewritten this last term as

$$\begin{aligned} \int_{t_0}^t dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'') &= \int_{t_0}^t dt''\int_{t_0}^{t''} dt' H_I(t'')H_I(t') \\ &= \int_{t_0}^t dt'\int_{t'}^{t} dt'' H_I(t'')H_I(t'), \end{aligned}$$

where in the first line, the range of integration is $t' \leq t''$, while in the second it is $t'' \geq t'$. Note that this expression is time-ordered too, so $H_I(t'')H_I(t') = T[H_I(t')H_I(t'')]$ for these limits of integration. It follows that the quadratic term can be written

$$\begin{aligned} (-i)^2\int_{t_0}^t dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'') &= \frac{(-i)^2}{2}\left[\int_{t_0}^t dt'\int_{t_0}^{t'} dt'' T[H_I(t')H_I(t'')] + \int_{t_0}^t dt'\int_{t'}^{t} dt'' T[H_I(t')H_I(t'')]\right] \\ &= \frac{(-i)^2}{2}\int_{t_0}^t dt'\int_{t_0}^{t} dt'' T[H_I(t')H_I(t'')]. \end{aligned}$$

We can play the same game for higher-order terms, and we'll get a symmetry factor of $n!$ to a term with n copies of $H_I(t')$. With the same limits of integration on dt', dt'', etc., these higher-order terms look a lot like the power series expansion of an exponential. This leads us to make the following definition.

[32] Easy to check. By the fundamental theorem of calculus (what a throwback), $i\frac{d}{dt}\left[1 + (-i)\int_{t_0}^t dt' H_I(t')U(t',t_0)\right] = H_I(t)U(t,t_0)$ and $U(t_0,t_0) = 1 + (-i)\int_{t_0}^{t_0} dt' H_I(t')U(t',t_0) = 1$, so it satisfies the boundary conditions.

Definition 8.5. Using time ordering, we find that U can be written compactly as

$$U(t,t_0) = T\exp\left\{-i\int_{t_0}^{t} dt' H_I(t')\right\},$$

which we call *Dyson's formula.* (Note that $U(t,t_0) = T\exp\{+i\int_{t_0}^{t} dt' L_I(t')\}$, in terms of the Lagrangian.) This is a formal result, but we usually just expand to some finite order in terms of the coupling constants which live in the interacting Hamiltonian H_I.

This is the last bit of machinery we need to start computing scattering amplitudes in quantum field theory!

Definition 8.6. The time evolution used in scattering theory is called the *S-matrix* (S for scattering). The S matrix is defined to be

$$S = \lim_{t\to\infty, t_0\to-\infty} U(t,t_0).$$

We will consider interactions where the final state $|f\rangle$ and the initial state $|i\rangle$ are well-separated from each other and are far away from the interaction. Therefore, the initial and final states $|i\rangle, |f\rangle$ behave like free particles, i.e. they are eigenstates of H_0.[33]

This should seem at least plausible: at late/early times, the particles are well-separated and don't feel the effect of each other. As they approach, they may interact before going their separate ways. The scattering amplitude is then

$$\lim_{t\to\infty, t_0\to-\infty} \langle f| U(t,t_0) |i\rangle = \langle f| S |i\rangle .$$

Note that there are some cases that need to be treated differently, like bound states. For instance, a proton and an electron at low energies could interact to form a hydrogen atom, $p + e^- \to$ the bound state (H). Here, the assumption that the particles end up well-separated is violated. It turns out that such solutions appear as poles in the S-matrix, but this is a more advanced topic and we won't discuss it further here.

Let's return to scalar Yukawa theory. Now, we'll drop the I subscripts and assume uniformly that we are in the interaction picture. Discarding the kinetic and mass terms from the free theory, we are left with the interaction Hamiltonian

$$\mathcal{H} = g\psi^*\psi\phi,$$

where ψ and ψ^* are (anti-)nucleons (e.g. a proton or neutron), and ϕ is a meson. What do each of these fields do?

- ϕ has a and $a^\dagger$ terms which destroy and create mesons, respectively.
- ψ has b and $c^\dagger$ terms, where b destroys a nucleon and $c^\dagger$ creates an anti-nucleon.
- ψ^* has $b^\dagger$ and c terms, where $b^\dagger$ creates a nucleon and c destroys an anti-nucleon.

Looking at the possible terms in the Hamiltonian, we can already see interesting behavior– we'll have terms where nucleon-anti-nucleon pairs are created and destroyed, e.g. $b^\dagger c^\dagger a$ which destroys a meson and produces a nucleon-anti-nucleon pair. This contributes to meson decay, $\phi \to \psi\bar{\psi}$.[34] What we recover is the leading order in g term in the S-matrix, and an interaction that schematically looks like Fig. 2.

At second order, S can include more complicated terms like

$$g^2(b^\dagger c^\dagger a)(a^\dagger c b),$$

which describes nucleon-anti-nucleon scattering. We can draw a nice diagram for this process too, seen in Fig. 3.

[33] This is a heuristically useful description but a little slippery in the details. A priori, there's no reason that eigenstates of the free Hamiltonian should be eigenstates of the interacting Hamiltonian. If you prefer, you can think of the scattering amplitude as the overlap (as measured by the inner product) between initial free particle states and final free particle states, with the possibility for some interaction in between. Even if we started with free particle eigenstates, our interaction is sure to evolve these states to some new ones, but we can look at the overlap between the time evolved versions of the free particle states $U(t,t_0)|i\rangle$ and the final free particle states we're interested in, $\langle f|$.

[34] When we talk about the fields, we use $*$, but when we denote antiparticles, we usually use the bar notation, e.g. an anti-ψ is a $\bar{\psi}$.

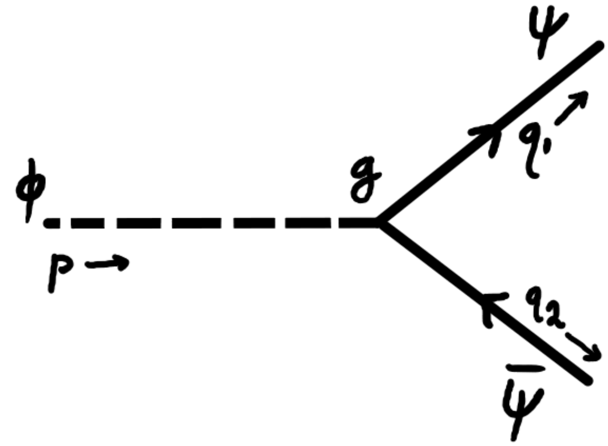

FIGURE 2. The Feynman diagram for meson decay, $\phi \to \psi\bar{\psi}$. Thinking perturbatively, this is the leading order behavior in the expansion of $\langle f|\, S\, |i\rangle$ where $|i\rangle \sim a^\dagger\, |0\rangle$, the one-meson state, and $|f\rangle \sim b^\dagger c^\dagger\, |0\rangle$, the state with a nucleon and an anti-nucleon.

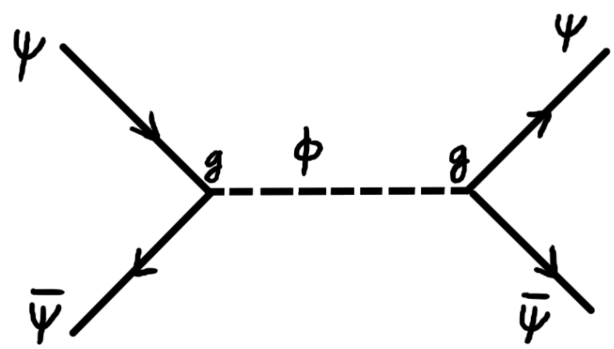

FIGURE 3. One Feynman diagram for nucleon-anti-nucleon scattering. Schematically, a nucleon and an anti-nucleon collide and annihilate into a meson, which then decays back to a nucleon and anti-nucleon. This isn't the only diagram at this order– we'll see that there's also a contributing diagram where the nucleon and anti-nucleon exchange a meson and then go on their way.

Returning to the case of meson decay, we have some ϕ meson going in with some defined momentum $\mathbf{p}$ as our initial state, and similarly we have $\psi, \bar{\psi}$ going out with some momenta $\mathbf{q}_1, \mathbf{q}_2$. We can write these states as

$$|i\rangle = \sqrt{2E_p}a^\dagger_{\mathbf{p}}\,|0\rangle$$

and

$$|f\rangle = \sqrt{4E_{q_1}E_{q_2}}b^\dagger_{\mathbf{q}_1}c^\dagger_{\mathbf{q}_2}\,|0\rangle\,.$$

To zeroth order there is no interaction and the scattering amplitude is zero, so to leading order, we have

$$\langle f|\, S\, |i\rangle = -ig\,\langle f|\int d^4x\psi^*(x)\psi(x)\phi(x)\,|i\rangle + O(g^2).$$

We'll compute this exactly next time and argue that the $O(g^2)$ corrections to this process are relatively small, arriving at our first quantum field theory scattering amplitude.

Lecture 9.

Tuesday, October 23, 2018

Today, we'll introduce Wick's formula and contractions, calculate some more scattering amplitudes, and maybe see our first Feynman diagrams for calculating amplitudes in a more convenient way.

First, we complete the calculation of the order g scattering amplitude from last time. We were interested in meson decay, where we prepared initial and final states

$$|i\rangle = \sqrt{2E_{\mathbf{p}}}a^\dagger_{\mathbf{p}}|0\rangle \tag{9.1}$$

and

$$|f\rangle = \sqrt{4E_{\mathbf{q}_1}E_{\mathbf{q}_2}}b^\dagger_{\mathbf{q}_1}c^\dagger_{\mathbf{q}_2}|0\rangle, \tag{9.2}$$

and we were interested in the scattering amplitude $\langle f|S|i\rangle$. To leading order, we found that

$$\langle f|S|i\rangle = -ig\langle f|\int d^4x\psi^*(x)\psi(x)\phi(x)|i\rangle + O(g^2), \tag{9.3}$$

and we'll now demonstrate how to compute this.

We know how to expand each of these fields in terms of their respective creation and annihilation operators, and we want to make sure that the initial state and final state are indeed proportional to each other so that this QFT amplitude will be reduced to a c-function of the four-momenta (i.e. it is just a number).

Note that when we put fields in, the creation and annihilation operators have to precisely cancel out the particles in the initial and final states. For instance, in the field ϕ we have both $a^\dagger_{\mathbf{p}}$ and $a_{\mathbf{p}}$ terms, but our initial state $|i\rangle$ already has an $a^\dagger$ in it. So the $a^\dagger$ bit of ϕ acting on $|i\rangle$ will produce a two-meson state proportional to $a^\dagger_{\mathbf{p}'}a^\dagger_{\mathbf{p}}|0\rangle$ which the ψs won't touch, and this means that the inner product of this two-meson state with $\langle 0|$ will be zero. Alternately, you can think of the $a^\dagger$ from ϕ as acting on the $\langle 0|$ on the left, since we can freely commute it through the ψs. That is, $a_{\mathbf{k}}|0\rangle = 0 \implies \langle 0|a^\dagger_{\mathbf{k}} = 0$, so in general any state with particles in it is going to be orthogonal to the vacuum. Our problem is therefore reduced to matching the operators in our fields with the operators in the initial and final states $|i\rangle$ and $|f\rangle$.

If we expand out the field ϕ, our matrix element now takes the form

$$\langle f|S|i\rangle = -ig\langle f|\int d^4x\psi^*(x)\psi(x)\int\frac{d^3k}{(2\pi)^3}\frac{\sqrt{2E_{\mathbf{p}}}}{\sqrt{2E_{\mathbf{k}}}}(a_{\mathbf{k}}a^\dagger_{\mathbf{p}}e^{-ik\cdot x} + a^\dagger_{\mathbf{k}}a^\dagger_{\mathbf{p}}e^{ik\cdot x})|0\rangle. \tag{9.4}$$

But as we've just argued, this second term is zero, and we can switch the $a_{\mathbf{k}}, a^\dagger_{\mathbf{p}}$ at the cost of a delta function $(2\pi)^3\delta^3(\mathbf{k}-\mathbf{p})$, which allows us to do the integral over d^3k.[35]

Now expanding the fields ψ^*, ψ, we get

$$\begin{aligned}
\langle f|S|i\rangle &= -ig\langle f|\int d^4x\psi^*\psi(x)e^{-ip\cdot x}|0\rangle \\
&= -ig\langle 0|\int\frac{d^4x}{(2\pi)^6}\frac{d^3k_1d^3k_2}{\sqrt{4E_{\mathbf{k}_1}E_{\mathbf{k}_2}}}\sqrt{4E_{\mathbf{q}_1}E_{\mathbf{q}_2}}c_{\mathbf{q}_2}b_{\mathbf{q}_1}(b^\dagger_{\mathbf{k}_1}e^{ik_1\cdot x} + c_{\mathbf{k}_1}e^{-ik_1\cdot x}) \\
&\quad\times(b_{\mathbf{k}_1}e^{-ik_2\cdot x} + c^\dagger_{\mathbf{k}_2}e^{ik_2\cdot x})e^{-ip\cdot x}|0\rangle.
\end{aligned}$$

From the fields ψ^* and ψ, only the $b^\dagger$ and $c^\dagger$ terms give a nonzero contribution, and taking the appropriate commutators gives us delta functions over the momenta (i.e. setting $\mathbf{q}_1 = \mathbf{k}_1, \mathbf{q}_2 = \mathbf{k}_2$). Using these delta functions to compute the k_1, k_2 integrals we find that

$$\begin{aligned}
\langle f|S|i\rangle &= -ig\langle 0|\int d^4xe^{i(q_1+q_2-p)\cdot x}|0\rangle \\
&= -ig[(2\pi)^4\delta^4(q_1+q_2-p)],
\end{aligned}$$

where this delta function simply imposes overall 4-momentum conservation. Note that this is a "matrix element," and not a probability yet. To actually turn this into a measurable probability (e.g. for an

[35] That is, we write $a_{\mathbf{k}}a^\dagger_{\mathbf{p}}|0\rangle = (a^\dagger_{\mathbf{p}}a_{\mathbf{k}} + (2\pi)^3\delta^3(\mathbf{k}-\mathbf{p}))|0\rangle = (2\pi)^3\delta^3(\mathbf{k}-\mathbf{p})|0\rangle$ since $a_{\mathbf{k}}$ kills the vacuum.

experiment), we must take the mod squared and integrate over the possible outgoing momenta q_1, q_2– we'll discuss this more later.[36]

Wick's theorem We'll now discuss *Wick's theorem* for a real scalar field. When we are working out S-matrix elements for more than one interaction, we will often need to compute quantities like

$$\langle f| T\{H_I(x_1)\ldots H_I(x_n)\} |i\rangle ,$$

the amplitude of some time-ordered product– remember that Dyson says we ought to be evolving our states in time with time-ordered products. Our lives would be easier if we could work in terms of normal-ordered products instead, where the as are on the RHS and the $a^\dagger$s are on the LHS. This would let us easily see which terms contribute to the final amplitude (e.g. which as precisely cancel particles created by $a^\dagger$s in the initial state $|i\rangle$, and vice versa for the outgoing state $|f\rangle$). In fact we can do this! Wick's theorem relates time-ordered products to normal-ordered products in a reasonably nice way.

Let's compute a simple example first. For a real scalar field ϕ, what is the time-ordered product $T\{\phi(x)\phi(y)\}$?

To do this computation, we write our scalar field as

$$\phi(x) \equiv \phi^+(x) + \phi^-(x)$$

where

$$\phi^+(x) \equiv \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}} a_{\mathbf{p}} e^{-ip\cdot x}$$

is the annihilation part of the field ϕ and

$$\phi^-(x) \equiv \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}} a^\dagger_{\mathbf{p}} e^{+ip\cdot x}$$

is the creation part of ϕ.[37] Now if we first consider the case $x^0 > y^0$, then $T\{\phi(x)\phi(y)\}$ takes the form

$$\begin{aligned}
T\{\phi(x)\phi(y)\} &= \phi(x)\phi(y)\\
&= (\phi^+(x)+\phi^-(x))(\phi^+(y)+\phi^-(y))\\
&= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^-(y) + \phi^-(y)\phi^+(x) + \phi^-(x)\phi^+(y) + [\phi^+(x), \phi^-(y)]\\
&=: \phi(x)\phi(y) : + D(x-y)
\end{aligned}$$

where we've collected terms with all the ϕ^+ terms to the right of the ϕ^- terms into the normal-ordered product $:\phi(x)\phi(y):$. If we had $y^0 > x^0$ instead, we would get

$$T\{\phi(x)\phi(y)\} =: \phi(x)\phi(y) : + D(y-x).$$

Putting these together, we see that

$$T\{\phi(x)\phi(y)\} =: \phi(x)\phi(y) : + \Delta_F(x-y), \tag{9.5}$$

where Δ_F is simply the Feynman propagator. It's important to note that while the time-ordered and normal-ordered products are both operators, their difference is Δ_F, a c-function.

Definition 9.6. We define a *contraction* of a pair of fields in a string (denoted by a square bracket between the two fields, or a curly overbrace here because of the limitations of my LaTeX formatting) to mean replacing the two contracted fields by their Feynman propagator. That is, $\overbrace{\phi(x)\phi(y)} \equiv \Delta_F(x-y)$.

For instance, we saw that

$$T\{\phi(x)\phi(y)\} =: \phi(x)\phi(y) : + \overbrace{\phi(x)\phi(y)}.$$

Theorem 9.7 (Wick's theorem)**.** *Time-ordered products of fields are related to normal-ordered products in the following way:*

$$T\{\phi(x_1)\ldots\phi(x_n)\} =: \phi(x_1)\ldots\phi(x_n) : + : \text{all possible contractions} : \tag{9.8}$$

[36] See Lecture 13 and 14 if you're impatient, or a would-be experimentalist. Or both.

[37] The signs here are an unfortunate convention having to do with these being "positive frequency" and "negative frequency" operators.

Note that "all possible contractions" here includes combinations of fields that are not fully contracted. For instance, the product $T\{\phi(x_1)\phi(x_2)\phi(x_3)\}$ will include terms like $\Delta_F(x_1 - x_2)\phi(x_3)$. Similarly $T\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\}$ includes $\Delta_F(x_1 - x_2) : \phi(x_3)\phi(x_4) :$ as well as the totally contracted $\Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4)$.

Example 9.9. Since all normal ordered terms kill the vacuum state, Wick's theorem allows us to immediately compute amplitudes like

$$\langle 0| T\{\phi_1 \dots \phi_4\} |0\rangle = \Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4) + \Delta_F(x_1 - x_3)\Delta_F(x_2 - x_4) + \Delta_F(x_1 - x_4)\Delta_F(x_2 - x_3).$$

The proof of Wick's theorem is by induction. Suppose it holds for $T\{\phi_2 \dots \phi_n\}$. Then (see textbooks for detail)

$$T\{\phi_1\phi_2 \dots \phi_n\} = (\phi_1^+ + \phi_1^-)(: \phi_2 \dots \phi_n + : \text{all contractions of } \phi_2 \dots \phi_n :).$$

The ϕ_1^- is okay where it is, while the ϕ_1^+ must be commuted to the RHS of the $\phi_2 \dots \phi_n$ terms. Each commutator past the x_k term in $\phi_2 \dots \phi_n$ gives us a $D(x_1 - x_k)$, which is equivalent to a contraction between ϕ_1 and ϕ_k.

Wick's theorem has some immediate consequences. For instance,

$$\langle 0| T\{\phi_1 \dots \phi_n\} |0\rangle = 0$$

if n is odd (since one ϕ is always left out of the contractions) and it is

$$\sum_{i_1,\dots,i_n} \Delta_F(x_{i_1} - x_{i_2})\Delta_F(x_{i_3} - x_{i_4}) \dots \Delta_F(x_{i_{n-1}} - x_{i_n})$$

if n is even, where the sum is taken over symmetric permutations of $i_1, \dots, i_n$.

Note that Wick's theorem also has a generalization to complex fields $\psi \in \mathbb{C}$, e.g.

$$T(\psi(x)\psi^*(y)) =: \psi(x)\psi^*(y) : +\Delta_F(x - y)$$

where the contraction of a ψ and ψ^* is a propagator, $\overbrace{\psi(x)\psi^*(y)} \equiv \Delta_F(x - y)$, and the contractions of two ψs or two ψ^*s is zero, $\overbrace{\psi(x)\psi(y)} = \overbrace{\psi(x)^*\psi^*(y)} = 0$.

Lecture 10.

Thursday, October 25, 2018

Last time, we computed the amplitude for meson decay to first order in g. Now let's apply Wick's theorem to nucleon scattering, $\psi(p_1)\psi(p_2) \to \psi(p_1')\psi(p_2')$. Our initial state looks like

$$|i\rangle = \sqrt{4E_{p_1}E_{p_2}} b^\dagger_{\mathbf{p}_1} b_{\mathbf{p}_2} |0\rangle \equiv |p_1, p_2\rangle$$

and our final state is

$$|f\rangle = \sqrt{4E_{p_1'}E_{p_2'}} b^\dagger_{\mathbf{p}_1'} b_{\mathbf{p}_2'} |0\rangle \equiv |p_1', p_2'\rangle .$$

We aren't interested in the trivial case where there's no scattering (i.e. the zeroth order term where the nucleons just go on their way without any interaction). Moreover, any single interaction $O(g)$ would produce a meson we don't want in our final state, so there is no order g contribution. What we're really interested in is the $O(g^2)$ term in $\langle f| (S - 1) |i\rangle$.

The amplitude takes the form

$$\frac{(-ig)^2}{2!} \int d^4x_1 d^4x_2 \langle p_1', p_2'| T\{\psi^*(x_1)\psi(x_1)\phi(x_1)\psi^*(x_2)\psi(x_2)\phi(x_2)\} |p_1, p_2\rangle .$$

Using Wick's theorem, we know there is a term of the form

$$: \psi^*(x_1)\psi(x_1)\psi^*(x_2)\psi(x_2) : \overbrace{\phi(x_1)\phi(x_2)}$$

in the time-ordered product. The contracted bit will make sure we have no issues with ϕ fields (since there are no ϕs in our initial and final states), while the normal-ordered part gives us ψs to annihilate the initial nucleons and ψ^*s to create the final nucleons. All other terms are zero.

We will ignore all terms involving $c, c^\dagger$ in the field expansions since they give zeroes (i.e. to this order, we don't need to worry about antiparticles). We ought to compute

$$\begin{aligned}\langle p_1', p_2' | : \psi^*(x_1)\psi(x_1)\psi^*(x_2)\psi(x_2) : |p_1, p_2\rangle = \int &\frac{d^3q_1 \dots d^3q_4 \sqrt{16E_{p_1} \dots E_{p_4}}}{(2\pi)^1 2\sqrt{2E_{q_1} \dots 2E_{q_4}}} \\ &\times \langle 0| b_{\mathbf{p}_1'} b_{\mathbf{p}_2'} \underbrace{b^\dagger_{\mathbf{q}_1} b^\dagger_{\mathbf{q}_2}}_{\psi^* s} \underbrace{b_{\mathbf{q}_3} b_{\mathbf{q}_4}}_{\psi s} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2} |0\rangle \\ &\times e^{i(q_1\cdot x_1 + q_2 \cdot x_2 - q_3 \cdot x_1 - q_2 \cdot x_2)}.\end{aligned}$$

Using commutation relations one can check (in a few lines) that this big mess of creation and annihilation operators simplifies to a slightly more manageable mess of delta functions we can integrate over and get rid of. That is,

$$\begin{aligned}\langle 0| b_{\mathbf{p}_1'} b_{\mathbf{p}_2'} b^\dagger_{\mathbf{q}_1} b^\dagger_{\mathbf{q}_2} b_{\mathbf{q}_3} b_{\mathbf{q}_4} b^\dagger_{\mathbf{p}_1} b^\dagger_{\mathbf{p}_2} |0\rangle = &\left[\delta^3(\mathbf{p}_1' - \mathbf{q}_2)\delta^3(\mathbf{p}_2' - \mathbf{q}_1) + \delta^3(\mathbf{p}_2' - \mathbf{q}_2)\delta^3(\mathbf{p}_1' - \mathbf{q}_1)\right] \\ &\times \left[\delta^3(\mathbf{q}_4 - \mathbf{p}_1)\delta^3(\mathbf{q}_3 - \mathbf{p}_2) + \delta^3(\mathbf{q}_4 - \mathbf{p}_2)\delta^3(\mathbf{q}_3 - \mathbf{p}_1)\right].\end{aligned}$$

If we now integrate over this, our delta functions give us several exponential terms:

$$\begin{aligned}\langle p_1', p_2' | : \psi^*(x_1)\psi(x_1)\psi^*(x_2)\psi(x_2) : |p_1, p_2\rangle = &\left[e^{i(p_1'\cdot x_2 + p_2' \cdot x_1)} + e^{i(p_2'\cdot x_2 + p_1'\cdot x_1)}\right] \\ &\times \left[e^{-i(p_1\cdot x_2 + p_2\cdot x_1)} + e^{-i(p_2 \cdot x_2 + p_1 \cdot x_1)}\right].\end{aligned}$$

Writing this all out, one can perform the x_1, x_2 integrals to get (surprise) even more delta functions. We also integrate over the internal momentum k and find as our final result

$$(-ig)^2 \left\{\frac{i}{(p_1 - p_1')^2 - m^2 + i\epsilon} + \frac{i}{(p_2' - p_1)^2 - m^2 + i\epsilon}\right\} (2\pi)^4 \delta^4(p_1 + p_2 - p_1' - p_2').$$

In fact, there are two terms here– one where the ψs exchange a meson and go on their ways, and one where the ψs exchange a meson and then cross over (so that what we thought was the first nucleon was actually the second). The meaning of this will be more obvious when we draw the Feynman diagrams, but we need both terms in order to ensure that the particles obey Bose-Einstein statistics (i.e. are indistinguishable). It should be clear that the delta function imposes conservation of overall momentum (i.e. the outgoing momentum is equal to the ingoing momentum, $p_1 + p_2 = p_1' + p_2'$).

Feynman diagrams This is basically the simplest interesting calculation we could have done, and using Wick's theorem to get there has given us a big mess. Surely there must be a better way, you say. And there is. We draw *Feynman diagrams* to keep track of the different possible Wick contractions, i.e. to represent the perturbative expansion of $\langle f| (S-1) |i\rangle$. We have a set of rules for how to draw the diagrams representing different processes, and can associate integrals to the diagrams.

Here are the rules.

- Draw an external line for each particle in the initial and final states $|i\rangle, |f\rangle$, assigning a four-momentum to each.
- For $\mathbb{C}$ fields we ought to add an arrow to label the flow of charge. Choose an in(out)-going arrow for (anti-)particles in $|i\rangle$, and the opposite convention holds for $|f\rangle$.
- Join the lines together with vertices as prescribed by the Lagrangian, i.e. making sure that the interaction has a corresponding term and that charge is conserved in each vertex.
- Assign a momentum k to each internal line i.
- Add a delta function corresponding to each vertex for momentum conservation, $(-ig)(2\pi)^4\delta^4(\sum_i k_i)$, where $\sum_i k_i$ is the sum of all 4-momenta flowing into the vertex and g is the coupling constant in the Lagrangian.
- For each internal line with a 4-momentum k, write a factor of the propagator for that particle, e.g. in Yukawa theory,

$$\int \frac{d^4k}{(2\pi)^4} D(k^2) \text{ where } D(k^2) = \begin{cases} \frac{i}{k^2 - m^2 + i\epsilon} & \text{for } \phi \\ \frac{i}{k^2 - \mu^2 + i\epsilon} & \text{for } \psi \end{cases}$$

Using the Feynman rules, we can draw the two $O(g^2)$ diagrams and immediately write down the amplitude for our nucleon scattering process: it is

$$\begin{aligned}\langle f|\,(S-1)\,|i\rangle &= (-ig)^2\int\frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i\epsilon}(2\pi)^8\{\delta^4(p_1-p_1'-k)\delta^4(p_2-p_2'+k)\\&\quad+\delta^4(p_1-p_2'-k)\delta^4(p_2+k-p_1')\}\\&= i(-ig)^2\left(\frac{1}{(p_1-p_1')^2-m^2+i\epsilon}+\frac{1}{(p_1-p_2')-m^2+i\epsilon}\right)(2\pi)^4\delta^4(p_1+p_2-p_1'-p_2').\end{aligned}$$

The diagrams are suggestive of an analogous classical scattering process, like billiard balls colliding elastically. If we like, we can say that this is like the nucleons exchanging a meson of 4-momentum k. However, note that this meson doesn't necessarily satisfy the relativistic dispersion relation $k^2 = m^2$. If it doesn't, it's called "off-shell" or a virtual particle, and the impact of virtual particle interactions is a purely quantum effect.

Conversely, the external legs of our diagram are forced to be *on-shell*– because these are outgoing particles (that one could really observe and measure in a detector, for example), they had better satisfy the relativistic dispersion relation. It's also important to recognize that while internal momenta are fixed by momentum conservation in "tree-level" diagrams, once we introduce loops into our Feynman diagrams all bets are off and we must integrate over all possible momenta for those virtual particles.

Lecture 11.

Saturday, October 27, 2018

Last time, we introduced the Feynman rules for drawing Feynman diagrams and computing scattering amplitudes, and it's good to check that these diagrams really do correspond to Wick contractions of our fields. Let's now make a canonical definition of the *amplitude* $\mathcal{A}_{fi}$, defined by

$$\langle f|\,(S-1)\,|i\rangle \equiv i\mathcal{A}_{fi}\,\underbrace{(2\pi)^4\delta^4(\sum_{j\in f}p_j-\sum_{j\in i}p_j)}_{\text{from translational invariance}}$$

where the i is included by convention to match with non-relativistic QM.

We should then refine the Feynman rules to compute the amplitude (stripping away the overall momentum-conserving delta function, since we will always get one). Here are our revised rules:

- Draw all possible diagrams with appropriate external legs given by $|i\rangle\,,|f\rangle$.
- Impose 4-momentum conservation at each vertex.
- Write a factor of the coupling $(-ig)$ at each vertex.
- For each internal line, add a factor of the propagator.
- Integrate over internal momenta $\int\frac{d^4k}{(2\pi)^4}$. (This is trivial for tree-level diagrams since the momenta are all fixed by momentum conservation, but these will be real integrals for diagrams with internal loops.)

Example 11.1. Consider the scattering process $\psi+\bar\psi\to\phi+\phi$ in scalar Yukawa theory. There are two diagrams for this, and both are of order $(-ig)^2$. We can write down the amplitude almost by inspection:

$$i\mathcal{A}_{if} = (-ig)^2\left[\frac{i}{(p_1-p_1')^2-\mu^2}+\frac{i}{(p_1-p_2')^2-\mu^2}\right]$$

Note we've dropped the $i\epsilon$s here since the denominators don't vanish.

Example 11.2. We can now consider our first loop diagram, $\phi\phi\to\phi\phi$. It's a $O(g^4)$ diagram, so we write down the amplitude for this diagram as

$$i\mathcal{A}_{if} = (-ig)^4\int\frac{d^4k}{(2\pi)^4}\frac{i}{k^2-\mu^2+i\epsilon}\frac{i}{(k-p_2')^2-\mu^2+i\epsilon}\frac{i}{(k+p_1'-p_1)^2-\mu^2+i\epsilon}\frac{i}{(k+p_1')^2-\mu^2+i\epsilon}.$$

We won't actually compute this integral, though we should note that at least it has a chance of converging since it goes as d^4k/k^8. These loop integrals can be tricky, and we'll revisit them in more detail next term in

Advanced QFT. Sometimes the integrals won't converge, and we'll need the machinery of renormalization to sweep away the infinities and get actual numbers out of our integrals.

Let's now consider ϕ^4 theory, with $\mathcal{H}_{int} = \frac{\lambda}{4!}\phi^4$. Now we have a single interaction vertex– it's a 4-point vertex, where for each vertex we pay a cost of $-i\lambda$. The other Feynman rules are the same. Note that there's no $1/4!$ factor in the final amplitude. To see why, consider the simplest diagram for $\phi\phi \to \phi\phi$ scattering.

$$i\mathcal{A}_{fi} \sim -\frac{i\lambda}{4!} \left\langle p_1', p_2' \right| : \phi(x)\phi(x)\phi(x)\phi(x) : \left| p_1, p_2 \right\rangle .$$

Generically, this is

$$\langle 0 | a_{\mathbf{p}_1'} a_{\mathbf{p}_2'} \dots a^\dagger_{\mathbf{p}_1} a^\dagger_{\mathbf{p}_2} | 0 \rangle$$

and so any one of the fields ϕ can annihilate or create the external particles. Therefore there are 4! ways of matching up the operators and commuting them so that we start and end with $|0\rangle$. You can get other combinatoric factors like this (often 2 or 4). Having a term $\lambda_n \phi^n / n!$ in the Lagrangian is conventional, though.

Let's consider now

$$\langle 0 | T\{\phi(x_1) \dots \phi(x_m) S\} | 0 \rangle ,$$

which we call a correlation function. This is analogous to the correlation functions we saw in Statistical Field Theory. It's a more elementary but less physical object than an S-matrix element. For brevity, denote

$$\phi_i \equiv \phi(x_i).$$

Now the nth term in the expansion for S gives

$$\frac{1}{n!}\left(\frac{-i\lambda}{4!}\right)^n \int d^4y_1 \dots d^4y_n \langle 0 | T\{\phi_1 \dots \phi_m \phi^4(y_1) \dots \phi^4(y_n)\} | 0 \rangle .$$

Wick's theorem tells us to contract all pairs of fields in all possible ways. As an example, consider the case $n = 1, m = 4$. Then we have a term

$$-\frac{i\lambda}{4!} \int d^4x \langle 0 | T\{\phi_1 \dots \phi_4 \phi^4(x)\} | 0 \rangle .$$

We're going to have to contract all the fields, since any uncontracted fields will kill the vacuum states after normal ordering. We could get a contraction where all the numbered ϕ fields contract with the xs, e.g.

$$-\frac{i\lambda}{4!} \int d^4x \, \overbrace{\phi_1\phi(x)} \, \overbrace{\phi_2\phi(x)} \, \overbrace{\phi_3\phi(x)} \, \overbrace{\phi_4\phi(x)}$$

and permutations of $\phi_1, \dots, \phi_4$. We could also contract two of the numbered ϕs,[38]

$$-\frac{i\lambda}{4!} \int d^4x \, \overbrace{\phi_1\phi_2} \, \overbrace{\phi_3\phi(x)} \, \overbrace{\phi_4\phi(x)} \, \overbrace{\phi(x)\phi(x)}$$

and permutations of contracting 2 ϕ_is. Finally, we'll have contractions of all the ϕ_is together, which look like

$$-\frac{i\lambda}{4!} \int d^4x \, \overbrace{\phi_1\phi_2} \, \overbrace{\phi_3\phi_4} \, \overbrace{\phi(x)\phi(x)} \, \overbrace{\phi(x)\phi(x)} .$$

The first of these gives us 4! terms of Feynman propagators $\Delta_F(x_i - x)$ (4 choices for x_1, 3 for x_2, and so on). There are 12 unique choices for which two fields ϕ_i, ϕ_j to contract with $\phi(x)$s, and 12 ways of pairing those fields ϕ_i, ϕ_j with $\phi(x)$s (4 choices for ϕ_i and then 3 choices for ϕ_j). Finally, there are 3 ways of pairing only $\phi(x)$s (e.g. take ϕ_1. We get 3 choices of $\phi_{i \neq 1}$ to pair it with, and the other contraction is then fixed).

So the first term gets $-i\lambda$, the second gets $-i\lambda/2$, and the last gets $-i\lambda/8$. Note that $\Delta_F(x - x) = \Delta_F(0)$ diverges, so these "bubble" diagrams will diverge badly. Our theory turns out to be renormalizable, but again this isn't always the case.

[38] This will give us a disconnected Feynman diagram.

Lecture 12.

Tuesday, October 30, 2018

Last time, we introduced the correlation functions

$$\langle 0 | T\{\phi(x_1)\dots\phi(x_m)S\} |0\rangle .$$

Let's consider the term with $m = 4$ and $n = 2$ (four numbered fields ϕ_i and two four-point vertices ϕ^4). That term looks like

$$\frac{1}{2}\left(\frac{-i\lambda}{4!}\right)^2 \int d^4x d^4y \langle 0| T\{\phi_1\phi_2\phi_3\phi_4\phi^4(x)\phi^4(y)\} |0\rangle .$$

As before, we claim that the most important contributions are the completely connected ones, and anything not totally contracted will vanish in the time-ordered product. One such contraction is

$$\overbrace{\phi_1\phi(x)}\,\overbrace{\phi_2\phi(x)}\,\overbrace{\phi_3\phi(y)}\,\overbrace{\phi_4\phi(y)}\,\overbrace{\phi(x)\phi(y)}\,\overbrace{\phi(x)\phi(y)}\,.$$

But we could get some distinct diagrams depending on how we connect up the dots. The Feynman rules for the first diagram give

$$\frac{(-i\lambda)^2}{2}\int d^4xd^4y\Delta_F(x_1-x)\Delta(x_2-x)\Delta_F(x_3-y)\Delta_F(x_4-y)\Delta_F^2(x-y).$$

Let's work out the combinatoric factors: there are four choices for which $\phi(x)$ goes with x_1 and three choices for which $\phi(x)$ goes with x_2, for a factor of 12. The same is true for x_3, x_4 and y. We get a factor of 2 for which of the remaining $\phi(y)$s the first $\phi(x)$ contracts with, and then the other is determined. We also get a factor 2! from interchange of x and y. The four $\phi(x)$s are identical, as are the four $\phi(y)$s, so we should add a factor of $(1/4!)^2$ to take care of that. Finally, we have $\binom{4}{2} = 1/2!$ choices of which ϕ_is to connect to $\phi(x)$s. Putting it all together we get

$$\frac{1}{2!}\times\left(\frac{1}{4!}\right)^2\times\underbrace{12}_{x_1,x_2\to x}\times\underbrace{12}_{x_3,x_4\to y}\times\underbrace{2}_{x\to y}\times\underbrace{2!}_{x\leftrightarrow y} = \frac{1}{2}.$$

The Feynman rules for the correlation functions of ϕ^4 theory are then given by

$$\langle 0| T\left\{\phi(x_1)\dots\phi(x_m)\exp\left(-\frac{i\lambda}{4!}\int d^4x\phi^4(x)\right)\right\}|0\rangle ,$$

which is equal to the sum of all diagrams with m external points and any number of internal vertices connected by propagator lines. In perturbation theory, we categorize the diagrams based on the number of powers of λ, i.e. the number of vertices in the diagram. For each diagram, there is one integral containing

- Each propagator from y to z, $\Delta_F(y-z)$
- Each vertex at x, $-i\lambda\int d^4x$,

and we divide by a symmetry factor. Since the propagator is an integral over momentum space, it's easier to express the Feynman rules in momentum space. Rather than integrating over all space d^4x we can equivalently just integrate a momentum-conserving delta function. Let's work out the momentum space Feynman rules:

- To each propagator from x to y, assign $e^{ip\cdot y}$ to the y vertex (where the arrow is going out) and $e^{-ip\cdot x}$ to the vertex x with arrows in.
- Associate $\frac{i}{p^2-m^2+i\epsilon}$ to the line itself (for a particle with mass m) and an integral over all momentum $\int\frac{d^4p}{(2\pi)^4}$.
- Thus the integral at a vertex becomes

  $$\int d^4xe^{-ip_1\cdot x+ip_2\cdot x-ip_3\cdot x+ip_4\cdot x} = (2\pi)^4\delta^4(p_1+p_3-p_2-p_4)$$

  where p_1, p_3 are flowing into the vertex, p_2, p_4 out. (There should also be a $-i\lambda$ for each vertex.)

However, as before the δ functions will make some of the momentum integrals trivial, and for each of these the $(2\pi)^4$ will cancel. We are left with the following momentum space rules:

- For each internal line associate a factor of $\frac{i}{p^2-m^2+i\epsilon}$.
- For each vertex associate a factor of $-i\lambda$.
- Impose four-momentum conservation at vertices, and overall.
- Integrate over undetermined momenta from internal lines, $\int \frac{d^4k}{(2\pi)^4}$.
- Divide by the appropriate symmetry factor.

Note that there isn't really a nice way to get the symmetry factors from looking at the Feynman diagrams– one must usually consider the Wick contraction to get these factors right.

Vacuum bubbles and connected diagrams What is the transition from the vacuum state to the vacuum state, $\langle 0| S |0\rangle$? In ϕ^4 theory, we get a sum of "vacuum bubbles," diagrams with *no external lines*. One should check (e.g. on the second example sheet) that the S-matrix element is simply the exponential of the various topologically distinct vacuum bubble diagrams. Weird!

In general we call the correlation function

$$\langle 0| T\{\phi(x_1)\dots\phi(x_m)S\} |0\rangle$$

an *m-point function*, and its value is the sum over diagrams with m external points. A typical diagram has some vacuum bubbles, e.g. at second order in ϕ^4 we have a disconnected diagram which looks like a line with a loop and the figure 8. Remarkably, the vacuum bubbles add to the same exponential as in the pure vacuum case. We'll discuss this more in detail next term, but there is an apparently sensible way of treating the vacuum bubbles.[39] Therefore we may write

$$\langle 0| T\{\phi(x_1)\dots\phi(x_m)S\} |0\rangle = (\textstyle\sum \text{connected diagrams}) \times \langle 0| S |0\rangle\,,$$

where connected means that every point in the diagram is connected to at least one external line.

Really, the issue here comes from the fact that the vacuum of the free theory is *not* the vacuum of the interacting theory.

Definition 12.1. Let $|\Omega\rangle$ be the vacuum of the *interacting* theory, normalized such that $H|\Omega\rangle = 0$ with $H = H_0 + H_{int}$ (n.b. $H_0|0\rangle = 0$) and $\langle\Omega|\,\Omega\rangle = 1$. Then we define

$$G^{(n)}(x_1\dots x_n) \equiv \langle\Omega| T\{\phi_H(x_1)\dots\phi_H(x_n)\} |\Omega\rangle\,.$$

We call these *Green's functions*.

We claim now that

$$\langle\Omega| T\{\phi_{1,H}\dots\phi_{m,H}\} |\Omega\rangle = \frac{\langle 0| T\{\phi_{1,I}\dots\phi_{m,I}S\} |0\rangle}{\langle 0| S |0\rangle}.$$

What this means is that the Green's functions are precisely given by the sum of connected diagrams with m external points– we need not worry too much about the vacuum bubbles and disconnected diagrams because removing the vacuum bubbles gets the behavior relative to the *interacting* vacuum right (and S evolves our interaction picture fields to Heisenberg picture fields). We'll do the proof next time.

Lecture 13.

Thursday, November 1, 2018

Last time, we claimed that

$$\langle\Omega| T\{\phi_{1,H}\dots\phi_{m,H}\} |\Omega\rangle = \frac{\langle 0| T\{\phi_{1,I}\dots\phi_{m,I}S\} |0\rangle}{\langle 0| S |0\rangle}.$$

That is, it suffices to consider only connected diagrams, since the vacuum bubbles add up to a multiplicative factor (namely, the vacuum energy) that can be factored out of the overall correlation function.

[39] I believe this is related to renormalization.

Proof. To prove this, let us expand the numerator on the RHS as

$$\langle 0| U(\infty,t_1)\phi_{1,I}U(t_1,t_2)\phi_{2,I}\dots U(t_{n-1},t_n)\phi_{n,I}U(t_n,-\infty)|0\rangle ,$$

and WLOG we label the fields to already be time-ordered, with $x_1^0 > x_2^0 > \dots > x_m^0$. That is, we've split up the overall time evolution operator $S = \lim_{t\to\infty, t_0\to-\infty} U(t,t_0)$ into intervals from $t_i = x_i^0$ to $t_{i+1} = x_{i+1}^0$ in order to write out the time ordering. We can then break up the time evolution operators as

$$U(t_1,t_2) = U(t_1,0)U(0,t_2)$$

so that the numerator becomes

$$\langle 0| U(\infty,0)\underbrace{[U(0,t_1)\phi_{1,I}U(t_1,0)]}_{\phi_{1,H}}[U(0,t_2)\phi_{2,I}U(t_2,0)]\dots U(t_{n-1},0)[U(0,t_n)\phi_{n,I}U(t_n,0)]U(0,-\infty)|0\rangle ,$$

or more compactly,

$$\langle 0| U(\infty,0)\phi_{1,H}\dots\phi_{n,H}U(0,-\infty)|0\rangle , \tag{13.1}$$

which is nothing more than a bunch of Heisenberg picture operators sandwiched between the vacuum states and a pair of time evolution operators.

Lemma 13.2. *For a general state* $|\psi\rangle$,

$$\lim_{t_0\to-\infty}\langle\psi| U(0,t_0)|0\rangle = \langle\psi|\Omega\rangle\langle\Omega|0\rangle . \tag{13.3}$$

Proof. First, note that

$$\langle\psi| U(0,t_0)|0\rangle = \langle\psi| e^{iHt_0}|0\rangle$$

since $U(0,t_0) = e^{iHt_0}e^{-iH_0t_0}$ and $H_0|0\rangle = 0$. Insert a complete set of interacting states $|p_1,\dots,p_n\rangle$. Then

$$\begin{aligned}\lim_{t_0\to-\infty}\langle\psi| U(0,t_0)|0\rangle &= \lim_{t_0\to-\infty}\langle\psi| e^{iHt_0}\left[|\Omega\rangle\langle\Omega| + \sum_{n=1}^{\infty}\int\prod_{j=1}^{n}\frac{d^3p_j}{2E_{p_j}(2\pi)^3}|p_1,\dots,p_n\rangle\langle p_1,\dots,p_n|\right]|0\rangle\\ &= \langle\psi|\Omega\rangle\langle\Omega|0\rangle\\ &\quad+\lim_{t_0\to-\infty}\sum_{n=1}^{\infty}\int\prod_{j=1}^{n}\frac{d^3p_j}{2E_{p_j}(2\pi)^3}e^{i\sum_{k=1}^{n}E_{p_k}t_0}\langle\psi|p_1,\dots,p_n\rangle\langle p_1,\dots,p_n|0\rangle .\end{aligned}$$

Note that in the first term $\langle\psi| e^{iHt_0}|\Omega\rangle\langle\Omega|0\rangle$, all nonzero powers of H from the exponential will kill the (interacting) vacuum state $|\Omega\rangle$ by definition, so the only thing that survives is the zeroth order term, $\langle\psi|\Omega\rangle\langle\Omega|0\rangle$. Luckily, the second term vanishes due to the Riemann-Lebesgue lemma: stated roughly, "for reasonable $f(x)$ (i.e. square-integrable), $\lim_{\mu\to\infty}\int_a^b f(x)e^{i\mu x}dx = 0$." With this second term gone, we conclude that

$$\lim_{t_0\to-\infty}\langle\psi| U(0,t_0)|0\rangle = \langle\psi|\Omega\rangle\langle\Omega|0\rangle .$$

☒

By the same reasoning,

$$\lim_{t_0\to\infty}\langle 0| U(t_0,0)|\psi\rangle = \langle 0|\Omega\rangle\langle\Omega|\psi\rangle . \tag{13.4}$$

Now we can apply our lemma so that our numerator (Eqn. 13.1) becomes

$$\langle 0| U(\infty,0)\phi_{1,H}\dots\phi_{n,H}U(0,-\infty)|0\rangle = \langle\Omega|\phi_{1,H}\dots\phi_{m,H}|\Omega\rangle\langle\Omega|0\rangle\langle 0|\Omega\rangle \tag{13.5}$$

and the denominator is just $\langle 0|\Omega\rangle\langle\Omega|0\rangle$.[40] Therefore

$$\frac{\langle 0| T\{\phi_{1,I}\dots\phi_{m,I}S]\}|0\rangle}{\langle 0| S|0\rangle} = \frac{\langle\Omega|\phi_{1,H}\dots\phi_{m,H}|\Omega\rangle\langle\Omega|0\rangle\langle 0|\Omega\rangle}{\langle\Omega|0\rangle\langle 0|\Omega\rangle} = \langle\Omega| T\{\phi_{1,H}\dots\phi_{m,H}\}|\Omega\rangle ,$$

as promised. Note we have put time ordering back in since we explicitly time-ordered the fields when we expanded out S. This completes the proof. ☒

[40] It's literally the same calculation– just take out the fields ϕ. Then the first factor is just $\langle\Omega|\Omega\rangle = 1$ by normalization of the interacting theory vacuum states and we're left with $\langle\Omega|0\rangle\langle 0|\Omega\rangle$.

In words, this tells us that we can do our calculations relative to the vacuum of the interacting theory $|\Omega\rangle$ rather than the vacuum of the free theory $|0\rangle$, which means (in terms of our perturbative expansion) that we need not consider vacuum bubbles when we compute our correlation functions.

Going back to our previous example, we say that to describe scattering in the interacting theory, our external states, e.g. $|p_1, p_2\rangle$, should come from the interacting theory. This means that we exclude loops on the external lines (a process we call "amputation").

Mandelstam variables In two-particle scattering processes, the same combinations of p_1, p_2, p_1', p_2' (ingoing and outgoing four-momenta) often appear, so it's useful to introduce the *Mandelstam variables* s, t, and u, defined as

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_1' + p_2')^2 \\ t &= (p_1 - p_1')^2 = (p_2 - p_2')^2 \\ u &= (p_1 - p_2')^2 = (p_2 - p_1')^2 \end{aligned}$$

where the squared here indicates a four-vector product (e.g. $(p_1 + p_2)^2 = (p_1^\mu + p_2^\mu)(p_{1\mu} + p_{2\mu})$.

Exercise 13.6. Show that the sum of the Mandelstam variables is

$$s + t + u = m_1^2 + m_2^2 + {m_1'}^2 + {m_2'}^2,$$

where m_1, m_2, m_1', m_2' are the masses of the initial and final particles, so the Mandelstam variables are not all independent.

WLOG, we can consider the initial particles in the center-of-mass frame, i.e. a frame in which the net 3-momentum is zero. Thus $\mathbf{p}_1 = -\mathbf{p}_2$. In this frame, s takes the simple form

$$s = (p_1 + p_2)^2 = (E_1 + E_2)^2.$$

Since s is a Lorentz scalar, it takes the same value in all frames. Therefore $\sqrt{s}$ is the center of mass energy, e.g. at the LHC we say that $\sqrt{s} = 13\,\text{TeV}$. In particular if $m_1 = m_2$, then by symmetry $E_1 = E_2 = \sqrt{s}/2$.

Cross sections and decay rates So far, $|i\rangle$ and $|f\rangle$ have been states of definite momenta. What happens in a realistic situation where our ingoing states are now some distribution (a density function) smeared over momenta?

To understand this, suppose we have a collision with $2 \to n$ scattering, i.e. we have two particles ingoing with momenta p_1, p_2 and n outgoing particles with momenta $q_1, \ldots, q_n$. Then the scattering amplitude is proportional to

$$\langle q_1 q_2 \ldots q_n |\, p_1 p_2\rangle\, (2\pi)^4 \delta^4(p_1 + p_2 - \sum_{i=1}^n q_i).$$

But probabilities are related to the amplitude squared, so it seems as if we've picked up an extra delta function in computing the physical probability of this interaction. The resolution is this– in reality, $|i\rangle, |f\rangle$ are very sharply peaked superpositions of momentum eigenstates. That is, our ingoing states take the form

$$|p_1 p_2\rangle_{in} = \int \frac{d^3\tilde{p}_1}{(2\pi)^3 2E_{\tilde{p}_1}} \frac{d^3\tilde{p}_2}{(2\pi)^3 2E_{\tilde{p}_2}} f_1(\tilde{p}_1) f_2(\tilde{p}_2)\, |\tilde{p}_1 \tilde{p}_2\rangle\,,$$

where $|\tilde{p}_1 \tilde{p}_2\rangle$ are the real four-momentum eigenstates.

If we suppose that the outgoing particles are also pure momentum eigenstates, then then our delta functions are soaked up by integrals when we try to compute the transition probability W. We then have

$$\begin{aligned} W = (2\pi)^8 \int & \frac{d^3\tilde{p}_1}{(2\pi)^3 2E_{\tilde{p}_1}} \frac{d^3\tilde{p}_2}{(2\pi)^3 2E_{\tilde{p}_2}} \frac{d^3 p_1'}{(2\pi)^3 2E_{p_1'}} \frac{d^3 p_2'}{(2\pi)^3 2E_{p_2'}} \\ & \times \left\{ |M|^2 f_1(\tilde{p}_1) f_1^*(p_1') f_2(\tilde{p}_2) f_2^*(p_2') \delta^4(\sum_i q_i - \tilde{p}_1 - \tilde{p}_2) \delta^4(\sum q_i - p_1' - p_2') \right\}. \end{aligned}$$

Note that what we have written as the square of the matrix element here is really

$$|M|^2 = \langle q_1 \ldots q_n |\, \tilde{p}_1 \tilde{p}_2\rangle \langle p_1' p_2' |\, q_1 \ldots q_n\rangle\,.$$

We'll clean this up later to write everything in terms of the physical values p and q rather than dummy variables $p', \tilde{p}$.

This expression for W is the transition probability for $2 \to n$ scattering to states of definite momentum $q_1 \dots q_n$. We can expand one of the delta functions in Fourier space to write

$$W = \int d^4x \int \frac{d^3\tilde{p}_1}{(2\pi)^3 2E_{\tilde{p}_1}} f_1(\tilde{p}_1) e^{i\tilde{p}_1\cdot x} \frac{d^3\tilde{p}_2}{(2\pi)^3 2E_{\tilde{p}_2}} f_2(\tilde{p}_2) e^{i\tilde{p}_2\cdot x}$$
$$\times \frac{d^3p_1'}{(2\pi)^3 2E_{p_1'}} f_1^*(p_1') e^{ip_1'\cdot x} \frac{d^3p_2'}{(2\pi)^3 2E_{p_2'}} f_2^*(p_2') e^{ip_2'\cdot x}$$
$$\times \delta^4(\sum_i q_i - p_1' - p_2').$$

Using the normalization we define the Fourier transform of the wavepacket,

$$|\psi_i\rangle \equiv \int \frac{d^3p}{(2\pi)^3\sqrt{2E_p}} f_i(p) e^{-ip\cdot x} |p\rangle .$$

What we've called the matrix element $|M|^2$ is still a function of $\tilde{p}_1, \tilde{p}_2, p_1', p_2', q_i$, but one can use the notion of sharp peaks (i.e. in our distributions f_i) to approximate $|M^2|$ by its value where $\tilde{p}_i = p_i' = p_i$. That is, our momentum distributions f_i are localized around some values p_i, so they behave similarly to delta functions and we can set all the dummy variables to the physical momenta p_1, p_2. Then the transition probability becomes

$$W = \int d^4x \frac{|\psi_1(x)|^2}{2E_1} \frac{|\psi_2(x)|^2}{2E_2} (2\pi)^4 \delta^4(\sum_i q_i - p_1 - p_2) |M|^2,$$

which means that the wavepacket in position space has some corresponding spread– like momentum, it is localized and not a single value. The total transition probability is a function of the spread in momentum $f_i(p)$ as well as the momenta themselves, $|M|^2$. Thus

$$\frac{dW}{d^4x} = \frac{|\psi_1(x)|^2}{2E_1} \frac{|\psi_2(x)|^2}{2E_2} (2\pi)^4 \delta^4(\sum_i q_i - p_1 - p_2) |M|^2,$$

where $|M|^2$ is now the actual matrix element $|\langle q_1 \dots q_n | \, p_1 p_2 \rangle|^2$. We'll complete this discussion next time.

Lecture 14.

Saturday, November 3, 2018

After today's lecture, we'll be able to solve all the problems on Example Sheet 2 (in principle). Remark: on question 10b on sheet 2, the answer is incorrect. It should read *"Find $\frac{d\sigma}{dt}$ in terms of g, s, t, m, and M."* Note that the matrix element $\mathcal{M}$ and the amplitude A_{f_i} are the same thing (e.g. in Prof. Allanach's notes). Whew.

Okay, moving on. Last time we wrote down

$$\frac{dW}{d^4x} = \frac{|\psi_1(x)|^2}{2E_1} \frac{|\psi_2(x)|^2}{2E_2} (2\pi)^4 \delta^4(\sum_i q_i - p_1 - p_2) |M|^2,$$

which is the *transition probability density per unit time*. It depends (perhaps very weakly) on x. Here's the picture we should imagine– we have a "beam" of particle 2, described in space by a wavefunction $|\psi_2(x)|^2$ and moving with velocity v. Thus the flux of particle 2 per unit area is $\phi = v|\psi_2(x)|^2$. In the rest frame of particle 1, we have a density of particle 1 $\rho = |\psi_1(x)|^2$, and it has some effective cross-sectional area $d\sigma$. Therefore we can rewrite this probability density as

$$\frac{dW}{d^4x} = d\sigma \cdot \phi \cdot \rho.$$

Equivalently we write the differential cross section as

$$d\sigma = \frac{(2\pi)^4 \delta^4(p_1 + p_2 - \sum_i q_i)}{\mathcal{F}} |M|^2$$

where $\mathcal{F} = 4E_1E_2v$ is the "flux factor." Thus $d\sigma$ is the effective cross-sectional area to scatter into final states of momenta $\{q_i\}$. If we now boost to the rest frame of particle 2, in this frame the four-momenta take the form

$$p_2^\mu = (m_2, 0), \quad p_1^\mu = (\sqrt{m_1^2 + |\mathbf{p}_1|^2}, \mathbf{p}_1).$$

The relative velocity $v = |\mathbf{p}_1|/E_1$, so in this frame the flux factor takes the form

$$\mathcal{F} = 4E_1|\mathbf{p}_1| = 4m_2\sqrt{E_1^2 - m_1^2} = 4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2},$$

where we have used the fact that in this frame $p_1 \cdot p_2 = E_1 m_2$. This is the correct Lorentz invariant definition of the flux factor.

In the massless limit, $m_1, m_2 \ll E_1, E_2$. This is the case for high-energy colliders like the LHC ($\sqrt{s} = 13\,\text{TeV}$, while $m_p \sim 1\,\text{GeV}$). In this limit, we therefore have

$$\mathcal{F} = 4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} \approx 4(p_1 \cdot p_2) \approx 2(m_1^2 + 2p_1 \cdot p_2 + m_2^2) = 2(p_1 + p_2)^2,$$

where we have added on and neglected mass terms rather freely in the limit where the masses are small compared to the $p_1 \cdot p_2$ term which is proportional to the energy E_1.

Then $\mathcal{F} \sim 2s$ where $s = (p_1 + p_2)^2$. To compute the total cross-section, we then sum over the $\{q_i\}$ in the correct manner to get

$$\sigma = \int \prod_{i=1}^{n} \left(\frac{d^3 q_i}{(2\pi)^3 2E_{q_i}} \right) \frac{|M|^2}{\mathcal{F}} (2\pi)^4 \delta^4(p_1 + p_2 - \sum_{i=1}^{n} q_i).$$

We call the integrals over $d^3 q_i$ "phase space integrals."

$2 \to 2$ **scattering** Let us specialize in the case of 2 to 2 scattering. What is the behavior of the differential cross-section, e.g. in terms of the Mandelstam variables? Let's look at the variations of σ with respect to

$$t = (p_1 - q_1)^2 = m_1^2 + {m_1'}^2 - 2E_{p_1}E_{q_1} + 2\mathbf{p}_1 \cdot \mathbf{q}_1.$$

Notice that

$$\frac{dt}{d\cos\theta} = 2|\mathbf{p}_1||\mathbf{q}_1|,$$

where $\cos\theta$ is the angle between $\mathbf{p}_1$ and $\mathbf{q}_1$. But θ is a frame-dependent quantity, so we must be a little careful what frame we're working in. Let us instead write the integration measure

$$\frac{d^3 q_2}{2E_{q_2}} = d^4 q_2 \delta(q_2^2 - {m_2'}^2)\theta(q_2^0)$$

with θ the step function. We proved this way back in Lecture 5, in a somewhat different form. What we wrote then was

$$\frac{d^3 q_2}{2E_{q_2}} = d^4 q_2 \delta((q_2^0)^2 - \mathbf{q}_2^2 - {m_2'}^2)|_{q_2^0 > 0}.$$

But this is clearly equivalent– just turn the q_2^0 condition into a step function and rewrite $(q_2^0)^2 - \mathbf{q}_2^2$ in terms of the four-momentum q_2^2. We then rewrite the $d^3 q_1$ integral in spherical coordinates for q_1:

$$\frac{d^3 q_1}{2E_{q_1}} = \frac{|\mathbf{q}_1|^2 d|\mathbf{q}_1|}{2E_{q_1}} d\cos\theta d\phi.$$

Since $E_{q_1}^2 + m_1^2 = {m_1'}^2 + |\mathbf{q}_1|^2 \implies 2E_{q_1}dE_{q_1} = |\mathbf{q}_1|d|\mathbf{q}_1|$ allows us to rewrite our expression for $\frac{d^3 q_1}{2E_{q_1}}$ (using the $dt/d\cos\theta$ expression) as

$$\frac{d^3 q_1}{2E_{q_1}} = \frac{1}{4|\mathbf{p}_1|} dE_{q_1} d\phi dt.$$

If we explicitly substitute our expressions for $d^3 q_1/2E_{q_1}$ and $d^3 q_2/2E_{q_2}$ into the expression for σ, we get

$$\sigma = \int \frac{1}{(2\pi)^2} \left(\frac{1}{4|\mathbf{p}_1|} dE_{q_1} d\phi dt \right) \left(d^4 q_2 \delta(q_2^2 - {m_2'}^2)\theta(q_2^0) \right) \frac{|M|^2}{\mathcal{F}} \delta^4(p_1 + p_2 - (q_1 + q_2)).$$

The ϕ integral is trivial– it cancels a factor of 2π. The q_2 integral is also trivial by the last delta function– since it just sets $q_2 = q_1 - p_1 - p_2$. (All the step function tells us is that the energy of the final state is non-negative.) We now take the derivative d/dt of both sides to get an expression for $d\sigma/dt$:

$$\frac{d\sigma}{dt} = \frac{1}{8\pi\mathcal{F}|\mathbf{p}_1|}\int dE_{q_1}|M|^2\delta((q_1-\sqrt{s})^2 - {m_2'}^2).$$

Expanding out the square we find that

$$(q_1-\sqrt{s})^2 - {m_2'}^2 = q_1^2 - 2q_1\cdot(p_1+p_2) + s - {m_2'}^2,$$

so our final expression is

$$\frac{d\sigma}{dt} = \frac{1}{8\pi\mathcal{F}|\mathbf{p}_1|}\int dE_{q_1}|M|^2\delta(s - {m_2'}^2 + {m_1'}^2 - 2q_1\cdot(p_1+p_2)).$$

Boosting now to the center of mass frame where $p_1^\mu = (\sqrt{|\mathbf{p}_1|^2+m_1^2}, \mathbf{p}_1)$ and $p_2^\mu = (\sqrt{|\mathbf{p}_1|^2+m_2^2}, -\mathbf{p}_1)$, we note that s is some constant of the collision,

$$s = \left(\sqrt{|\mathbf{p}_1|^2+m_1^2} + \sqrt{|\mathbf{p}_1|^2+m_2^2}\right)^2.$$

We can solve for $|\mathbf{p}_1|$ as an exercise (see the end of this section) to find

$$|\mathbf{p}_1| = \frac{\lambda^{1/2}(s,m_1^2,m_2^2)}{2\sqrt{s}}$$

where

$$\lambda(x,y,z) \equiv x^2+y^2+z^2-2xy-2xz-2yz.$$

We therefore find that

$$\mathcal{F} = 2\lambda^{1/2}(s,m_1^2,m_2^2).$$

With our expressions for $|\mathbf{p}_1|$ and $\mathcal{F}$ firmly in hand, we can plug them back into our expression for $d\sigma/dt$, we get

$$\frac{d\sigma}{dt} = \frac{|M|^2}{16\pi\lambda(s,m_1^2,m_2^2)(1/2\sqrt{s})}\int dE_{q_1}\delta(s - {m_2'}^2 + {m_1'}^2 - 2q_1\cdot(p_1+p_2)).$$

Since we are in the center-of-mass frame, $p_1+p_2 = (m_1+m_2,0,0,0) = (\sqrt{s},0,0,0)$, and so

$$\begin{aligned}\frac{d\sigma}{dt} &= \frac{|M|^2}{16\pi\lambda(s,m_1^2,m_2^2)(1/2\sqrt{s})}\int dE_{q_1}\delta(s - {m_2'}^2 + {m_1'}^2 - 2E_{q_1}\sqrt{s})\\ &= \frac{|M|^2}{16\pi\lambda(s,m_1^2,m_2^2)(1/2\sqrt{s})}\int d\tilde{E}_{q_1}\frac{1}{2\sqrt{s}}\delta(s - {m_2'}^2 + {m_1'}^2 - \tilde{E}_{q_1})\\ &= \frac{|M|^2}{16\pi\lambda(s,m_1^2,m_2^2)}.\end{aligned}$$

In the massless limit (a common approximation) we have $t = (p_1-q_1)^2 - 2p_1\cdot q_1 = -2|\mathbf{p}_1||\mathbf{q}_1|(1-\cos\theta)$, and the total cross section is

$$\sigma_{tot} = \int_{-4|\mathbf{p}_1||\mathbf{q}_1|}^{0} dt\frac{d\sigma}{dt}.$$

In the center-of-mass frame $|\mathbf{p}_1| = |\mathbf{q}_1| = \sqrt{s}/2$ so $\frac{dt}{d\cos\theta} = \frac{s}{2}$. Defining the differential solid angle element $d\Omega$ by

$$d\Omega \equiv d\cos\theta d\phi$$

(a frame-dependent quantity) we find that

$$\frac{d\sigma}{d\Omega} = \frac{s}{4\pi}\frac{d\sigma}{dt} = \frac{|M|^2}{64\pi^2 s}$$

for particles with masses much less than the collision energy.[41]

We can also consider decay rates, which we treat much the same way. Take the initial state to be a sharply peaked superposition of momentum-space eigenstates. Our transition probability density is

$$\frac{dW}{d^4x} = \frac{|\psi(x)|^2}{2E_p}|M|^2(2\pi)^4\delta^4(p - \sum_i q_i),$$

where $\psi(x)$ is the space-time wavefunction of the decaying particle. dW/d^x is then the chance of finding the decaying particle per unit volume. We can equivalently define the differential decay rate $d\Gamma$ such that

$$\frac{dW}{d^4x} = |\psi(x)|^2 \times d\Gamma.$$

Thus

$$\Gamma = \frac{1}{2E_p}\int \prod_{i=1}^{n}\left(\frac{d^3q_i}{(2\pi)^3 2E_{q_i}}\right)|M|^2(2\pi)^4\delta^4(p - \sum_{i=1}^{n} q_i).$$

Note that Γ is *not* Lorentz invariant, as it goes as $1/E$ of the decaying particle. The standard convention is to define Γ in the rest frame of the decaying particle. The lifetime of a particle is given by

$$\tau = 6.58 \times 10^{-25}\ \text{seconds} \times \frac{1\,\text{GeV}}{\Gamma}.$$

To link this back to our previous discussion of nucleon scattering, $\psi\psi \to \psi\psi$, we computed two diagrams for this process. We found that the matrix element was

$$iM = (-ig)^2\left\{\frac{1}{t-m^2} + \frac{1}{u-m^2}\right\},$$

with t and u the standard Mandelstam variables.

Non-lectured aside– solving for $|\mathbf{p}_1|$ We have

$$s = \left(\sqrt{|\mathbf{p}_1|^2 + m_1^2} + \sqrt{|\mathbf{p}_1|^2 + m_2^2}\right)^2.$$

To solve for $|\mathbf{p}_1|$, let us expand out

$$s = (|\mathbf{p}_1|^2 + m_1^2) + (|\mathbf{p}_1|^2 + m_2^2) + 2\sqrt{(|\mathbf{p}_1|^2 + m_1^2)(|\mathbf{p}_1|^2 + m_2^2)}.$$

We move all the terms outside the square root to the LHS to get

$$\frac{s - 2|\mathbf{p}_1|^2 - m_1^2 - m_2^2}{2} = \sqrt{(|\mathbf{p}_1|^2 + m_1^2)(|\mathbf{p}_1|^2 + m_2^2)}$$

and square again to get rid of all the square roots. We can then expand the left side in a useful way, writing

$$\frac{(s - m_1^2 - m_2^2)^2}{4} + |\mathbf{p}_1|^4 - |\mathbf{p}_1|^2(s - m_1^2 - m_2^2) = (|\mathbf{p}_1|^2 + m_1^2)(|\mathbf{p}_1|^2 + m_2^2)$$

or equivalently

$$\frac{(s - m_1^2 - m_2^2)^2}{4} + |\mathbf{p}_1|^4 - |\mathbf{p}_1|^2(s - m_1^2 - m_2^2) = |\mathbf{p}_1|^4 + (m_1^2 + m_2^2)|\mathbf{p}_1|^2 + m_1^2m_2^2.$$

The $|\mathbf{p}_1|^4$ terms cancel, as do the $(m_1^2m_2^2)|\mathbf{p}_1|^2$s, so we are left with

$$s|\mathbf{p}_1|^2 = \frac{(s - m_1^2 - m_2^2)^2}{4} - m_1^2m_2^2.$$

A little rearranging yields

$$|\mathbf{p}_1|^2 = \frac{(s - m_1^2 - m_2^2)^2 - 4m_1^2m_2^2}{4s} = \frac{\lambda(s, m_1^2, m_2^2)}{4s}. \quad \boxtimes$$

[41]Solid angle is the generalization of angles in the plane. A normal angle measured in radians corresponds to an arc length subtended by that angle on a circle of unit radius. In the same way, solid angle (measured in steradians) can be thought of as a surface area on a 2-sphere of unit radius, so that the total solid angle for a sphere is 4π.

Now we want to solve for $\mathcal{F}$. Note that

$$2p_1 \cdot p_2 = (p_1 + p_2)^2 - m_1^2 - m_2^2 = s - m_1^2 - m_2^2.$$

Then we can get $\mathcal{F}$ by writing

$$\begin{aligned}\mathcal{F} &= 4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} \\ &= 2\sqrt{(2p_1 \cdot p_2)^2 - 4m_1^2 m_2^2} \\ &= 2\sqrt{(s - m_1^2 - m_2^2)^2 - 4m_1^2 m_2^2} \\ &= 2\lambda^{1/2},\end{aligned}$$

where we have recognized λ from the first calculation for $|\mathbf{p}_1|$. ☒

Lecture 15.

Tuesday, November 6, 2018

Today, we'll introduce spinors, the mathematical framework describing the behavior of fermions! We'll start to show explicitly why spin 1/2 is different than spin 0.[42]

Now, so far we've only considered scalar fields ϕ. Under a Lorentz transformation, these transform as

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

Most particles have an intrinsic angular momentum, which we call *spin*, and fields with spin have a non-trivial Lorentz transformation. For instance, spin 1 particles (i.e. *vector fields*) come with an index μ and transform as

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

In general fields can transform as $\phi^a \to D^a{}_b(\Lambda)\phi^b(\Lambda^{-1}x)$, where we say the $D^a{}_b$ form a representation of the Lorentz group. These might be familiar from *Symmetries, Fields and Particles*, but to give a quick overview, a representation D of a group g is a map from that group to a space of linear transformations (usually taken to be matrices) which preserves the group multiplication. That is, it satisfies

$$D(\Lambda_1\Lambda_2) = D(\Lambda_1)D(\Lambda_2)$$
$$D(\Lambda^{-1}) = (D(\Lambda))^{-1}$$
$$D(I) = I.$$

To find the representations, we look at the Lorentz algebra by considering infinitesimal Lorentz transformations. If we write

$$\Lambda^\mu{}_\nu = \delta^\mu_\nu + \epsilon\omega^\mu{}_\nu + O(\epsilon^2),$$

then the property that Λ preserves the inner product on four-vectors implies that $\omega_{\mu\nu}$ is a 4×4 antisymmetric matrix. In particular this means it has $\frac{4\times 3}{2} = 6$ independent components, corresponding to the three rotations and three Lorentz boosts.

We may introduce a basis of six 4×4 matrices, which we will label by four indices

$$(M^{\rho\sigma})^{\mu\nu} = \eta^{\rho\mu}\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu},$$

where these matrices are antisymmetric in ρ, σ and in μ, ν. We take ρ, σ to specify which matrix we are looking at. Lowering the index ν, we take μ, ν to specify the row and column respectively. Therefore

$$(M^{\rho\sigma})^\mu{}_\nu = \eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu.$$

[42] It's pretty cool to learn about this in Cambridge, where Dirac actually discovered the behavior of spin 1/2 particles.

Example 15.1. The basis vector $(M^{01})^\mu{}_\nu$ is given by

$$(M^{01})^\mu{}_\nu = \begin{pmatrix} 0 & +1 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This generates a boost in the x^1 direction (it mixes up x^1 and t).

Similarly, the basis vector $(M^{12})^\mu{}_\nu$ takes the form

$$(M^{12})^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This generates rotations in the $(x^1 - x^2)$ plane.

Note that when we lower ν in order to write the generators as matrices, the matrix may not explicitly look antisymmetric! We can now write

$$\omega^{\mu\nu} = \frac{1}{2}(\Omega_{\rho\sigma} M^{\rho\sigma})^\mu{}_\nu$$

where these Ms are the generators of the group of Lorentz transformations and Ω is some set of antisymmetric parameters.

Definition 15.2. The *Lorentz algebra* is a set of relations between matrices M defined by the bracket

$$[M^{\rho\sigma}, M^{\tau\nu}] = \eta^{\sigma\tau} M^{\rho\nu} - \eta^{\rho\tau} M^{\sigma\nu} + \eta^{\rho\nu} M^{\sigma\tau} - \eta^{\sigma\nu} M^{\rho\tau}.$$

The spinor representation means that we search for other matrices satisfying the Lorentz algebra.

Definition 15.3. We define the Clifford algebra (in any number of dimensions we like, though four is the most useful for our purposes) as a set of matrices γ^μ such that

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} I,$$

where we have defined the anticommutator $\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu$. In four dimensions, the γ^μ are a set of 4×4 matrices with $\mu = 0, 1, 2, 3$.

We need to find 4 matrices which anticommute, and such that $(\gamma^i)^2 = -I \forall i \in \{1, 2, 3\}$ and $(\gamma^0)^2 = I$. The simplest representation is in terms of 4×4 matrices. A common choice is the *chiral* or *Weyl representaion*, where we take

$$\gamma^0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix},$$

where the σ^i are the usual 2×2 Pauli matrices. As a quick refresher, the Pauli matrices are

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

They satisfy the commutation and anticommutation relations

$$[\sigma^i, \sigma^j] = 2i e^{ijk} \sigma^k \text{ and } \{\sigma^i, \sigma^j\} = 2\delta^{ij} I_2.$$

Note that the γ matrices under any similarity transformation $U\gamma^\mu U^{-1}$ (where U is an invertible constant matrix) also forms an equally good basis.

We now define

$$S^{\rho\sigma} \equiv \frac{1}{4}[\gamma^\rho, \gamma^\sigma] = \frac{1}{2}\gamma^\rho\gamma^\sigma - \frac{1}{2}\eta^{\rho\sigma}$$

by the Clifford algebra. We'll make the following claims: first,

$$[S^{\mu\nu}, \gamma^\rho] = \gamma^\mu \eta^{\nu\rho} - \gamma^\nu \eta^{\rho\mu}.$$

Second, using the previous claim and the definition of S, we can prove (e.g. on the example sheet) that S satisfies the commutation relation

$$[S^{\rho\sigma}, S^{\tau\nu}] = \eta^{\sigma\tau} S^{\rho\nu} - \eta^{\rho\tau} S^{\sigma\nu} + \eta^{\rho\nu} S^{\sigma\tau} - \eta^{\sigma\nu} S^{\rho\tau}.$$

But this is precisely the relations that the Lorentz group generators satisfy, and so S provides a representation of the Lorentz algebra.[43]

We now introduce a four-component *Dirac spinor* $\psi_\alpha(x), \alpha \in \{1,2,3,4\}$. The spinor then transforms under Lorentz transformations as

$$\psi^\alpha(x) \to S[\Lambda]^\alpha{}_\beta \psi^\beta(\Lambda^{-1}x).$$

Here,

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \text{ and } \Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right)$$

are both 4×4 matrices.

Is the spinor representation equivalent to the usual vector representation? No– one can look at specific Lorentz transformations to see this. For instance, the rotations $i, j \in \{1,2,3\}$ give

$$\begin{aligned} S^{ij} &= \frac{1}{4}\left[\gamma^i, \gamma^j\right] \\ &= \left[\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix}\right] \\ &= \frac{-i}{2}\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \end{aligned}$$

If we write $\Omega_{ij} = -\epsilon_{ijk}\phi^k$, where ϕ^k is a vector specifying a rotation axis, e.g. $\Omega_{12} = -\phi^3$. Then

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) = \begin{pmatrix} e^{i\boldsymbol{\phi}\cdot\boldsymbol{\sigma}/2} & 0 \\ 0 & e^{i\boldsymbol{\phi}\cdot\boldsymbol{\sigma}/2} \end{pmatrix}.$$

Therefore a rotation about the x^3 axis can be written as $\phi = (0,0,2\pi)$, and then

$$S[\Lambda] = \begin{pmatrix} e^{i\sigma^3\pi} & 0 \\ 0 & e^{i\sigma^3\pi} \end{pmatrix}. = -I_4.$$

Therefore a rotation of 2π takes $\psi_\alpha(x) \to -\psi_\alpha(x)$. This is different from the vector representation, where

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right) = \exp\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2\pi & 0 \\ 0 & -2\pi & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = I_4,$$

as expected. So indeed spinors see the full $SU(2)$ rotational symmetry, and not just the $SO(3)$ symmetry of the ordinary Lorentz group.

What about boosts? Let us take

$$S^{0i} = \frac{1}{2}\begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}$$

and write our boost parameter $\Omega_{0i} = -\Omega_{i0} \equiv \chi_i$. Then

$$S[\Lambda] = \begin{pmatrix} e^{-\boldsymbol{\chi}\cdot\boldsymbol{\sigma}/2} & 0 \\ 0 & e^{-\boldsymbol{\chi}\cdot\boldsymbol{\sigma}/2} \end{pmatrix}.$$

For rotations, $S[\Lambda]$ is unitary since $S[\Lambda]S[\Lambda]^\dagger = I$, but for boosts this is *not* the case.

It turns out there are no finite-dimensional unitary representations of the Lorentz group: this is because the matrices

$$S[\Lambda] = \exp\left[\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right]$$

are only unitary if the $S^{\mu\nu}$ are anti-hermitian, $(S^{\mu\nu})^\dagger = -S^{\mu\nu}$. But

$$(S^{\mu\nu})^\dagger = -\frac{1}{4}[\gamma^{\mu\dagger}, \gamma^{\nu\dagger}]$$

[43] At this point, Professor Allanach made a slight digression to read from an interview with Dirac conducted by an USAmerican journalist. It's entertaining reading and can be found here: `http://sites.math.rutgers.edu/~greenfie/mill_courses/math421/int.html`

can be anti-hermitian if all the γ^μs are either all hermitian or all anti-hermitian. However, this can't be arranged, since $\gamma^0)^2 = I \implies \gamma^0$ has real eigenvalues (and cannot be anti-hermitian), whereas $(\gamma^i)^2 = -I \implies \gamma^i$ has purely imaginary eigenvalues, and therefore cannot be hermitian.

Lecture 16.

Thursday, November 8, 2018

Today we will construct a LI action of spinor fields. Suppose we have a complex field ψ, with

$$\psi^\dagger(x) = (\psi^*)^T(x).$$

Is $\psi^\dagger(x)\psi(x)$ a Lorentz scalar? We'll check how it transforms. In general we have

$$\psi^\dagger(x)\psi * x) \to \psi^\dagger(\Lambda^{-1}x)\underbrace{S[\Lambda]^\dagger S[\Lambda]}_{\neq 1}\psi(\Lambda^{-1}x),$$

which is not quite what we want, since S is not unitary. Since $\gamma^0 = (\gamma^0)^\dagger$ is hermitian and $\gamma^i = -(\gamma^i)^\dagger$ is antihermitian in our representation, we have

$$\gamma^0\gamma^\mu\gamma^0 = (\gamma^\mu)^\dagger \implies (S^\mu\nu)^\dagger = -\frac{1}{4}[\gamma^{\mu\dagger}, \gamma^{\nu\dagger}] = -\gamma^0 S^{\mu\nu}\gamma^0.$$

(Note that Greek indices run from 0 to 3 here, while Latin indices are $1, 2, 3$. As they should be.)

Thus

$$S[\Lambda]^\dagger = \exp(\frac{1}{2}\Omega_{\mu\nu}(S^{\mu\nu})^\dagger) = \gamma^0 S[\Lambda]^{-1}\gamma^0,$$

which we get by using $(\gamma^0)^2 = 1$ repeatedly.

Definition 16.1. With this in mind, we define a *Dirac adjoint* of ψ:

$$\bar\psi(x) \equiv \psi^\dagger(x)\gamma^0.$$

We now claim that $\bar\psi(x)\psi(x)$ *is* a Lorentz scalar. Writing explicitly,

$$\begin{aligned}\bar\psi(x)\psi(x) &= \psi^\dagger(x)\gamma^0\psi(x)\\ &\to \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger\gamma^0 S[\Lambda]\psi(\Lambda^{-1}(x)\\ &= \psi^\dagger(\Lambda^{-1}x)\gamma^0\psi(\Lambda^{-1}(x)\\ &= \bar\psi(\Lambda^{-1}x)\psi(\Lambda^{-1}x). \quad \boxtimes\end{aligned}$$

Moreover, we claim that $\bar\psi(x)\gamma^\mu\psi(x)$ is a Lorentz vector. Under a Lorentz transformation, it transforms as

$$\bar\psi(\Lambda^{-1}x)S[\Lambda]^\dagger\gamma^\mu S[\Lambda]\psi(\Lambda^{-1}(x).$$

If this is to be a Lorentz vector, we must have

$$S[\Lambda]^{-1}\gamma^\mu S[\Lambda] = \Lambda^\mu_\nu\gamma^\nu.$$

Now we know that

$$\Lambda^\mu_\nu = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right)^\mu_{\ \nu}$$

and

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right),$$

so infinitesimally we have

$$(M^{\rho\sigma})^\mu_\nu\gamma^\nu = -[S^{\rho\sigma}, \gamma^\mu].$$

But from the definition of M, we have on the LHS

$$(\eta^{\rho\mu}\delta^\sigma_\nu - \eta^{\sigma\mu}\delta^\rho_\nu)\gamma^\nu = \eta^{\rho\mu}\gamma^\sigma - \gamma^\rho\eta^{\sigma\mu} = -[S^{\rho\sigma}, \gamma^\mu],$$

which we proved previously.

Now we'll claim that

$$S = \int d^4x \, \underbrace{\bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x)}_{\mathcal{L}_D}$$

is a LI action, where $\mathcal{L}_D$ is the *Dirac Lagrangian*. This action describes a free spinor field, and it has some strange properties. If we look at the mass dimension of the field with $[m] = 1$, we find that $[\psi] = [\bar{\psi}] = \frac{3}{2}$. We can now vary $\psi, \bar{\psi}$ independently to get the equations of motion. Varying ψ, we find that

$$(i\gamma^\mu \partial_\mu - m)\psi = 0,$$

which is known as the *Dirac equation.* Note that this equation is only first-order in ∂_μ, whereas the scalar field yielded a second-order equation in ∂_μ. One arrives at a similar equation of $\bar{\psi}$ after an integration by parts:

$$i\partial_\mu \gamma^\mu \bar{\psi} + m\bar{\psi} = 0.$$

Let us now introduce the *slash notation*:

$$A_\mu \gamma^\mu = \gamma_\mu A^\mu = \not{A}.$$

Hence the Dirac equation is written

$$(i\not{\partial} - m)\psi = 0.$$

Note that the Dirac equation mixes up different components of ψ, but each individual component solves the Klein-Gordon equation:

$$\begin{aligned}(i\not{\partial} + m)(i\not{p} - m)\psi = 0 &\implies -(\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + m^2)\psi = 0 \\ &\iff -(\frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\partial_\mu \partial_\nu + m^2)\psi = 0 \\ &\iff -(\partial_\mu \partial^\mu + m^2)\psi = 0.\end{aligned}$$

Remember, we should think of the spinor as secretly four components with a non-trivial transformation under rotations.

Now in our representation (the chiral representation), $S[\Lambda]$ is block diagonal. It takes the form

$$S[\Lambda] = \begin{cases} \begin{pmatrix} e^{i\boldsymbol{\phi}\cdot\boldsymbol{\sigma}/2} & 0 \\ 0 & e^{i\boldsymbol{\phi}\cdot\boldsymbol{\sigma}/2} \end{pmatrix} & \text{for rotations,} \\ \begin{pmatrix} e^{-\boldsymbol{\chi}\cdot\boldsymbol{\sigma}/2} & 0 \\ 0 & e^{-\boldsymbol{\chi}\cdot\boldsymbol{\sigma}/2} \end{pmatrix} & \text{for boosts.} \end{cases}$$

From *Symmetries*, we might recall that since the representation takes a block diagonal form, it is *reducible*, i.e. it decomposes into two *irreducible* representations acting on U_L, U_R, where we now write

$$\psi = \begin{pmatrix} U_L \\ U_R \end{pmatrix}$$

with U_L, U_R some 2-component $\mathbb{C}$ objects. We call U_L and U_R (where L, R stand for left and right) *Weyl* or *chiral spinors.* They transform identically under rotations,

$$U_{L,R} \to e^{i\boldsymbol{\phi}\cdot\boldsymbol{\sigma}/2} U_{L,R}$$

but oppositely under boosts,

$$\begin{aligned} U_L &\to e^{-\boldsymbol{\chi}\cdot\boldsymbol{\sigma}/2} U_L \\ U_R &\to e^{+\boldsymbol{\chi}\cdot\boldsymbol{\sigma}/2} U_R. \end{aligned}$$

In group theory language, we say that U_L is in the $(1/2, 0)$ representation of the Lorentz group, while U_R is in the $(0, 1/2)$ representation (where the Lorentz group $SO(1,3) \simeq SU(2) \times SU(2)$). A general spinor is in the direct product space,

$$\psi = (1/2, 0) \oplus (0, 1/2).$$

The Weyl equation Let us now decompose the Dirac Lagrangian $\mathcal{L}_D$ in terms of Weyl spinors. Thus

$$\mathcal{L}_D = \bar{\psi}(i\partial\!\!\!/ - m)\psi = iU_L^\dagger \sigma^\mu \partial_\mu U_L + iU_R^\dagger \bar{\sigma}^\mu \partial_\mu U_R - m(U_L^\dagger U_R + U_R^\dagger U_L),$$

where $\sigma^\mu \equiv (I, \boldsymbol{\sigma}), \bar{\sigma}^\mu \equiv (I, -\boldsymbol{\sigma})$. We observe that the kinetic terms separate entirely– it is only the mass term which mixes U_L and U_R. A massive spinor requires both U_L and U_R in general, but a massless fermion only requires a single one (e.g. U_L). This leads us to write

$$i\sigma^\mu \partial_\mu U_L = 0,$$
$$i\bar{\sigma}^\mu \partial_\mu U_R = 0,$$

which are known as Weyl's equations.

Naïvely, we expect that since U_L and U_R each have two complex components, our count of the real degrees of freedom should come out to $2 \times 2 \times 2 = 8$. But it turns out this is not quite right. In classical mechanics, the number of degrees of freedom are typically given by

$$\#\text{ d.o.f.} = \frac{1}{2} \times (\text{dimensionality of phase space.}).$$

In field theory, we discuss instead the d.o.f. per spacetime point. For a real scalar ϕ, the conjugate momentum is $\Pi_\phi = \dot{\phi} \implies \#\text{ d.o.f.} = \frac{1}{2} \times (2) = 1$. However, for a spinor we have $\Pi_\psi = \psi^\dagger$, not $\dot{\psi}$. Therefore we get 4 complex components = 8 real degrees of freedom is ψ, but no extra in $\psi^\dagger$. The upshot is that for spinors,

$$\#\text{ d.o.f.} = \frac{1}{2}(8) = 4.$$

We can choose spin $\uparrow$ or spin $\downarrow$, and consider particles or antiparticles, so $2 \times 2 = 4$. We'll explore what happens to the extra degrees of freedom next time.

Lecture 17.

Saturday, November 10, 2018

We saw in the chiral representation that the spinor representation $S[\Lambda]$ was block diagonal, but this is not always true. As far as we are concerned, any repn related to the original γ^μs by a similarity transformation is equally good,

$$\gamma^\mu \to U\gamma^\mu U^{-1}.$$

However, what we would like is a repn independent way to define the Weyl spinors. As it turns out, we can do this by defining the matrix γ^5 as

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3.$$

One can check (quick exercise) that γ^5 satisfies

$$\{\gamma^\mu, \gamma^5\} = 0 \text{ and } (\gamma^5)^2 = I.$$

In the chiral repn, γ^5 takes the form

$$\gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}.$$

Let us define the projection operators

$$P_L \equiv \frac{1}{2}(I - \gamma^5)$$
$$P_R \equiv \frac{1}{2}(I + \gamma^5).$$

Note this convention is slightly different than the one in David Tong's notes. Under these definitions, one can see that

$$P_L^2 = P_L, \quad P_R^2 = P_R, \quad \text{and } P_L P_R = 0.$$

From the explicit form of γ^5, it's clear that these projection operators select the corresponding Weyl spinors.

Definition 17.1. We define a *left-handed spinor* as

$$\psi_L = P_L \psi,$$

and similarly a *right-handed spinor* as

$$\psi_R = P_R \psi.$$

For instance, neutrinos are left-handed spinors, while the right-handed up quark is a right-handed spinor.

Handedness can be more physically defined in terms of whether a particle's spin is parallel or antiparallel to its velocity. Note that *massive* left-handed and right-handed spinors can be transformed into one another by a Lorentz boost, since we can boost into the rest frame of a massive particle (for instance) or boost past it into a frame where it appears to be moving in the opposite direction. This is not the case for the massless Weyl spinors we defined before– you can never catch up with or overtake a massless particle, so its handedness is invariant under Lorentz transformations.

One can now construct new tensors using γ^5, e.g.

$$\bar{\psi}(x)\gamma^t\psi(x) \to_{LT} \bar{\psi}(\Lambda^{-1}x)S[\Lambda]^{-1}\gamma^5 S[\Lambda]\psi(\Lambda^{-1}x).$$

We call such a quantity a *pseudoscalar,* and one can check that $[S_{\mu\nu}, \gamma^t] = 0$, so our pseudoscalar under a Lorentz transformation transforms to

$$\bar{\psi}(\Lambda^{-1}x)\gamma^5\psi(\Lambda^{-1}x).$$

However, it's not quite a scalar because of a subtle point we'll come to shortly. Similarly we can define

$$\bar{\psi}(x)\gamma^5\gamma^\mu\psi(x),$$

which we call an *axial vector*. These are distinguished from regular scalars and vectors by their behavior under a parity transformation.

We say that ψ_L, ψ_R are related by *parity*, i.e. a flip in handedness. So far, all the Lorentz transformations we've looked at were continuously connected to the identity (i.e. they are in the connected component of $O(3,1)$). In fact, there are also two discrete symmetries we can consider which leave four-vector inner products fixed.[44]

- Time reversal $T : x^0 \to -x^0, x^i \to x^i$
- Parity $P : x^0 \to x^0, x^i \to -x^i$.

Note that P involves flipping all three spatial components, since changing the sign of only one or two is equivalent to a rotation (and therefore that transformation is in the same connected component).

Under P, rotations don't change sign, but boosts do. That is,

$$\begin{cases} U_\pm \to e^{i\boldsymbol{\phi}\cdot\boldsymbol{\sigma}/2}U_\pm \text{ under rotations,} \\ U_\pm \to e^{\pm\boldsymbol{\chi}\cdot\boldsymbol{\sigma}/2}U_\pm \text{ under boosts.} \end{cases}$$

Therefore we see that P exchanges left- and right-handed spinors:

$$P : \psi_{L/R}(\mathbf{x}, t) \to \psi_{R/L}(-\mathbf{x}, t).$$

For a Dirac spinor, this is implemented by

$$P : \psi(\mathbf{x}, t) \to \gamma^0\psi(-\mathbf{x}, t).$$

The quantity $\bar{\psi}\psi(\mathbf{x}, t) \to \bar{\psi}\psi(-\mathbf{x}, t)$, so $\bar{\psi}\psi$ transforms under P like a scalar. Meanwhile, $\bar{\psi}\gamma^\mu\psi(\mathbf{x}, t)$ transforms as

$$\bar{\psi}\gamma^\mu\psi(\mathbf{x}, t) \to \begin{cases} \bar{\psi}\gamma^0\psi(-\mathbf{x}, t) & \mu = 0 \\ \bar{\psi}\gamma^0\gamma^i\gamma^0\psi(-\mathbf{x}, t) & \mu = i \\ = -\bar{\psi}\gamma^i\psi(-\mathbf{x}, t). \end{cases}$$

[44] The connected components of the Lorentz group have the structure of V_4, the Klein four-group. All these transformations do is switch around regions between parts of the light cone, if you like.

Coupling	Name	Number
$\bar{\psi}\psi$	scalar	1
$\bar{\psi}\gamma^\mu\psi$	vector	4
$\bar{\psi}S^{\mu\nu}\psi$	tensor	$4 \times 3/2 = 6$
$\bar{\psi}\gamma^5\psi$	pseudoscalar	1
$\bar{\psi}\gamma^5\gamma^\mu\psi$	pseudovector	4

TABLE 1. The different bilinears we can write down for spinors. They can be characterized by how they transform under Lorentz transformations and parity.

Therefore this transforms as a vector under P, with the spatial part flipping sign. Some other combinations we can cook up are the transformation of

$$\begin{aligned}\bar{\psi}\gamma^t\psi(\mathbf{x},t) &= \bar{\psi}\gamma^0\gamma^5\gamma^0\psi(-\mathbf{x},t)\\ &= -\bar{\psi}\gamma^5\psi(-\mathbf{x},t),\end{aligned}$$

so the minus sign here leads us to call this a pseudoscalar, while $\bar{\psi}\gamma^5\gamma^\mu\psi(\mathbf{x},t)$ transforms as

$$\begin{aligned}\bar{\psi}\gamma^5\gamma^\mu\psi(\mathbf{x},t) &\to \bar{\psi}\gamma^0\gamma^5\gamma^\mu\gamma^0\psi(-\mathbf{x},t)\\ &= \begin{cases} -\bar{\psi}\gamma^5\gamma^0\psi(-\mathbf{x},t) & \mu = 0\\ +\bar{\psi}\gamma^5\gamma^i\psi(-\mathbf{x},t) & \mu = i.\end{cases}\end{aligned}$$

The different sorts of bi-linears (combos of a $\bar{\psi}$ and ψ) we can cook up is summarized in Table 1. We can add extra terms to the Lagrangian using γ^5, but such terms break P symmetry. Nature in fact uses thes, e.g. terms in the Lagrangian like

$$\mathcal{L} = \frac{g}{2}W_\mu\bar{\psi}\gamma^\mu(1-\gamma^5)\psi.$$

For instance, the W boson is a vector field which only couples to left-handed fermions.

Definition 17.2. A theory which puts ψ_L, ψ_R on equal footing is called *vector-like.* If they appear differently/have different interactions, the theory is called *chiral.*

It turns out the Standard Model is a chiral theory– Chien-Shiung Wu famously observed in 1956 that parity symmetry is explicitly violated in beta decays.

Symmetries and currents of spinors We might now like to understand what the corresponding conserved quantities are that correspond to spinors. For instance, under spacetime translation we have

$$x^\mu \to x^\mu - \alpha\epsilon^\mu,$$

so

$$\Delta\psi = \epsilon^\mu\partial_\mu\psi \implies T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi - \eta^{\mu\nu}\mathcal{L}_D.$$

This comes from applying Noether's theorem straightforwardly to the Dirac Lagrangian. We get a conserved current when the equations of motion are obeyed, so we can impose them on $T^{\mu\nu}$. This doesn't help us in the bosonic case where the equations of motion are second-order, but for spinors it does because the equatiosn are first order. Thus

$$(i\partial\!\!\!/ - m)\psi = 0 \implies \text{we can set } \mathcal{L}_D = 0 \text{ in } T^{\mu\nu}.$$

Therefore

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi.$$

We can write $\mathcal{L}_D$ in a more symmetric way by splitting it up and integrating by parts,

$$S = \int d^4x \frac{1}{2}\bar{\psi}(i\partial\!\!\!/^{\rightarrow} - m)\psi + \frac{1}{2}\bar{\psi}(-i\partial\!\!\!/^{\leftarrow} - m)\psi = \frac{1}{2}\bar{\psi}(i\partial\!\!\!/^{\leftrightarrow} - m)\psi$$

where the $\leftrightarrow$ indicates a symmetrization. Thus we can write

$$T^{\mu\nu} = \frac{i}{2}\bar{\psi}(\gamma^\mu\partial^\nu - \gamma^\nu\partial^\mu)\psi.$$

Lecture 18.

Tuesday, November 13, 2018

Today, we'll continue our discussion of symmetries and currents of spinors. The spacetime translations give rise to a stress-energy tensor $T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi$. We can also consider the current associated to Lorentz transformations:

$$\psi^\alpha \to S[\Lambda]^\alpha{}_\beta \psi^\beta (x^\mu - \omega^\mu{}_\nu x^\nu)$$

where $\omega^\mu{}_\nu = \frac{1}{2}\Omega_{\rho\sigma}(M^{\rho\sigma})^\mu{}_\nu$. If we take $(M^{\rho\sigma})^\mu{}_\nu = \eta^{\rho\mu}\delta^\sigma_\nu - \eta^{\sigma\mu}\delta^\rho_\nu$, we get $\omega^{\mu\nu} = \Omega^{\mu\nu}$. Thus the variation in ψ^α is

$$\delta\psi^\alpha = -\omega^\mu{}_\nu x^\nu \partial_\mu \psi^\alpha + \frac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\alpha{}_\beta \psi^\beta.$$

Substituting for $\Omega^{\mu\nu}$ and pulling out $\omega^{\mu\nu}$, we find that

$$\delta\psi^\alpha = -\omega^{\mu\nu}[x_\nu \partial_\mu \psi^\alpha - \frac{1}{2}(S_{\mu\nu})^\alpha{}_\beta \psi^\beta].$$

Similarly, the variation in $\bar{\psi}$ is given by

$$\delta\bar{\psi}_\alpha = -\omega^{\mu\nu}[x_\nu \partial_\mu \bar{\psi}_\alpha + \frac{1}{2}\bar{\psi}_\beta (S_{\mu\nu})^\beta{}_\alpha].$$

Note that this last term comes with a plus sign for $\bar{\psi}$. Applying Noether's theorem, we get the conserved currents

$$(J^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} - i\bar{\psi}\gamma^\mu S^{\rho\sigma}\psi.$$

The first two terms are the same as in the scalar case, but we get an extra term which will give us the properties of spin 1/2 after quantization.

For example, the last term for $(J^0)^{ij}$ is given by

$$\begin{aligned}(J^0)^{ij} &= -i\bar{\psi}\gamma^0 S^{ij}\psi \\ &= \frac{1}{2}\epsilon^{ijk}\psi^\dagger \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \psi,\end{aligned}$$

where we have written the second line in the chiral repn, used the commutation relations of the γ matrices, and take $i, j, k \in \{1, 2, 3\}$.

There are also internal vector-like symmetries,

$$\psi \to e^{i\alpha}\psi \implies \delta\psi = i\alpha\psi.$$

Thus the conserved current here is

$$j^\mu_V = \bar{\psi}\gamma^\mu\psi,$$

with the conserved quantity

$$Q = \int d^3x \bar{\psi}\gamma^0\psi = \int d^3x \psi^\dagger\psi.$$

This has the interpretation of conserved electric charge and particle number.

Finally, we have axial symmetries. In the $m = 0$ limit, we can do $\psi \to e^{i\alpha\gamma^5}\psi$, which rotates LH/RH spinors in opposite directions. The conserved axial vector current is then

$$j^\mu_A = \bar{\psi}\gamma^\mu\gamma^5\psi.$$

Plane wave solutions We'd like to solve the Dirac equation,

$$(i\partial\!\!\!/ - m)\psi) = 0.$$

In particular, we will look for solutions of the form

$$\psi = u_p e^{-ip\cdot x}.$$

Substituting into the Dirac equation (using the chiral repn for γ^μ), we have

$$(p\!\!\!/ - mI)u_{\mathbf{p}} = \begin{pmatrix} -m & p_\mu\sigma^\mu \\ p_\mu\bar{\sigma}^\mu & -m \end{pmatrix} u_{\mathbf{p}} = 0. \tag{18.1}$$

Let us now claim that the solution is

$$u_{\mathbf{p}} = \begin{pmatrix} \sqrt{p \cdot \sigma} & \xi \\ \sqrt{p \cdot \bar{\sigma}} & \xi \end{pmatrix}$$

for any constant two-component spinor ξ, normalized such that $\xi^\dagger \xi = 1$.

Proof. Let us suppose that $u_{\mathbf{p}} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and substitute into Eqn. 18.1. Then we get

$$(p \cdot \sigma) u_2 = m u_1 \tag{18.2}$$
$$(p \cdot \bar{\sigma}) u_1 = m u_2. \tag{18.3}$$

Indeed, either of these implies the other since

$$\begin{aligned}(p \cdot \sigma)(p \cdot \bar{\sigma}) &= p_0^2 - p_i p_j \sigma^i \sigma^j \\ &= p_0^2 - p_i p_j \underbrace{\frac{1}{2}\{\sigma^i, \sigma^j\}}_{\delta^{ij}} \\ &= p_\mu p^\mu = m^2.\end{aligned}$$

So multiplying the first by $p \cdot \bar{\sigma}$ gives the second, for instance. Now we try the solution

$$u_1 = (p \cdot \sigma)\xi'$$

for some 2-spinor ξ' to find that

$$u_2 = \frac{1}{m}(\partial \cdot \bar{\sigma})(p \cdot \sigma)\xi' = m\xi'.$$

What this tells us is that any vector of the form

$$u_{\mathbf{p}} = A \begin{pmatrix} (p \cdot \sigma)\xi' \\ m\xi' \end{pmatrix}$$

is a solution to 18.1 with A a constant. To make this look more symmetric, we choose $A = 1/m$ and $\xi' = \sqrt{p \cdot \bar{\sigma}}\xi$,with ξ constant. Then

$$u_{\mathbf{p}} = \begin{pmatrix} \sqrt{p \cdot \sigma}\xi \\ \sqrt{p \cdot \bar{\sigma}}\xi \end{pmatrix}.$$

⊠

Example 18.4. Let's take a massive spinor in its rest frame, mass m and $\mathbf{p} = 0$. Then

$$u_{\mathbf{p}} = \sqrt{m}\begin{pmatrix} \xi \\ \xi \end{pmatrix}$$

for any ξ. Under spatial rotations, ξ transforms to

$$\xi \to e^{i\boldsymbol{\sigma} \cdot \boldsymbol{\phi}/2}\xi,$$

and after quantization, ξ will describe spin. For instance, $\xi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ will be a spin $\uparrow$ along the z-axis.

What happens if we boost the particle along the x^3 axis? We get

$$u_{\mathbf{p}} \begin{pmatrix} \sqrt{E - p^3}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \sqrt{E + p^3}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} \to_{E \to p^3, m \to 0} \sqrt{2E}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

and for spin down we have

$$u_{\mathbf{p}} \begin{pmatrix} \sqrt{E + p^3}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \sqrt{E - p^3}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} \to_{m \to 0} \sqrt{2E}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

We've been a little slick in rewriting the nonvanishing component of $p \cdot \sigma$, since the argument of the square root is technically a matrix. But we won't dwell on this too much except to say that it seems to be well-defined for our purposes.

Note that there are also negative frequency solutions to the Dirac equation. We simply take the ansatz of $\psi = v_{\mathbf{p}} e^{+ip\cdot x}$ and get some similar solutions $v_{\mathbf{p}} = \begin{pmatrix} \sqrt{p\cdot\sigma}\eta \\ -\sqrt{p\cdot\bar{\sigma}}\eta \end{pmatrix}$, with η a two-component spinor where $\eta^\dagger \eta = 1$.

Definition 18.5. The *helicity* operator projects angular momentum along the direction of motion:

$$h = \hat{\mathbf{p}} \cdot \mathbf{s} = \frac{1}{2}\hat{p}_i \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}.$$

The massless spin $\uparrow$ particle has $h = +1/2$, while the massless spin $\downarrow$ particle has $h = -1/2$, as one might expect.

Quantizing the Dirac field The Dirac field admits a quantization– if we throw around some creation and annihilation operators like we as field theorists are wont to do, we find that the field can be written as

$$\psi(\mathbf{x}) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left[b^s_{\mathbf{p}} u^s_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} + c^{s\dagger}_{\mathbf{p}} v^s_{\mathbf{p}} e^{-i\mathbf{p}\cdot\mathbf{x}} \right]$$

and $\psi^\dagger$ is similar,

$$\psi(\mathbf{x})^\dagger = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left[b^{s\dagger}_{\mathbf{p}} u^{s\dagger}_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} + c^{s}_{\mathbf{p}} v^{s\dagger}_{\mathbf{p}} e^{-i\mathbf{p}\cdot\mathbf{x}} \right].$$

Lecture 19.

Thursday, November 15, 2018

Today we'll continue our discussion of quantizing the Dirac field, which describes the behavior of spinors (e.g. spin 1/2 particles). For bosons, we had commutation relations. In fermionic quantization, we instead require anti-commutation relations between the creation and annihilation operators. That is, defining the anti-commutator as

$$\{A, B\} \equiv AB + BA,$$

we have the following anti-commutation relations.

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

We now claim these are equivalent to the following anti-commutation relations for the creation and annihilation operators:

$$\{b^r_{\mathbf{p}}, b^{s\dagger}_{\mathbf{q}}\} = \{c^r_{\mathbf{p}}, c^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3 \delta^{rs}\delta^3(\mathbf{p}-\mathbf{q}),$$

with all other anticommutators vanishing. (See Sheet 3, Q6 for the computation.)

The corresponding Hamiltonian is

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = i\psi^\dagger\dot{\psi} - i\bar{\psi}\gamma^0\partial_0\psi - i\bar{\psi}\gamma^i\partial_i\psi + m\bar{\psi}\psi.$$

One finds that the first two terms cancel, so we are left with the Hamiltonian

$$\mathcal{H} = \bar{\psi}(-i\gamma^i\partial_i + m)\psi.$$

We can now plug in the expansions for $\psi, \bar{\psi}$ which we wrote down last lecture and use the anti-commutation relations for operators (plus some results on products of spinors) to show that

$$\begin{aligned} u^{r\dagger}_{\mathbf{p}} u^s_{\mathbf{p}} = v^{r\dagger}_{\mathbf{p}} v^s_{\mathbf{p}} &= 2p_0\delta^{rs}, \\ u^{r\dagger}_{\mathbf{p}} v^s_{\mathbf{p}} = v^{r\dagger}_{\mathbf{p}} u^s_{\mathbf{p}} &= 0. \end{aligned}$$

We'll show this on Sheet 3, Q7. Thus the Hamiltonian can be rewritten (after normal ordering) as

$$H = \int \frac{d^3p}{(2\pi)^3} E_p \sum_{s=1}^{2} (b_{\mathbf{p}}^{s\dagger} b_{\mathbf{p}}^s + c_{\mathbf{p}}^{s\dagger} c_{\mathbf{p}}^s.$$

The issue with trying to use commutation relations is that they produce a minus sign in one of the terms in the Hamiltonian, meaning that one can reduce the energy of a state by creating a particle. In words, the vacuum state becomes unstable, and we get an explosion of particles as the energy decreases without bound. This isn't really physical so anti-commutation relations are the way to go.

The spinor field also leads to *Fermi-Dirac statistics*. Note that

$$b_{\mathbf{p}}^s \left|0\right\rangle = 0 = c_{\mathbf{p}}^s \left|0\right\rangle.$$

Although $b_{\mathbf{p}}^s, c_{\mathbf{p}}^r$ have anti-commutation relations, the Hamiltonian H has the usual commutation relations with them (check this):

$$[H, b_{\mathbf{p}}^{r\dagger}] = E_{\mathbf{p}} b_{\mathbf{p}}^{r\dagger}$$
$$[H, b_{\mathbf{p}}^{r}] = -E_{\mathbf{p}} b_{\mathbf{p}}^{r}.$$

Let's set up the state

$$\left|\mathbf{p}, r\right\rangle \equiv b_{\mathbf{p}}^{r\dagger} \left|0\right\rangle.$$

Then by the anti-commutation relations, the two-particle state obeys

$$\left|\mathbf{p}_1, r_1; \mathbf{p}_2, r_2\right\rangle = -\left|\mathbf{p}_2, r_2; \mathbf{p}_1, r_1\right\rangle,$$

which precisely means the state is antisymmetric under exchange of particles.

Let's now pass to the Heisenberg picture. We have $\psi(x)$ satisfying $\frac{\partial \psi}{\partial t} = i[H, \psi]$, which is solved by

$$\psi_\alpha(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} (b_{\mathbf{p}}^s u_{\mathbf{p}_\alpha}^s e^{-ip\cdot x} + c_{\mathbf{p}}^{s\dagger} v_{\mathbf{p}_\alpha}^s e^{+ip\cdot x}).$$

$\psi_\alpha^\dagger$ is similar, but with daggers everywhere:

$$\psi_\alpha^\dagger(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} (b_{\mathbf{p}}^{s\dagger} u_{\mathbf{p}_\alpha}^{s\dagger} e^{+ip\cdot x} + c_{\mathbf{p}}^{s} v_{\mathbf{p}_\alpha}^{s\dagger} e^{-ip\cdot x}).$$

In analogy with the Feynman propagator $\Delta(x-y) = [\phi(x), \phi(y)]$, let us now define

$$iS_{\alpha\beta}(x-y) = \{\psi_\alpha(x), \bar{\psi}_\beta(y)\}.$$

In what follows, we'll drop the indices α, β but should remember that S is really a 4×4 matrix since α and β index over the four spinor components. We substitute in our expressions for $\psi, \bar{\psi}$ from above to find

$$iS(x-y) = \sum_{r,s} \int \frac{d^3q d^3p}{(2\pi)^6 \sqrt{4E_p E_q}} \left(\{b_{\mathbf{p}}^s, b_{\mathbf{q}}^{r\dagger}\} u_{\mathbf{p}}^s \bar{u}_{\mathbf{q}}^r e^{-i(p\cdot x - q\cdot y)} + \{c_{\mathbf{p}}^{s\dagger}, c_{\mathbf{q}}^r\} v_{\mathbf{p}}^s \bar{v}_{\mathbf{q}}^r e^{i(p\cdot x - q\cdot y)} \right). \tag{19.1}$$

Using the anticommutation relations of b, c we have

$$iS(x-y) = \int \frac{d^3p}{(2\pi)^3 2E_p} \left(\sum_{s=1}^{2} u_{\mathbf{p}\alpha}^s \bar{u}_{\mathbf{p}\beta}^s e^{-ip\cdot(x-y)} + v_{\mathbf{p}\alpha}^s \bar{v}_{\mathbf{p}\beta}^s e^{+ip\cdot(x-y)} \right).$$

We see that (Sheet 3, Q5) these $u\bar{u}$ terms become $(\not{p} + m)_{\alpha\beta}$, so the overall expression becomes

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

where $\not{\partial}_x = \gamma^\mu \frac{\partial}{\partial x^\mu}$ and $D(x-y) = \int \frac{d^3p}{(2\pi)^3 2E_p} e^{-p\cdot(x-y)}$.

Some comments on how to interpret this.

(a) For spacelike separations $(x-y)^2 < 0, D(x-y) - D(y-x) = 0$. In bosonic theory, we made a big deal of this, since it ensured that the commutator was zero for spacelike separations. We interpreted this as saying that our theory was causal– the propagator from point x to y precisely cancels the contribution from propagating from y to x when x and y are spacelike separated. What can we say about the anticommutation relation in our spinor theory? We have

$$\{\psi_\alpha(x), \bar{\psi}_\beta(y)\} = 0 \quad \forall (x-y)^2 < 0.$$

However, it turns out that all observables are bilinear in $\psi, \bar{\psi}$ (cf. our table 1) and so the observables do commute at spacelike separations. The theory is still causal.

(b) Away from singularities (e.g. poles in the S-matrix), we have

$$(i\partial\!\!\!/_x - m)S(x-y) = 0.$$

The proof is by direct computation:

$$\begin{aligned}(i\partial\!\!\!/_x - m)S(x-y) &= (i\partial\!\!\!/_x - m)(i\partial\!\!\!/_x + m)[D(x-y) - D(y-x)] \\ &= -(\partial_x^2 + m^2)[D(x-y) - D(y-x)] \\ &= 0 \text{ using } p^2 = m^2 \text{ on-shell.}\end{aligned}$$

In going from the first to the second line, we have also cancelled the slashes (see Sheet 3 Q3): $\partial\!\!\!/\partial\!\!\!/ = \partial^2$.

These taken together allow us to write down the Feynman propagator for spinors. A similar calculation gives

$$\langle 0| \psi_\alpha(x)\bar{\psi}_\beta(y) |0\rangle = \int \frac{d^3p}{(2\pi)^3 2E_p}(p\!\!\!/ + m)_{\alpha\beta} e^{-ip\cdot(x-y)}.$$

We had almost the same thing in the bosonic case, but without the $p\!\!\!/ + m$. Similarly,

$$\langle 0| \bar{\psi}_\beta(x)\psi_\alpha(y) |0\rangle = \int \frac{d^3p}{(2\pi)^3 2E_p}(p\!\!\!/ - m)_{\beta\alpha} e^{-ip\cdot(x-y)}.$$

One should check this (as one of many, many exercises).

Let us now define

$$S_{F,\alpha\beta}(x-y) \equiv \langle 0| T\{\psi_\alpha(x)\bar{\psi}_\beta(y)\} |0\rangle = \begin{cases} \langle 0| \psi_\alpha(x)\bar{\psi}_\beta(y) |0\rangle & : x^0 > y^0 \\ -\langle 0| \bar{\psi}_\beta(y)\psi_\alpha(x) |0\rangle & : y^0 > x^0. \end{cases} \tag{19.2}$$

Note the minus sign is required for Lorentz invariance. When $(x-y)^2 < 0, \{\psi(x), \bar{\psi}(y)\} = 0$ and so T as defined is Lorentz invariant. The same is true for strings of fermionic operators inside the time ordering T: they anti-commute, so $[\phi_1, \psi] = 0, \{\psi_1, \bar{\psi}_2\} = 0, [\phi_1, \phi_2] = 0$ sum up the relations of bosonic and fermionic fields.

We also pick up the sign flip when computing normal-ordered products, which will affect Wick's theorem. Here,

$$: \psi_1\psi_2 := - : \psi_2\psi_1 :$$

We can write the contraction

$$S_F(x-y) \equiv \overbrace{\psi(x)\bar{\psi}(y)},$$

so the time-ordered version gives us

$$T[\psi(x)\bar{\psi}(y)] =: \psi(x)\bar{\psi}(y) : + \overbrace{\psi(x)\bar{\psi}(y)}.$$

Wick's theorem is then as before, but with extra minus signs, e.g.

$$: \psi_1^+\psi_2^- := -\psi_2^-\psi_1^+.$$

If we want to contract ψ_1 with $\bar{\psi}_3$ in the expression

$$: \psi_1\psi_2\bar{\psi}_3\psi_4 :$$

then we have to anti-commute $\bar{\psi}_3, \psi_2$ to get

$$: \overbrace{\psi_1\psi_2\bar{\psi}_3}\psi_4 := - : \overbrace{\psi_1\bar{\psi}_3}\psi_2\psi_4 := -\overbrace{\psi_1\bar{\psi}_3} : \psi_2\psi_4 :$$

The Feynman propagator for spinors is therefore

$$S_F(x-y) = i\int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{\not p + m}{p^2 - m^2 + i\epsilon},$$

which satisfies

$$(i\not\partial - m)S_F(x-y) = i\delta^4(x-y),$$

so S_F is a Green's function of the Dirac equation.

Lecture 20.

Saturday, November 17, 2018

Today we will discuss fermionic Yukawa theory. Before, we looked at Yukawa theory in a simplified scalar case. However, nucleons are really fermions. The interactions between fermions and a scalar particle are governed by the Yukawa interaction,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \underbrace{\frac{\mu^2}{2}\phi^2}_{\text{scalar mass}} + \bar\psi(i\not\partial - m)\psi - \underbrace{\lambda\phi\bar\psi\psi}_{\text{Yukawa interaction}}.$$

From the kinetic terms, we see that $[\phi] = 1$ and $[\psi] = [\bar\psi] = 3/2$, which is a bit unusual. We conclude that $[\lambda] = 0$ which means that this coupling is *marginal*.

Let us consider again the process of nucleon-nucleon scattering, $\psi\psi \to \psi\psi$. Now we must keep track of spin indices as well as momentum. Our initial state is a two-particle state

$$|i\rangle = \sqrt{2E_p}\sqrt{2E_q} b^{s\dagger}_{\mathbf{p}} b^{r\dagger}_{\mathbf{q}} |0\rangle$$

and our final state is similar,

$$|f\rangle = \sqrt{2E_{p'}}\sqrt{2E_{q'}} b^{s'\dagger}_{\mathbf{p}'} b'^{r\dagger}_{\mathbf{q}'} |0\rangle$$

As before, we will disregard the $O(\lambda^0)$ term where the particles do not interact, and there is no $O(\lambda)$ term as in the scalar case. Therefore the leading order interesting behavior is the $O(\lambda^2)$ term:

$$\langle f| (S-1) |i\rangle = \langle f| \frac{(-i\lambda)^2}{2!} \int d^4x_1 d^4x_2 T\left[\bar\psi(x_1)\psi(x_1)\phi(x_1)\bar\psi(x_2)\psi(x_2)\phi(x_2)\right] |i\rangle . \tag{20.1}$$

All fields are in the interaction picture as usual. We'll use Wick's theorem to compute the time ordering. The contribution to the scattering then comes from the contraction

$$: \bar\psi(x_1)\psi(x_1)\bar\psi(x_2)\psi(x_2) : \overbrace{\phi(x_2)\phi(x_1)}.$$

Thus the ψs will annihilate the $|i\rangle$ and the $\bar\psi$s will create $\langle f|$. In order to put the normal-ordered bit in the right order, we must anticommute $\bar\psi(x_2)$ past $\psi(x_1)$, picking up a sign flip for our trouble. We therefore get the interaction

$$\begin{aligned} I \equiv&: \bar\psi_\alpha(x_1)\psi_\alpha(x_1)\bar\psi_\beta(x_2)\psi_\beta(x_2) : b^{s\dagger}_{\mathbf{p}} b^{r\dagger}_{\mathbf{q}} |0\rangle \\ &= -\int \frac{d^3k_1 d^3k_2}{(2\pi)^6 2\sqrt{E_{k_1}E_{k_2}}} \left[\bar\psi_\alpha(x_1)U^m_{\mathbf{k}_1,\alpha}\right]\left[\bar\psi_\beta(x_2)U^n_{\mathbf{k}_2,\beta}\right] e^{-i(k_1x_1+k_2x_2)} b^m_{\mathbf{k}_1} b^n_{\mathbf{k}_2} b^{s\dagger}_{\mathbf{p}} b^{r\dagger}_{\mathbf{q}} |0\rangle . \end{aligned}$$

Here, the Us are the planar wave solutions from before, and the square bracket means that we contract over the four spinor indices. The other terms i the expansion go to zero because they have either $\langle 0| c^\dagger = 0$ or $c|0\rangle = 0$. Looking at the creation and annihilation operators, we can rewrite as

$$\begin{aligned} b^m_{\mathbf{k}_1} b^n_{\mathbf{k}_2} b^{s\dagger}_{\mathbf{p}} b^{r\dagger}_{\mathbf{q}} |0\rangle &= (b^m_{\mathbf{k}_1}\{b^n_{\mathbf{k}_2} b^{s\dagger}_{\mathbf{p}}\} b^{r\dagger}_{\mathbf{q}} - b^m_{\mathbf{k}_1} b^{s\dagger}_{\mathbf{p}}\{b^n_{\mathbf{k}_2} b^{r\dagger}_{\mathbf{q}}\}) |0\rangle \\ &= (\{b^m_{\mathbf{k}_1} b^{r\dagger}_{\mathbf{q}}\}\{b^n_{\mathbf{k}_2} b^{s\dagger}_{\mathbf{p}}\} - \{b^m_{\mathbf{k}_1} b^{s\dagger}_{\mathbf{p}}\}\{b^n_{\mathbf{k}_2} b^{r\dagger}_{\mathbf{q}}\}) |0\rangle , \end{aligned}$$

where we have used the fact that annihilation operators kill the vacuum state and anticommutators are just c-numbers. Thus this whole expression becomes delta functions,

$$(2\pi)^6[\delta^3(\mathbf{k}_2 - \mathbf{p})\delta^3(\mathbf{k}_1 - \mathbf{q})\delta^{ns}\delta^{mr} - \delta^3(\mathbf{k}_1 - \mathbf{p})\delta^3(\mathbf{k}_2 - \mathbf{q})\delta^{ms}\delta^{nr}] |0\rangle .$$

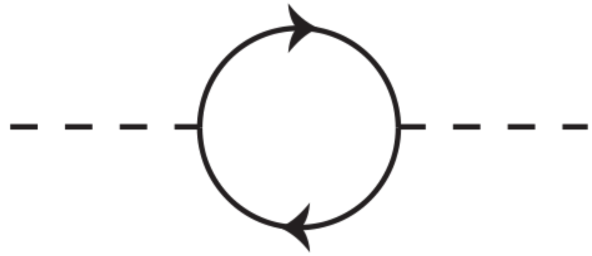

FIGURE 4. The closed fermion loop diagram. Image from Wikipedia.

Hence I simplifies somewhat to

$$I = -\frac{1}{2\sqrt{E_p E_q}} \left\{ [\bar{\psi}(x_1) U^r_{\mathbf{q}}][\bar{\psi}(x_2) U^s_{\mathbf{p}}] e^{-i(q\cdot x_1 + p\cdot x_2)} - [\bar{\psi}(x_1) U^s_{\mathbf{p}}][\bar{\psi}(x_2) U^r_{\mathbf{q}}] e^{-i(p\cdot x_1 + q\cdot x_2)} \right\} |0\rangle \,.$$

We still need to apply the final state to get

$$-\frac{2\sqrt{2E_{p'}E_{q'}}}{2\sqrt{E_p E_q}} \langle 0|\, b^{r'}_{\mathbf{q}'} b^{s'}_{\mathbf{p}'} \left\{ [\bar{\psi}(x_1) U^r_{\mathbf{q}}][\bar{\psi}(x_2) U^s_{\mathbf{p}}] e^{-i(q\cdot x_1 + p\cdot x_2)} - [\bar{\psi}(x_1) U^s_{\mathbf{p}}][\bar{\psi}(x_2) U^r_{\mathbf{q}}] e^{-i(p\cdot x_1 + q\cdot x_2)} \right\} |0\rangle \,.$$

From here, we pass to the integral representation, writing

$$-\frac{2\sqrt{2E_{p'}E_{q'}}}{2\sqrt{E_p E_q}} \langle 0| \left[\int \frac{d^3k_1 d^3k_2}{2\sqrt{E_{k_1}E_{k_2}}(2\pi)^6} b^{r'}_{\mathbf{q}'} b^{s'}_{\mathbf{p}} [\bar{U}^m_{k_1} U^r_{\mathbf{q}}] b^{m\dagger}_{\mathbf{k}_1} b^{n\dagger}_{\mathbf{k}_2} [\bar{U}_{k_2^n} u^s_{\mathbf{p}}] e^{i(k_1 x_1 + k_2 x_2) - i(q\cdot x_1 + p\cdot x_2)} - \dots \right]$$

where the ... indicates a similar term with p and q switched, r and s switched. But the bs pair up nicely to give us delta functions:

$$b^{r'}_{\mathbf{q}'} b^{s'}_{\mathbf{p}} b^{m\dagger}_{\mathbf{k}_1} b^{n\dagger}_{\mathbf{k}_2} = (2\pi^6) \langle 0| \left(\delta^{s'm} \delta^{r'n} \delta^3(\mathbf{p} - \mathbf{k}_1) \delta^3(\mathbf{q}' - \mathbf{k}_2) - \dots \right)$$

where ... is the same with m, n switched and $\mathbf{k}_1, \mathbf{k}_2$ switched. After applying delta functions (check the Tong notes for this) the expression cleans up in the same way that the Feynman diagrams would have shown us, but keeping track of spins. Writing $\langle f| (S-1) |i\rangle = iM(2\pi)^4 \delta^4(p+q-p'-q')$, we get

$$M = -(-i\lambda)^2 \left\{ \frac{[\bar{U}^{s'}_{\mathbf{p}'} U^r_q][\bar{U}^{r'}_{\mathbf{q}'} u^s_{\mathbf{p}}]}{(q'-p)^2 - \mu^2 + i\epsilon} - \frac{[\bar{u}^{r'}_{\mathbf{q}'} U^r_{\mathbf{q}}][\bar{U}^{s'}_{\mathbf{p}'} U^s_{\mathbf{p}}]}{(p'-p)^2 - \mu^2 + i\epsilon} \right\}. \tag{20.2}$$

That was a lot of work. What can Feynman tell us about this matrix element? We have some momentum space Feynman rules for fermion amplitudes. They are as follows.

- Dirac fermions preserve fermion number, so the arrows must not clash (e.g. no two arrows into a vertex).
- Incoming fermions get a momentum and a spinor index, $U^s_{\mathbf{p}}$. Outgoing fermions get $\bar{U}^s_{\mathbf{p}}$ instead.
- Incoming antifermions get $\bar{v}^s_{\mathbf{p}}$ and outgoing antifermions, $v^s_{\mathbf{p}}$.
- We impose 4-momentum conservation at each vertex.
- Each three-point vertex with a dashed line (the scalar) gets a factor of $(-i\lambda)$.
- Internal lines for ψs with spinor indices α going to β get propagators

$$\frac{i(\not{p} + m)_{\beta\alpha}}{p^2 - m^2 + i\epsilon},$$

where these α, β indices are contracted at vertices with other propagators or with external spinors. (The indices are contracted in the opposite direction to the fermion number arrows.)

For example, for the closed fermion loop in Fig. 4 we get an amplitude which goes as

$$-\bar{\psi}_\alpha(x)\,\psi_\alpha(x)\bar{\psi}_\beta(y)\,\psi_\beta(y) = -\psi_\beta(y)\bar{\psi}_\alpha(x)\,\psi_\alpha(x)\bar{\psi}_\beta(y)\,.$$

So we get an additional minus sign for a fermion loop, as well as the usual $d^4k/(2\pi)^4$ and fermion propagators.

Lecture 21.

Tuesday, November 20, 2018

Here's an important note about Wick's theorem on spinor fields. We can only contract spinor fields in the order $\overbrace{\psi(x)\bar{\psi}(y)}$. So far, we've looked at simple couplings like $\lambda\phi\bar{\psi}\psi$. What if we inserted a γ^5 to get

$$\mathcal{L}_{int} = -\lambda\phi\bar{\psi}_\alpha\gamma^5_{\alpha\beta}\psi_\beta?$$

We simply pick up a γ^5 in the interaction. Thus the three-point interaction is proportional to $-i\gamma^5_{\alpha\beta}\lambda$. Note that this interaction only preserves P symmetry if ϕ is also a pseudoscalar, i.e. if $P\phi(t,\mathbf{x}) = -\phi(t,-\mathbf{x})$.

We might then ask how to deal with spin indices when computing our physical observables like $|M|^2$ or σ. It turns out that in most experiments (e.g. at the LHC) the beams are prepared with random initial spin states, so when calculating observables it suffices to average over the spins. In scattering of two spin 1/2 particles, for instance, we would sum $\frac{1}{4}\sum_{s,r}$ where the 1/4 accounts for the four combinations $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$. The final states will also have some spin states, but we can take care of this by summing over final spins to get the cross-section variables.

Note also that for $\psi\psi \to \psi\psi$ scattering, the matrix element is $M \equiv A - B$ where A and B are the two different terms. Therefore the square of the matrix element is

$$\overline{|M|}^2 = \overline{|A|^2} + \overline{|B|^2} - \overline{A^\dagger B} - \overline{B^\dagger A}$$

where a bar indicates averaging/summing over spins. Here,

$$A = \frac{\lambda^2[\bar{u}^{s'}_{\mathbf{p}'}u^r_{\mathbf{q}}][\bar{u}^{r'}_{\mathbf{q}'}u^s_{\mathbf{p}}]}{u-\mu^2+i\epsilon}.$$

However, note that we can rewrite

$$[\bar{u}^{s'}_{\mathbf{p}'}u^r_{\mathbf{q}}] = [\bar{u}^r_{\mathbf{q}}u^{s'}_{\mathbf{p}'}]$$

since $(\gamma^0)^\dagger = \gamma^0$. Now

$$\overline{|A|^2} = \frac{\lambda^4}{4}\sum_{r,s,r',s'}\frac{\bar{u}^{s'}_{\mathbf{p}',\alpha}u^r_{\mathbf{q},\alpha}\bar{u}^r_{\mathbf{q},\beta}u^{s'}_{\mathbf{p}',\beta}}{(u-\mu^2)^2}\bar{u}^{r'}_{\mathbf{q}',\gamma}u^s_{\mathbf{p},\gamma}\bar{u}^s_{\mathbf{p},\delta}u^{r'}_{\mathbf{q}',\delta}.$$

After summing over r (cf. Example Sheet 3), we get a pair of traces,

$$\overline{|A|^2} = \frac{\lambda^4}{4}\frac{\mathrm{Tr}[(\not{\mathbf{p}}'+m)(\not{q}+m)}{(u-\mu^2)^2}\mathrm{Tr}(\not{\mathbf{q}}'-m)(\not{p}+m)].$$

In the high-energy (low-mass) limit as $\mu, m \to 0$, we get

$$\overline{|A|^2} = \frac{\lambda^4}{4u^2}\mathrm{Tr}[\not{\mathbf{p}}'\not{q}]\mathrm{Tr}[\not{\mathbf{q}}'\not{p}].$$

Similarly $\overline{|B|^2}$ comes out to

$$\overline{|B|^2} = \frac{\lambda^4}{4t^2}\mathrm{Tr}[\not{q}'\not{q}]\mathrm{Tr}[\not{p}'\not{p}].$$

We can also do the same for the cross-terms to find

$$\begin{aligned}\overline{A^\dagger B} &= \frac{\lambda^4}{4ut}\sum_{r,s,s'r'}\{\bar{u}^r_{\mathbf{q},\beta}u^{s'}_{\mathbf{p}',\beta}\bar{u}^s_{\mathbf{p},\alpha}u^{r'}_{\mathbf{q}',\alpha}\bar{u}^{r'}_{\mathbf{q}',\gamma}u^r_{\mathbf{q},\gamma}\bar{u}^{s'}_{\mathbf{p}',\delta}u^s_{\mathbf{p},\delta}\}\\ &= \frac{\lambda^4}{4ut}\mathrm{Tr}(\not{p}\not{p}'\not{q}\not{q}').\end{aligned}$$

Can we do this without going through the Wick way? Let us write down some Feynman rules for $\overline{|M|^2}$ with fermions.

- $\mathbb{C}$ conjugation switches the initial and final momenta.

- Fermion lines with identical momenta are joined on the LHS. A closed fermion line is given by a trace over γ matrices (after spin sum/average), with any γ matrices at vertices placed in the correct position in the grace. Fermion lines are followed backwards (against the arrows).

Applying the Feynman rules, we can read off the traces:

$$\overline{|M|^2} = \frac{\lambda^4}{4}\left\{\frac{\mathrm{Tr}(\not{q}\not{q}')\mathrm{Tr}(\not{p}\not{p}')}{t^2} + \frac{\mathrm{Tr}(\not{q}'\not{p})\mathrm{Tr}(\not{p}'\not{q})}{u^2} - \frac{2\mathrm{Re}\mathrm{Tr}\not{p}\not{p}'\not{q}\not{q}'}{ut}\right\}$$

We can rewrite these traces as dot products, and moreover we know some good properties of the Mandelstam variables:

$$\begin{aligned} s &= (p+q)^2 = (p'+q')^2 \implies p\cdot q = p'\cdot q' = s/2 \\ t &= (p-p')^2 = (q-q')^2 \implies p\cdot p' = q\cdot q' = -t/2 \\ u &= (p-q')^2 = (q-p')^2 \implies p\cdot q' = p'\cdot q = -u/2. \end{aligned}$$

In terms of dot products, the matrix element is

$$\frac{\lambda}{4}\left[\frac{4(q\cdot q')4p\cdot p'}{t^2} + \frac{4(q'\cdot p)4(p'\cdot q)}{u^2} - \frac{8}{ut}(p\cdot q'p'\cdot q + p\cdot p'q\cdot q' - p\cdot qp'\cdot q')\right].$$

Thus the matrix element reduces to

$$\overline{|M|^2} = \lambda^4\left\{1 + 1 - \frac{u^2+t^2-s^2}{2ut}\right\}.$$

In terms of the differential cross-section, we now have

$$\frac{d\sigma}{dt} = \frac{\overline{|M|^2}}{16\pi\lambda(s, m_1^2, m_2^2)}.$$

But in this limit $m_1 = m_2 = 0$ and $\lambda(s,0,0) = s^2$. $t = 2|\mathbf{p}|\mathbf{p}'|(\cos\theta - 1)$ in the center-of-mass frame, and $|\mathbf{p}| = |\mathbf{p}'| = \sqrt{s}/2$. Therefore

$$\frac{dt}{d\cos\theta} = 2|\mathbf{p}||\mathbf{p}'| = s/2$$

and we find that

$$d\Omega = d\cos\theta d\phi \implies \frac{d\sigma}{d\Omega} = \frac{s}{4\pi}\frac{d\sigma}{dt} = \frac{\overline{|M|^2}}{64\pi^2 s} = \frac{3\lambda^4}{64\pi^2 s}.$$

If we now integrate the final state over the hemisphere of solid angle (since the particles are identical), we find that the total cross-section is

$$\sigma = \frac{3\lambda^4}{32\pi s}.$$

Note that $[\lambda] = 0$ and $[s] = 2$, so indeed this quantity has a mass dimension of -2, i.e. has dimensions of area.

Lecture 22.

Thursday, November 22, 2018

As a quick reminder, there is no lecture on Saturday! The lecture has been rescheduled to Monday at 1 PM in MR2. Also, a correction to Examples Sheet 3, Q8: s^2 in the numerator should be $(s-4m^2)^2$.

Quantum electrodynamics (QED) It's time now to quantize the free electromagnetic field A_μ. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{22.1}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ is the *field strength tensor*. If we compute the equations of motion by the usual Euler-Lagrange procedure, we find that

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}\right) = 0 = \partial_\mu F^{\mu\nu}.$$

From its definition in terms of A_μ, we see that $F_{\mu\nu}$ therefore satisfies the *Bianchi identity*, i.e.

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0. \tag{22.2}$$

From this simple-looking Lagrangian, we can recover all of Maxwell's equations. We'll need to be a little careful about signs in our definition of 3-vectors, so let us write the potential as $A^\mu = (\phi, \mathbf{A})$ such that $\mathbf{A} = (A^1, A^2, A^3)$. Then the electric field is

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \dot{\mathbf{A}}$$

where

$$\boldsymbol{\nabla} \equiv \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right) = \partial_i.$$

We also define the magnetic field as

$$\mathbf{B} = \boldsymbol{\nabla} \wedge \mathbf{A},$$

where the wedge product is really just telling us to take the curl of $\mathbf{A}$ like we learned in freshman electrodynamics. Thus

$$\mathbf{E} = (F_{01}, F_{02}, F_{03}) = (-F^{01}, -F^{02}, -F^{03}).$$

Looking at the definition of $\mathbf{B}$, we see that if $\mathbf{B} = (B_1, B_2, B_3)$, then for instance

$$B_3 = \partial_1 A^2 - \partial_2 A^1 = -\partial_1 A_2 + \partial_2 A_1 = -F_{12}.$$

All in all, the field strength tensor allows us to recover all the elements of the electric and magnetic fields,

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}.$$

Here, μ indexes over rows and ν indexes over columns. The Bianchi identity then reads

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \text{ and } \dot{\mathbf{B}} = -\boldsymbol{\nabla} \wedge \mathbf{E},$$

while the equations of motion give

$$\boldsymbol{\nabla} \cdot \mathbf{E} = 0 \text{ and } \dot{\mathbf{E}} = \boldsymbol{\nabla} \wedge \mathbf{B}.$$

These are precisely Maxwell's equations in vacuum.

Now let's quantize the field. Our Lagrangian has no mass term, which is as we expect since the photon is a massive vector field (spin 1) A_μ. However, there are four components of A_μ, so it seems that we have 4 real degrees of freedom ($\mu = 0, 1, 2, 3$) even though the photon only has two polarization states! How do we resolve this? We make the following observations.

- A_0 is not dynamical since $F_{\mu\nu}$ is antisymmetric– there's no kinetic term for it in the Lagrangian $\mathcal{L}$. In particular, given $A_i(\mathbf{x}, t_0)$ and $\dot{A}_i(\mathbf{x}, t_0)$, A_0 is fully determined. For if $\boldsymbol{\nabla} \cdot \mathbf{E} = 0$ then

  $$\nabla^2 A_0 + \boldsymbol{\nabla} \cdot \dot{\mathbf{A}} = 0,$$

  with solution[45]

  $$A_0(\mathbf{x}, t_0) = \int \frac{d^3x' \boldsymbol{\nabla} \cdot \dot{\mathbf{A}}(\mathbf{x}', t_0)}{4\pi|\mathbf{x} - \mathbf{x}'|}. \tag{22.3}$$

  So A_0 isn't independent– there are really only three real degrees of freedom.
- There is a large symmetry group of transformations of the form

  $$A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x)$$

  with λ such that $\lim_{|\mathbf{x}| \to \infty} \lambda(x) = 0$. Under such a transformation, we find that

  $$F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_n u\lambda) - \partial_\nu(A_\mu + \partial_\mu \lambda) = \partial_m u A_\nu - \partial_n u A_\mu$$

  since partial derivatives commute, so $F_{\mu\nu}$ and therefore $\mathcal{L}$ is invariant under these transformations. Equivalently, we may write the equations of motion as $\eta_{\mu\nu}\partial_\rho F^{\rho\nu} = 0$ or

  $$(\eta_{\mu\nu}\partial_\rho\partial^\rho - \partial_\mu\partial_\nu)A^\nu = 0.$$

[45] This is just the Green's function for ∇^2.

Note that the operator in parentheses is not invertible– it annihilates functions of the form $\partial^\nu \lambda(x)$. Therefore there is a redundancy in our description of the vector field A^ν, which we call a *gauge symmetry*. The existence of a gauge symmetry is equivalent to the statement that there is no unique solution for A_μ; it is only determined up to adding $\partial_\mu \lambda(x)$.

We say that the configuration space for A_μ is then *foliated*[46] by *gauge orbits*[47]. We can draw these as lines in configuration space, such that all states on a given line represent the same physical state. To actually compute things, we usually take a representative from a gauge orbit ("fix the gauge"), and in general we should choose a gauge that any A_μ can be put into. Here are some examples.

(a) Loren(t)z[48] gauge:

$$\partial_\mu A^\mu = 0.$$

This is a suitable gauge, as one can always put A_μ into this form. The proof is straightforward– if we have an A_μ such that $\partial_\mu A^\mu = f(x)$, we simply define $\tilde{A}_\mu = A_\mu + \partial_\mu \lambda(x)$ where $\partial_m u \partial^\mu \lambda(x) = \partial^2 \lambda = -f(x)$. We can solve this by our usual Green's function tricks for ∂^2,[49] and then $\partial_\mu \tilde{A}^\mu = 0$. However, this gauge still does not completely specify A_μ (remember what I said about the orbits being hypersurfaces), as we can always add to our new A_μ any $\partial_\mu \tilde{\lambda}$ such that $\partial^2 \tilde{\lambda} = 0$, of which there are infinitely many choices for $\tilde{\lambda}$. The advantage of this gauge is that it is Lorentz invariant, and often convenient when we want to write propagators and other quantities in a manifestly covariant way.

(b) Coulomb gauge/radiation gauge:

$$\boldsymbol{\nabla} \cdot \mathbf{A} = 0.$$

We can always put A^μ into this form by similar arguments to the Lorenz gauge, but with the regular three-dimensional Laplacian rather than the full ∂^2 operator. Referring back to Eqn. 22.3, we see that this gauge sets $A_0 = 0$ in vacuum. The advantage of this gauge is that it makes manifest the two physical degrees of freedom, i.e. the two polarization states of the photon. However, we lose Lorentz invariance.

Having resolved the question of the extra degrees of freedom in A_μ, let us now proceed to write the Hamiltonian for the EM field. What are the conjugate momenta? The first is

$$\pi_0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = 0,$$

since there is no $\dot{A}_0 = \dot{\phi}$ component in any of the fields, while the others are

$$\pi_i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = -\dot{A}_i + \partial_i A^0 = F^{i0} = E^i.$$

After a little algebra, we see that the Hamiltonian takes the form

$$H = \int d^3x (\pi_i \dot{A}_i - \mathcal{L}) = \int d^3x \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) - A_0(\boldsymbol{\nabla} \cdot \mathbf{E}).$$

In this last term, A_0 (being non-dynamical) just acts like a Lagrange multiplier and sets $\boldsymbol{\nabla} \cdot \mathbf{E} = 0$.

If we work in the Lorenz gauge, $\partial_\mu A^\mu = 0$, the equations of motion then become

$$\partial_\mu \partial^\mu \mathbf{A} = 0.$$

Let us note that in this gauge, we can write a new Lagrangian with an extra term,

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

---

[46] separated into hypersurfaces

[47] everywhere the gauge freedom takes you from a given starting point, i.e. for a particular choice of A_μ, the corresponding orbit O is $O = \{A_\mu + \partial_\mu \lambda(x) : \lim_{|\mathbf{x}| \to \infty} \lambda(x) = 0\}$

[48] This gauge condition is actually named for Ludvig Lorenz, and not Hendrik Lorentz of Lorentz invariance. Somewhat confusingly, this gauge condition is *Lorentz* invariant, hence the confusion. See `https://en.wikipedia.org/wiki/Lorenz_gauge_condition`

[49] The operator ∂^2 is like the Laplacian ∇^2, but it has a time derivative in it too. In Minkowski space it takes the simple form $-\partial_t^2 + \partial_x^2 + \partial_y^2 + \partial_z^2$ and is known as the d'Alembert operator or the d'Alembertian, denoted by $\Box$. Of course, we might want to do QFT in curved spacetime, so then ∂^2 takes some other more complicated form.

Thus

$$\partial_\mu F^{\mu\nu} + \partial^\nu(\partial_\mu A^\mu) = 0 \iff \partial_\mu \partial^\mu A^\nu = 0.$$

It will be convenient to work with this new Lagrangian and only impose $\partial_\mu A^\mu = 0$ later, at the operator level. Thus we can write a general Lagrangian as

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

where $\alpha = 1$ is known as *Feynman gauge* and $\alpha = 0$[50] is called *Landau gauge*.

Lecture 23.

Tuesday, November 27, 2018

The make-up lecture that was supposed to take place on Monday was cancelled, so we'll have an extra lecture instead on Thursday at the regular time.

Last time, we started discussing Lorenz gauge, where

$$\partial_\mu A^\mu = 0.$$

We wrote down a new Lagrangian

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

noting that this new theory does not have the gauge symmetry of the original, and additionally, A_0, A_i are now all dynamical. Thus

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = -\frac{\partial_\mu A^\mu}{\alpha}$$

(which is now nonzero) and

$$\pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = -\dot{A}^i + \partial^i A^0$$

as before.

If we now apply the commutation relations

$$[A_\mu(\mathbf{x}), A_\nu(\mathbf{y})] = [\pi_\mu(\mathbf{x}), \pi_\nu(\mathbf{y})] = 0$$

and

$$[A_\mu(\mathbf{x}), \pi_\nu(\mathbf{y})] = -i\delta^3(\mathbf{x}-\mathbf{y})\eta_{\mu\nu},$$

we can expand out the field in terms of polarization vectors $\epsilon_\mu^{(\lambda)}$ and creation and annihilation operators $a_{\mathbf{p}}^\lambda, a_{\mathbf{p}}^{\lambda\dagger}$. A_μ takes the form

$$A_\mu(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2|\mathbf{p}|}}\sum_{\lambda=0}^{3}\left(\epsilon_\mu^{(\lambda)}(\mathbf{p})a_{\mathbf{p}}^\lambda e^{i\mathbf{p}\cdot\mathbf{x}} + \epsilon_\mu^{(\lambda)*}(\mathbf{p})a_{\mathbf{p}}^{\lambda\dagger}e^{-i\mathbf{p}\cdot\mathbf{x}}\right).$$

Note that the polarization vectors have a Lorentz index μ since A_μ must transform like a vector field. WLOG we may pick $\epsilon_\mu^{(0)}$ to be timelike and $\epsilon_\mu^{(i)}$ to be spacelike, with normalization

$$\epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'}.$$

Moreover we shall choose $\epsilon_\mu^{(1)}, \epsilon_\mu^{(2)}$ to be transverse polarizations, i.e. $\epsilon^{(1)} \cdot p = 0$ and $\epsilon^{(2)} \cdot p = 0$. We now choose $\epsilon^{(3)}$ to be the longitudinal polarization. For a photon traveling in the x^3 direction, the momentum is

[50] This looks bad, I know. What we'll see is that in the end, something sensible happens when we try to work out the photon propagator. Really, we should think of taking the $\alpha \to 0$ limit as forcing $\partial_m u A^\mu \to 0$. In the path integral context, as $\alpha \to 0$ there is an increasingly high energy cost (i.e. an exponential damping in the factor e^{iS}) to having any $\partial_\mu A^\mu$ coupling. In any case, we should first find the photon propagator in terms of α, and then set α to zero (or one, if we're working in Feynman gauge).

simply $p^\mu = |\mathbf{p}|(1,0,0,1)$, so the polarization vectors take the simple form

$$\epsilon^{(0)} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \epsilon^{(1)} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \epsilon^{(2)} = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \epsilon^{(3)} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

and we can Lorentz boost or rotate to get photons traveling in other directions.

We also write the conjugate momentum in terms of creating and annihilation operators.

$$\pi^\mu(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\sqrt{\frac{|\mathbf{p}|}{2}}\sum_{\lambda=0}^{3}\left((\epsilon^{(\lambda)}(\mathbf{p}))^\mu a^\lambda_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} - (\epsilon^{(\lambda)}(\mathbf{p}))^{*\mu} a^{\lambda\dagger}_\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}}\right).$$

Now we get the following commutation relations:

$$[a^\lambda_\mathbf{p}, a^{\lambda'}_\mathbf{q}] = [a^{\lambda\dagger}_\mathbf{p}, a^{\lambda'\dagger}_\mathbf{q}] = 0$$

and

$$[a^\lambda_\mathbf{p}, a^{\lambda'\dagger}_\mathbf{q}] = -\eta^{\lambda\lambda'}(2\pi)^3\delta^3(\mathbf{p}-\mathbf{q}).$$

The minus sign in front of the delta function may seem okay for for $\lambda = 1,2,3$ but a bit strange for the timelike polarization. Somehow, timelike γs are different.

We define a ground state $|0\rangle$ by

$$a^\lambda_\mathbf{p}|0\rangle = 0$$

as usual, and then the various momentum states are

$$|\mathbf{p},\lambda\rangle = a^{\lambda\dagger}_\mathbf{p}|0\rangle.$$

Now, this is totally fine for $\lambda = 1,2,3$ but if we take $\lambda = 0$, we get

$$\langle p,\lambda=0|\, q,\lambda=0\rangle = \langle 0|\, a^0_\mathbf{p} a^{0\dagger}_\mathbf{q}\,|0\rangle = -(2\pi)^3\delta^3(\mathbf{p}-q),$$

which appears to be a state of negative norm. Now, a Hilbert space with negative norm states is usually problematic– in particular, the probabilistic interpretation of QM goes out the window. In our case, the constraint equation comes to the rescue.

Let us switch to the Heisenberg picture and see what becomes of this polarization.

(a) Initially, we might think we could just impose the gauge condition. But note that $\partial_\mu A^\mu = 0$ doesn't work since $\pi^0 = -\partial_\mu A^\mu$, so the commutation relationships cannot be obeyed (i.e. if π^0 vanishes then all its commutators are zero).

(b) We could impose a condition on Hilbert space, such that for physical states $|\psi\rangle$ we have $\partial_\mu A^\mu|\psi\rangle = 0$. But this is too strong. For suppose we split up A_μ into to parts,

$$A^+_\mu(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2|\mathbf{p}|}}\sum_{\lambda=0}^{3}\epsilon^{(\lambda)}_\mu(\mathbf{p}) a^\lambda_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}}$$

and

$$A^-_\mu = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2|\mathbf{p}|}}\sum_{\lambda=0}^{3}\epsilon^{(\lambda)*}_\mu(\mathbf{p}) a^{\lambda\dagger}_\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}}.$$

Then we see that $\partial_\mu A^{\mu+}|0\rangle = 0$, which is okay, but $\partial_\mu A^{\mu-}|0\rangle \neq 0$, which tells us that $|0\rangle$ is not physical. Therefore this doesn't work.

(c) Finally, let us say that physical states $|\psi\rangle$ are defined by

$$\partial_\mu A^{+\mu}(\mathbf{x})|\psi\rangle = 0 \tag{23.1}$$

so that

$$\langle\psi'|\,\partial_\mu A^\mu\,|\psi\rangle = 0$$

for all physical states $|\psi\rangle, |\psi'\rangle$. Eqn. 23.1 is known as the *Gupta-Bleuler condition*.

By the linearity of 23.1, we see that the physical states $\{|\psi\rangle\}$ span a Hilbert space. Moreover, we can decompose a generic state into its transverse components $|\psi_T\rangle$ and its timelike (and/or longitudinal) components $|\phi\rangle$:

$$|\psi\rangle = |\psi_T\rangle |\phi\rangle .$$

Then

$$\partial_\mu A^{+\mu} |\psi\rangle = 0 \iff (a^3_{\mathbf{k}} - a^0_{\mathbf{k}}) |\phi\rangle = 0$$

Check this! This means that physical states contain timelike/longitudinal pairs only. That is,

$$|\phi\rangle = \sum_{n=0}^{\infty} C_n |\phi_n\rangle$$

where n indexes over n timelike/longitudinal pairs in general.

Therefore

$$\langle \phi_m | \phi_n \rangle = \delta_{m0}\delta_{n0},$$

so states with transverse-longitudinal (TL) pairs have zero norm. These zero norm states are treated as an equivalence class– two states which differ only in the TL pairs are treated as physically equivalent.

This only makes sense if observables don't depend on $|\phi_n\rangle$, e.g. if our Hamiltonian is

$$H = \int \frac{d^3p}{(2\pi)^3} |\mathbf{p}| \left(\sum_{i=1}^{3} a^{i\dagger}_{\mathbf{p}} a^i_{\mathbf{p}} - a^{0\dagger}_{\mathbf{p}} a^0_{\mathbf{p}} \right),$$

then since

$$(a^3_{\mathbf{k}} - a^0_{\mathbf{k}}) |\psi\rangle = 0 \implies \langle\psi| a^{3\dagger}_{\mathbf{p}} a^3_{\mathbf{p}} |\psi\rangle = \langle\psi| a^{0\dagger}_{\mathbf{p}} a^0_{\mathbf{p}} |\psi\rangle .$$

That is, as long as timelike and longitudinal pieces cancel in the Hamiltonian H, we only get physical contributions from transverse states. In general this cancellation works for any gauge-invariant operator evaluated on physical states.

Now it's time for us to write down the photon propagator! It is

$$\langle 0| T[A_\mu(\mathbf{x}) A_\nu(\mathbf{y})] |0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i}{p^2 + i\epsilon} \left[\eta_{\mu\nu} + (\alpha - 1)\frac{p_\mu p_\nu}{p^2} \right] e^{-ip\cdot(x-y)}$$

for a general gauge α. Note that in Feynman gauge ($\alpha = 1$), the propagator takes a particularly simple form– we just get the $\frac{-i\eta_{\mu\nu}}{p^2+i\epsilon}$ term. One can check that if we do any physical calculation in full generality leaving the α in, there are no αs in the final result.

Can we introduce interactions and sources into our theory? Sure we can. Let's first write down the Maxwell Lagrangian with a source,

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu A_\mu.$$

The equations of motion are $\partial_\mu F^{\mu\nu} = j^\nu$, and we see that this implies

$$\partial_\nu j^\nu = \partial_\nu \partial_\mu F^{\mu\nu} = 0,$$

so j^ν is a conserved current. Now the Dirac Lagrangian tells us that our theory of spin 1/2 fermions,

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

has an internal symmetry $\psi \to e^{-i\alpha}\psi, \bar{\psi} \to e^{i\alpha}\bar{\psi}$ with $\alpha \in \mathbb{R}$. This yields a current $j^\mu = \bar{\psi}\gamma^\mu\psi$. So we take this current from the Dirac Lagrangian and put it straight into the Maxwell Lagrangian to couple our photon to fermions.

Doing so yields the Lagrangian for quantum electrodynamics,

$$L_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\partial\!\!\!/ - m)\psi - e\bar{\psi}_\alpha \gamma^\mu_{\alpha\beta} A_\mu \psi_\beta.$$

Here, α, β are spinor indices and μ is a Lorentz index. e is a coupling constant determining the strength of the coupling of the photon to our fermion (e.g. an electron). We therefore have the kinetic terms describing a massless spin one particle in $F_{\mu\nu}$, the Dirac kinetic terms for a massive spin 1/2 particle in $\bar{\psi}(i\partial\!\!\!/ - m)\psi$, and a coupling term which tells us that (as we are well aware) photons and electrons can interact.

Lecture 24.

Thursday, November 29, 2018

We previously wrote down the coupling of the electromagnetic force to fermions, and said that the theory of quantum electrodynamics (QED) is therefore given by the Lagrangian

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{\partial} - m)\psi - e\bar{\psi}\gamma^{\mu}A_{\mu}\psi \tag{24.1}$$

where we have suppressed spinor indices. Note that gauge invariance in pure electromagnetism allowed us to get rid of the two extra degrees of freedom in the photon polarizations, leaving us with the two physical polarization states we expect from a massless spin 1 particle. Is the same true now that we have a coupling to fermions in $\mathcal{L}_{QED}$? Let us rewrite $\mathcal{L}_{QED}$ suggestively as

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi, \tag{24.2}$$

where D_μ is the *covariant* derivative given by

$$D_\mu\psi \equiv (\partial_\mu + ieA_\mu)\psi. \tag{24.3}$$

This sort of construction should look really familiar from the last few *Symmetries* lectures– we can reasonably hope that our new covariant derivative D_μ will live up to its name and transform correctly under gauge transformations.

In fact, it turns out that $\mathcal{L}_{QED}$ is invariant under gauge transformations, but both the gauge field A_μ *and* the spinor field ψ have to transform:

$$A_\mu(x) \to A_\mu(x) + \partial_\mu\lambda(x) \tag{24.4}$$

$$\psi(x) \to e^{-ie\lambda(x)}\psi(x) \tag{24.5}$$

$$\bar{\psi}(x) \to e^{+ie\lambda(x)}\bar{\psi}(x). \tag{24.6}$$

Note that these are local symmetries, i.e. $\lambda(x)$ depends explicitly on the spacetime point x! This is different from the global symmetry, where the field is transformed everywhere in the same way (e.g. by a factor $e^{ie\tilde{\lambda}}$, with $\tilde{\lambda}$ a constant). Let us now show that the covariant derivative transforms like the spinor field under gauge transformations.

Proof. By direct computation, the covariant derivative transforms as follows:

$$\begin{aligned} D_\mu\psi = (\partial_\mu + ieA_\mu)\psi &\to (\partial_\mu + ie(A_\mu + \partial_\mu\lambda(x))(e^{-ie\lambda(x)}\psi) \\ &= (-ie\partial_\mu\lambda e^{-ie\lambda}\psi + e^{-ie\lambda}\partial_\mu\psi) + (ieA_\mu e^{-ie\lambda}\psi + ie\partial_\mu\lambda e^{-ie\lambda}\psi) \\ &= e^{-ie\lambda}\partial_\mu\psi + ieA_\mu e^{-ie\lambda}\psi \\ &= e^{-ie\lambda}(\partial_\mu ieA_\mu)\psi = e^{-ie\lambda}D_\mu\psi. \end{aligned}$$

Therefore the covariant derivative $D_\mu\psi$ transforms like the spinor field ψ. Moreover $\not{D}$ is the same thing up to contraction with the gamma matrices γ^μ, and the gamma matrices are independent of the gauge (they are just some representation of the Clifford algebra), so $\not{D}$ also transforms like ψ. ⊠

We already checked that the Maxwell term was invariant since it only involves A_μ, so now we see that

$$\bar{\psi}(i(\not{D} - m)\psi \to (e^{ie\lambda}\bar{\psi})(e^{-ie\lambda}(i\not{D} - m)\psi) = \bar{\psi}(i(\not{D} - m)\psi$$

is also invariant under gauge transformations and therefore the entire QED Lagrangian is invariant, as we claimed.

From the QED Lagrangian, we see that the coupling constant e has the interpretation of electric charge since the equations of motion are

$$\partial_\mu F^{\mu\nu} = ej^\nu.$$

In pure electromagnetism, j^0 was just the electric charge density, but as a quantum operator we have instead

$$\begin{aligned} Q &= -e \int d^3x \bar{\psi}\gamma^0\psi \\ &= -e \int \frac{d^3p}{(2\pi)^3}(b_{\mathbf{p}}^{s\dagger}b_{\mathbf{p}}^s - c_{\mathbf{p}}^{s\dagger}c_{\mathbf{p}}^s) \\ &= -e(\text{\# of particles} - \text{\# of anti-particles}). \end{aligned}$$

Let us also note that while there is a single factor of e in our Lagrangian, actual cross-sections depend on the squares of matrix elements and so we commonly define

$$\alpha \equiv \frac{e^2}{4\pi} \tag{24.7}$$

to be a factor we call the *fine-structure constant*, and it has a numerical value measured to be about $1/137$.[51]

We can now discuss a similar problem, a theory with a gauge field A_μ coupling to a complex scalar ϕ. For a real scalar field, there is no suitable current to couple to, but for the complex scalar, introducing a coupling turns out to be doable. The appropriate covariant derivative is

$$D_\mu\phi \equiv \partial_\mu\phi - ieqA_\mu\phi, \tag{24.8}$$

where q is the charge of the scalar ϕ in units of e. For instance, the sup squark (supersymmetric partner of the up quark) has $q = +2/3$. Here, if the scalar field ϕ transforms as

$$\phi(x) \to e^{ieq\lambda(x)}\phi(x), \tag{24.9}$$

then it follows that

$$D_\mu\phi = \partial_\mu\phi - ieqA_\mu\phi \to \partial_\mu(e^{ieq\lambda}\phi) - ieqA_\mu(e^{ieq\lambda}\phi) = e^{ieq\lambda}D_\mu\phi \tag{24.10}$$

by a quick application of the Leibniz rule. Therefore the Lagrangian

$$\mathcal{L} = \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)(D_\mu\phi)^\dagger \tag{24.11}$$

is gauge invariant (the dagger flips the sign on $e^{ieq\lambda}$). If we now look at the interacting part of this Lagrangian and expand out terms a bit, we have

$$\mathcal{L}_{int} = ieq(\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi)^\dagger\phi)A_\mu + e^2q^2A_\mu A^\mu\phi^\dagger\phi. \tag{24.12}$$

This sort of Lagrangian is a good model for photons interacting with charged pions at low energies, $E \lesssim 100\,\text{MeV}$. At these energies, the pion "looks fundamental" to the photon, which is only sensitive to length scales on the order of its de Broglie wavelength. In reality, the pion is made up of a quark and anti-quark (e.g. $\pi^+ = u + \bar{d}$), and high-energy photons can "see" the component quarks with their fractional charges.

The Lagrangian now has a conserved current

$$j_\mu = ieq[(D_\mu\phi)^\dagger\phi - \phi^\dagger D_\mu\phi]],$$

which is gauge invariant.

In general this process is known as minimal coupling– in order to introduce a coupling between a $U(1)$ gauge field and any number of fields ϕ^a (which can be fermionic or bosonic), we consider how the fields transform under the gauge transformation and promote the partial derivatives in the kinetic terms to covariant derivatives so that

$$\partial_\mu\phi^a \to D_\mu\phi^a \equiv \partial_m u\phi^a - ieq_{(a)}\lambda\phi^a,$$

[51] An interesting aside from dimensional analysis. Recall that $[\psi] = 3/2$. Looking at the Maxwell term, we see that terms like $(\partial_\mu A_\nu)^2$ must have mass dimension 4, so the gauge field A_μ has mass dimension $[A_\nu] = 1$ like the scalar field in the scalar Yukawa coupling. But then if we look at the QED coupling term $-e\bar{\psi}\gamma^\mu A_\mu\psi$, we see that the coupling constant e must have mass dimension zero. (Of course, the gamma matrices are just collections of numbers so they do not contribute to the overall dimension). But this means that the fine structure constant α is itself a dimensionless number.

Now, a question– where did this value come from? It's dimensionless, so it is independent of our unit system, but it doesn't appear to be an integer or any mathematically significant constant like π or e. In the words of Richard Feynman, "It has been a mystery ever since it was discovered more than fifty years ago, and all good theoretical physicists put this number up on their wall and worry about it."

where λ comes from the gauge transformation $A_\mu \to A_\mu + \partial_\mu \lambda$.

We may also consider Feynman diagrams with some external photons in them (picture later).These diagrams end up giving ups polarization vectors in the final scattering amplitudes, $\epsilon_\mu^{(\lambda)}(k)$ for ingoing photons and $\epsilon_\mu^{(\lambda)*}(k)$ for outgoing photons. Typically we don't bother to resolve the polarization states (e.g. in the detectors of colliders), so we simply average over the initial polarizations and sum over the final polarizations. Here, λ indexes over the polarization modes, though we should keep in mind that only two are physical (the two transverse modes).

Moreover, when we perform the sum over polarizations, we sometimes need to compute sums of the form

$$\sum_\lambda \epsilon_\mu^{(\lambda)}(k)\epsilon_\nu^{(\lambda)}(k).$$

For instance, consider a matrix element corresponding to a diagram with one external outgoing photon with a momentum k,

$$M(k) = \epsilon_\mu^{(\lambda)*}(k) M^\mu, \tag{24.13}$$

where we have simply pulled out the polarization out of the matrix element M and written the rest with some index to be contracted over as M^μ. Then the physical amplitude corresponding to this process is the matrix element squared–

$$\begin{aligned} |M|^2 &\propto \sum_\lambda |\epsilon_\mu^{(\lambda)*}(k) M^\mu(k)|^2 \\ &= \sum_\lambda \epsilon_\mu^{(\lambda)*} \epsilon_\nu^{(\lambda)} M^\mu(k) M^{\nu *}(k) \end{aligned}$$

Now, it turns out that in QED amplitudes, we can simply replace this sum over polarizations with $-\eta_{\mu\nu}$. The reason for this is as follows. WLOG, let us take the photon to be traveling in the x^3 direction so that $k^\mu = (k, 0, 0, k)$ and then the transverse modes are simply

$$\begin{aligned} \epsilon^{(1)\mu} &= (0, 1, 0, 0) \\ \epsilon^{(2)\mu} &= (0, 0, 1, 0). \end{aligned}$$

Then our sum over polarizations becomes

$$\begin{aligned} \sum_\lambda |\epsilon_\mu^{(\lambda)*}(k) M^\mu(k)|^2 &= |M^1(k)|^2 + |M^2(k)|^2 \\ &= -|M^0(k)|^2 + |M^1(k)|^2 + |M^2(k)|^2 + |M^3(k)|^2 \\ &= \eta_{\mu\nu} M^\mu(k) M^{\nu *}(k), \end{aligned}$$

since the timelike and longitudinal polarizations cancel.